

AGENCY, AUTHORITY & ACCOUNTABILITY IN AUGMENTED DEMOCRACY

Anthony Ugochukwu Nwokoye

Department of Philosophy

Nnamdi Azikiwe University, Awka, Anambra State

au.nwokoye@unizik.edu.ng

ABSTRACT

Augmented democracy is a concept that proposes using technology to make political systems more inclusive and responsive. A key proposed benefit is its potential to grant a political voice to those who might otherwise be excluded from participation by acting on their behalf. In theory, this wider and more constant inclusion of perspectives could lead to better-informed and more representative collective decisions. The rise of computational agents that can summarize, represent, and act on citizens' political preferences tagged "digital twins" presses democratic political theory in new ways. This paper examines the idea of using digital twins; computer copies built from a person's data to act on their behalf in democratic processes. It focuses on three main questions: whether digital twins can truly reflect a person's will agency, whether they should have the power to make decisions authority, and who is responsible when something goes wrong accountability. Thus, the central problem addressed in this paper is the legitimacy of representation in augmented democracy; whether computational agents, designed as digital twins, can genuinely speak and act on behalf of citizens without undermining the foundational principles of autonomy, accountability, and agency. The paper employs the analytic method of philosophy to investigate competing accounts of representation and deliberation, tests their implications in the context of digital twins. The paper concludes that while digital twins could help improve democracy in some ways, they must be carefully designed to protect human agency and ensure fairness. Without the right safeguards, they risk undermining the very values they claim to support.

INTRODUCTION

Imagine a staff meeting at Nnamdi Azikiwe University, where many of the lecturers "speaking" aren't actually there. Instead, they are digital copies of the lecturers. These digital copies are computer programs built from a teacher's online profile, their past research, student reviews, and things they've said before; especially from the University CIRMS (Central Information Referencing and Management System). These programs can write short opinion papers; give advice on new courses or rules, even vote on simple, everyday issues - but only if the real lecturer gave them permission to. This can go a long way to help lecturers who are busy or on adjunct to participate in the decision making. Their digital copy can show up for them and share their opinions and expertise, making the meetings better with more viewpoints. However, there is an issue that these digital copies are like a "black box," that is, no one knows how they work, or if a private company owns the software, it could be a problem. The program might not truly represent what the teacher thinks. It could also allow a small group of tech experts to secretly control the discussion and influence the outcomes.

The problems with this idea of digital twins aren't just about the technology or how practical it is; they're about what's right and what's wrong. It raises philosophical questions on What does it mean for someone or something to speak for you? What does "good representation" actually look like? Who is responsible when something goes wrong? This is especially tricky when computer programs, which can process huge amounts of data and give instant answers, are doing the representing. The age-long discussions about what makes power legitimate, who has authority, and how we hold people accountable need to be revisited in order to understand this new reality. Thus, the aim of this seminal paper is to create clear, useful guidelines for evaluating and managing this new form of "augmented democracy," focusing specifically on digital twins.

DEMOCRATIC THEORY: REPRESENTATION AND DELIBERATION

Democratic theory is fundamentally a means of justifying power. It seeks to answer why the authority of the state is binding on free and equal individuals. The foundational principle of democracy lies in the notion of popular sovereignty; the idea that legitimate authority is derived from the people. However, the problem with popular sovereignty is on how exactly the will of the people constitutes legitimate authority. In response to this, two central ideas are very important; democratic representation and public deliberation. In the classical and contemporary periods, philosophers have tried to provide frameworks for understanding these democratic processes, which the modern technological advancements ought to implement.

Jurgen Habermas championed a deliberative turn in democratic theory and situates legitimacy not in the mere aggregation of preferences, that is, vote-centric models, rather he places emphasis on the process of communicative reason or action. According to Habermas, a norm can only claim validity if all affected by it could potentially agree to it as participants in a practical discourse.¹ This “discourse principle” is operationalized through the concept of an “ideal speech situation,” a counterfactual presupposition of communication characterized by the absence of coercive power, equal opportunities for participation, and the sincere pursuit of understanding.² In this way of thinking, a rule or decision is seen as legitimate or truly right if it was made through a fair process. This process must be open to everyone, include all relevant voices, and be based on rational discussion where the best idea wins, not the loudest voice or the most powerful person. So, for something to be considered legitimate, the method used to create it is actually more important than the final result. The method itself must be something you can defend as being truly fair and reasoned.

This idea that democracy is really about people working together to solve problems has a strong connection to the American philosopher John Dewey. For Dewey, democracy is far more than a political mechanism; it is a “form of associated living,” a community engaged in continuous collective inquiry.³ In other words, democracy wasn't just about voting. He saw it as a way of life, where a community is constantly learning and working things out together. His famous solution to democracy's problems was simply “more democracy,” which he didn't mean as more elections, but as the improved “method of intelligence” the systematic application of observation, discussion, and experimentation to social problems.⁴ Thus, the primary task of democratic institutions, is to cultivate an informed and engaged public capable of this form of intelligence. Dewey's focus shifts the theoretical concern from moments of decision to the ongoing, background processes of communication, education, and the formation of public opinion. A legitimate democracy requires institutions, that is, a free press, civic associations, schools—that nurture this public intelligence, ensuring that collective decisions are the product of informed deliberation rather than mere impulse or external manipulation.

Modern thinkers like Amy Gutmann and Dennis Thompson built on this idea to focus on how societies full of people with different, strong beliefs can still make decisions together. They argue that the essence of deliberative democracy lies in the requirement of reciprocity. Citizens and officials must justify their demands to one another by seeking reasons that can be accepted by those who are otherwise opposed.⁵ In other words, if one wants to support a law or policy, one can't just demand it. One has to explain one's reasoning to people who disagree with the policy in a way they can understand and accept. This process demands accessibility, that is, the reasons must be public, not esoteric or sectarian, and it aims for a provisional moral justification that allows for collective binding decisions despite ongoing disagreement.⁶ The reasons must be public and based on common ground, not on private or religious

beliefs that others can't share. The goal isn't to get everyone to agree, but to find a good enough reason to make a temporary decision that everyone can live with, even if they still disagree. This makes the decision legitimate. So, a law is seen as more fair and rightful if it comes from this process of people giving respectful, understandable reasons to one another, rather than just from a majority vote.

However, deliberative model of democracy has certain limitations. One of the major issues is the issue of scale, especially in a complex modern society with millions of people, which made it nearly impossible to practice the old Greek style of direct democracy, where every citizen could gather to debate and vote. Thus, representation becomes a fundamental aspect of modern democracy. As political theorist Bernard Manin argues, modern government is inherently representative; it is defined by the selection of rulers by the ruled at regular intervals, which is a different principle altogether from ancient democracy.⁷ In short, representation is the essential link that makes it possible for huge numbers of people to have a voice in their own government. It's the practical bridge between the public and the power that governs them.

Suffice to note is the fact that using representatives creates a new problem: Does it break the promise of autonomy by creating a separate group of professional politicians? A famous defense for this system comes from James Madison in the *Federalist Papers*. Madison didn't see representation as a failure of democracy. Instead, he saw it as a useful filter. His argument was that electing a small group of citizens to govern can actually improve public decision-making. He argues, "the public views, are refined and enlarged...it may well happen that the public voice, pronounced by the representatives of the people, will be more consonant to the public good than if pronounced by the people themselves."⁸ Thus,

these representatives can think about issues more deeply and from a broader, national perspective than the general public might. They can cool down heated emotions and help avoid the tyranny of passionate factions. In this view, the goal is that the final decision made by representatives will actually be wiser and more in line with the true public good than a direct vote by the people themselves would have been. So, representation isn't a loss of democracy, but a tool to make it work better on a large scale.

Another issue associated with deliberative democracy stemmed from the work of Iris Marion Young, where she pointed out that focusing only on calm, logical debate can actually leave people out. This model, she contends, privileges communication styles historically associated with educated men while ignoring the ways other marginalized groups often communicate such as personal stories, passionate speeches, protests, and everyday forms of conversation.⁹ For a democracy to be truly fair, Young argues it must welcome all these different ways of talking and sharing perspectives; not just one rigid style of debate. This criticism forces demands a rethink on what "good" participation looks like. It's not just about making rational arguments; it's about making sure everyone can actually be heard in their own way. True representation means not just having a seat at the table, but also having voices valued when one speaks.

This issue between the deliberative ideal and the practical necessity of representation finds a potential synthesis in the "systemic turn" in deliberative theory. Instead of expecting a single forum; a parliament or a citizen's assembly to fulfill all the demands of deliberation, this approach, championed by Jane Mansbridge, views the entire democratic ecosystem as a deliberative system.¹⁰ Jane Mansbridge suggests the need to stop looking for one perfect meeting where all the talking happens. Instead, one should see the whole of society; the news media, protests, courts, and congress as one big "deliberative system" where each part has a different job.¹¹

The activists and social movements are the alarm bells. They identify problems and demand attention, often through passionate protests and rallies, not calm debate. The media takes those issues, spreads them, and helps everyone talk about them. Expert agencies and courts then handle the complex, technical details of those problems. Finally, parliaments and congresses are the places where representatives have to give their final reasons for a decision and make it official. In this view, a healthy democracy needs all of these parts working together. The disruptive energy of protests and the careful reasoning of lawmakers are both essential; they just play different roles in the same system.

In this bigger picture, the whole system is considered legitimate because each part does a different job, and they all try to work together. The point isn't that every single step is perfectly rational, but that ideas and concerns can travel from the streets to the halls of power. For example, a loud protest might not be a calm debate, but it works like an alarm bell. It can force an important issue onto the agenda of politicians, who then have to debate it and give public reasons for their decisions. This way of looking at things is more practical. It accepts that not all politics will be polite or perfectly reasoned. Instead, it judges a democracy's health by one main thing: can the entire system take a public concern and turn it into a justified, actionable decision?

AUGMENTED DEMOCRACY

Over the last ten years, the idea of "augmented democracy" has appeared as a new and controversial way to think about political representation. This idea sits at the crossroads of AI, digital identities, and the theory of how democracy should work. Using AI programs often called "digital twins" to act for citizens could fix classic issues like scaling up democracy to large populations, getting more people to participate, and making governments more responsive to people's needs. On the other hand, there is fear that

this technology could actually weaken our civic freedoms, create a government run only by technical experts, and make it unclear who is responsible for decisions when they are made by complex, hidden systems.

The idea of "augmented democracy" was first widely introduced by César Hidalgo in a 2018 TED talk "A bold idea to replace politicians," where he asked whether democracy could be made more responsive if each citizen were represented by a personalized digital twin that aggregates her preferences and casts votes on her behalf.¹² This idea is compelling because it seems to solve a major problem; it could allow everyone to participate equally in democracy, no matter how busy or disadvantaged they are, making the system much fairer.

However, this idea runs into a deep philosophical problem about what true representation really means. While a digital twin might mechanically cast votes for you, which is a form of formal representation, it is questionable if it can truly capture your lived experience and your ability to discuss, change your mind, and take action in the real world. Hanna Pitkin's classic account of political representation distinguishes between formalistic, descriptive, and substantive dimensions of representation.¹³ Formalistic; having someone or something authorized to act for you. Descriptive; having a representative who looks like you or shares your background. Substantive; having a representative who truly acts in your best interests. A digital twin promises to maximize formalistic representation (a one-to-one mapping of citizens to agents), yet it remains unclear whether such agents can truly capture the lived, deliberative agency of a citizen.

As Hannah Arendt argued, politics is fundamentally about people showing up, speaking, and acting together. This raises a critical question: does being represented by an algorithm count as truly participating in political life, or does it simply replace genuine

human action with a convenient simulation?¹⁴ In other words, a digital twin is great at the first type as it can be a perfect, one-to-one mechanical vote-caster for you. But it's unclear if it can ever truly capture human lived experience or the ability of humans to think, discuss, and change human mind, which is the real human agency. Thus, Hidalgo's proposal raises critical question Does having an algorithm represent you count as truly participating in political life? Or does it just replace the real, human work of showing up and engaging with others with a convenient simulation?

Evangelos Pournaras, moved the conversation forward by proposing a specific technical system. In a 2019 paper, he suggested using blockchain technology to create a secure, unchangeable record that verifies a citizen's presence and participation in city decisions.¹⁵ This shifts the focus from Hidalgo's big-picture idea to a practical question: how do we build trust in this digital system? Blockchain prevents data from being tampered with, and his "proof of witness presence" acts like a digital stamp, making it impossible to falsely claim you participated in a debate or vote.¹⁶ From a philosophical view, Pournaras is trying to solve a problem of knowledge and authority: How can we be sure that a digital vote or opinion is real and hasn't been faked? His system successfully uses technology to guarantee that a digital interaction happened and is secure.

However, it runs into a deeper issue: Does having a cryptographically perfect record of a vote mean that genuine, meaningful political participation took place? A secure record is not the same thing as the human experience of showing up, bearing witness, and passionately advocating for a cause. Pournaras's system ensures the data is reliable, but it mistakenly equates having a political voice with having data integrity. This is where the thinker Derrida's idea is helpful. He argued that a record in an

archive, no matter how perfect and secure, is not the same thing as the actual, lived experience of being there and speaking up.¹⁷ This reduces the rich, complex reality of civic engagement to nothing more than a secure data point.

In 2020, a thinker named Will Stedden placed the idea of augmented democracy within a broader system he called "intentional capitalism."¹⁸ His idea was that using digital agents wouldn't just change politics; it could also reshape how decisions are made in business and city planning. Stedden's main contribution was philosophical. He saw a spectrum where human intention on one end is gradually mediated by algorithms on the other. This idea runs into a classic problem critiqued by philosophers like Habermas: the danger of instrumental reason¹⁹. This is the risk that everything, including our values and desires, gets treated as just another input for a machine to efficiently optimize and maximize. Politics becomes a technical problem to be solved, not a human conversation to be had. Stedden views augmentation as a tool for liberation. But Habermas would warn that unless this system is built within structures that require people to explain, discuss, and justify their decisions to one another, it won't be truly democratic. Instead of freeing humans, it could simply create a new, more efficient form of rule by technocrats and their algorithms.

In 2021, Paul Burgess made important progress in the philosophical debate by mapping out how algorithms could be used in democracy. He did this by creating four detailed scenarios, from AIs that simply help set meeting topics to systems that cast votes entirely on a citizen's behalf.²⁰ His key contribution was showing that "augmentation" isn't one single thing—it can happen at different levels, and each level comes with its own set of risks and benefits. Using clear, logical philosophy, Burgess asked a basic question: What does it truly mean for someone (or something) to legitimately represent you? He connected this to the ideas of philosopher Joseph Raz, who argued that a legitimate authority is

one that helps you follow the reasons and goals you already believe in.²¹ This becomes a useful test for a digital twin: Does it just robotically carry out your orders? Or does it actually help you make better decisions that align with your deeper values, especially when you're too busy or lack information? If it does the latter, it could be a legitimate form of assistance. In this view, the goal of augmented democracy isn't to replace people, but to act as a tool that helps them be more thoughtful and effective participants in their own governance.

In a 2022 essay, Daniela Piana agrees that using technology could make democracy more open and let more people participate. However, she warns that this might come at a steep price, that is, the loss of real civic freedom. She argues that the very systems designed to represent us could become a new, subtle form of control. Even if we agree to use them, the technology itself could be built in a way that quietly guides our choices, makes dissent harder, and makes existing social inequalities even worse.²² This idea connects to a classic distinction between two types of freedom; Positive liberty, which is the idea that the power to actually participate and control your life, and Negative liberty, which means freedom from being interfered with or controlled by others.²³ Piana suggests that while augmented democracy might give us more positive liberty (more ways to participate), it could destroy our negative liberty by making us dependent on complex, opaque systems that ultimately control us. Therefore, the biggest ethical challenge is to ensure these tools are designed to empower people, not to dominate them. Her crucial point is that simply having more ways to participate does not automatically mean we are more free.

By 2024, the conversation changed dramatically with the arrival of advanced AI like ChatGPT (large language models or LLMs). In their paper “LLMs as Agents for Augmented Democracy,” Jairo Gudiño-Rosero, Umberto Grandi, and César Hidalgo tested LLMs

as digital-twin agents, capable of simulating citizen preferences in deliberative tasks.²⁴ Their experiments showed that LLMs are very good at gathering, refining, and presenting political positions in a way that looks convincingly human. However, this raises serious philosophical questions. The first is a category mistake; a basic error in how we define something. Is it correct to treat an AI, which is essentially a sophisticated pattern-matching machine, as if it were a thinking, reasoning person with its own intentions?²⁵ It might translate to giving AI qualities it doesn't actually possess. The second major risk involves bias and injustice. These AIs are trained on existing data from the internet and books. This means they learn from a world where certain groups of people are already overrepresented and others are marginalized. If humans use these AIs to speak for citizens, they might simply repeat and amplify those existing biases, making them seem more authoritative.²⁶ The system could claim to be more inclusive while actually silencing the very voices it claims to help, effectively hiding old problems behind a shiny new technological mask.

In the same light, a review by Christopher Summerfield and his team looks at the big picture of how powerful, generative AI will affect democracy. Their conclusion is that while using AI could make democratic processes more efficient and help more people participate, the dangers are very real.²⁷ These include the risk of manipulation, power being concentrated in the hands of a few, and a serious loss of public trust. Their key philosophical point is that democracy is more than just a machine for making decisions; it is fundamentally a system that runs on trust. If citizens don't trust the system, any AI-powered "upgrade" will fail. Most importantly, this kind of trust cannot be manufactured by a machine. As philosopher Bernard Williams argued, genuine trust must be built through shared human practices of honesty, transparency, and, most crucially, holding people and institutions accountable for their actions.²⁸

By 2025, the idea of using AI to augment democracy moved from theory to real life. A report from Harvard showed how activists in countries with oppressive governments are using AI tools to organize and strengthen their pro-democracy movements.²⁹ At the same time, the Italian Parliament began experimenting with AI systems to help draft and improve new laws, testing the concept in an actual government setting.³⁰ These real-world examples show both the great potential and the serious danger of this technology. On one hand, it can be a powerful tool to help silenced communities fight for their rights. On the other hand, in stable democracies, it could turn the deep, human process of debate into a cold, mechanical bureaucracy. This all comes down to a timeless philosophical question: Is technology just a neutral tool, or does it come with its own built-in values? The philosopher Heidegger suggested that technology is never just a passive instrument; it's actually a way of seeing and shaping the world.³¹ In this light, augmented democracy isn't a neutral fix. It promotes a specific vision of politics, one where citizens are reduced to data points, and the legitimacy of our decisions is handled by computational proxies instead of people.

DIGITAL TWIN

A digital twin is a computational replica of a person, system, or object that mirrors its key features, behaviors, and preferences in a digital environment. It's built to act like the real thing in a digital space. In an augmented democracy, this means creating a digital copy of a citizen. This copy learns about you from your data like your articles, how you've voted before, or your stated opinions. Its main job is to then act as your stand-in or proxy in political discussions or decision-making. It's more than just a static profile; it's a model that learns and tries to predict how you would react to new policies or debates. It is not merely a database entry but a dynamic, learning model that can simulate how the person might respond to policies, debates, or decisions. The promise is that it can

give people a continuous political presence without requiring constant participation; the risk is that it may misrepresent or be manipulated, raising ethical questions about agency, authority, and authenticity.

A digital twin for a person is a computer program built to act as their stand-in. For it to truly be their digital twin, three things must be true:

1. It is *built from that person's own data* (like their answers to surveys, their social media posts, or records of their past behavior).
2. It is *designed to act or "speak" for them*. Its job is to create statements, suggestions, or votes that aim to reflect what the person would likely think or decide.
3. It is *used in civic or political situations* with the goal that other people (or institutions) will accept its actions as if they came directly from the person it represents.

Firstly, a digital twin isn't a general AI. It's a unique model built to copy one specific person. It learns from that individual's own information like their answers to questions, things they've written online, how they've voted before, or how they usually make decisions. The goal is to make the twin reflect that person's core beliefs, how they think, and what they prefer. The twin uses a person's personal data to act like a digital version of the person, so it's not just a random program.

Secondly, this isn't just a file that stores facts. A digital twin is built to do things. It creates outputs like statements, policy choices, votes, or suggestions that are meant to act as your stand-in. The big question here is about true representation: Is the twin really doing what you want, or are its choices being skewed by its programmers or its own design? As rightly noted by John Searle that just because a machine can combine symbols doesn't mean it understands what

they mean. So, is the twin truly speaking for you, or is it just a clever fake?

Thirdly, the digital twin only has power if other people and institutions like governments, councils, or other citizens accept its actions as if they came directly from the person it represents. For example, it could vote for you in a university meeting, send an opinion to the local government, or take part in a discussion forum. This brings up issues of power and permission: Do these groups officially accept the twin's actions? A political scholar named Hanna Pitkin stressed that representation only works if the audience agrees to it. If no one accepts the digital twin's role, it's irrelevant. But if they do, it becomes a powerful tool that lets a person be "present" in civic life without being there physically.

The idea of a "digital twin" which is a computer copy of a person that can act for them in politics seems like a handy tool for modern democracy. But when assessed through the eyes of philosophy, might break the very ideas that make democracy work. Rousseau believed that for a government to be legitimate, every single citizen must be directly involved in creating the general will—the community's common good, not just what individuals want.³² He hated the idea of representatives because he thought they would corrupt the people's true will. A digital twin is the worst kind of representative. It acts based on your past selfish interests, locking them in forever. It can't do the most important thing: let you change your mind, talk with your neighbors, and together figure out what's best for everyone. For Rousseau, democracy is an active process, and a digital twin turns citizens into passive data points.³³

On the other hand, Habermas said that laws are only legitimate if they come from free, open, and rational public discussion where the best argument wins.³⁴ This process, called communicative action, is what makes a decision right.³⁵ A digital twin can't have a real

conversation. It can't listen to a new argument, be persuaded, or explain its reasoning. It can only spit out a pre-programmed response. This isn't communication; it's just a calculation. If the tech behind these twins is secret or owned by a company, it lets hidden power and control sneak into public debates, which destroys trust.³⁶ Also, Arendt saw politics as the space for action; the amazing human ability to start something new together in public through speech and deeds.³⁷ This ability to begin anew is called natality.³⁸ A digital twin replaces unpredictable human action with predictable, calculated behavior. It kills the chance for a surprise, a new idea, forgiveness, or a promise. Your unique story and voice get replaced by a computer profile. For Arendt, this destroys the freedom and excitement that is the whole point of political life.³⁹

AGENCY, AUTHORITY, AND ACCOUNTABILITY IN THE AGE OF THE DIGITAL TWIN

The idea of an "augmented democracy," using digital twins, offers an appealing vision of a political system that includes more people, works more efficiently, and responds more quickly to citizens' needs.⁴⁰ It suggests a future where no one is left out of decision-making because they are too busy, lack resources, or cannot physically participate. In this future, public choices could be informed by ongoing input from the entire population. However, this research carefully shows that such a technological solution requires us to rethink some of the most basic principles of democratic theory: agency, authority, and accountability. The key question of legitimacy; whether computer programs can honestly represent people without weakening individual freedom and moral responsibility is an ethical issue.

The issue of agency is the first and most basic challenge. In a democracy, agency means the ability to act in politics in a conscious, intentional, and genuine way. It is not only about casting a vote but also about forming opinions through experience, discussion, and ethical thinking. The paper addresses this problem

by pointing out an important difference between formal agency and real agency, building on ideas from political thinker Hanna Pitkin. A digital twin is very good at formal agency. It is a powerful tool for processing data, created from a person's "articles, past votes, or stated views."⁴¹ It can use algorithms to predict a person's likely choice based on past behavior, creating a seemingly accurate copy of a decision. In this way, it offers a kind of agency—a substitute that takes action.

Additionally, this paper draws from Jürgen Habermas's concept of communicative action, where legitimacy comes from a process of "free, open, and rational public discussion where the best argument wins."⁴² Agency, from this perspective, is used through exchanging reasons. A digital twin "can't have a real conversation. It can't listen to a new argument, be persuaded, or explain its reasoning." Its result is a calculation, not a justified communication. So, while the twin may "act", it does not have "agency" in the deeper, philosophical sense needed for democratic legitimacy. It risks creating a system where citizens, comforted by the ease of representation, become inactive data points, their real agency fading as their digital substitutes perform an empty, algorithmic imitation of politics.

If a digital twin's claim to agency is weak, then the question of whether it should have authority becomes even more urgent. Authority in a democracy is the rightful power to make decisions that others must follow. The paper explores this by looking at two competing reasons for authority: one based on efficiency and inclusion, and another based on moral reasoning and forming a shared public will. The argument *for* giving authority to digital twins is mainly practical. As explained in the paper, the twin can address practical issues of size and participation. In a large, modern society of millions, direct democracy is not possible. Representation has always been the necessary solution, as noted by

Bernard Manin. Digital twins are presented as the logical, technological next step of this idea—an extremely efficient, personalized system of representation.⁴³ Will Stedden see this as a tool for freedom, a way to smoothly integrate citizen intention into the workings of government and even the economy.⁴⁴ The authority of the twin, from this view, comes from the clear permission of the user (its principal) and its ability to produce optimal results based on that user's recorded preferences.

A closer examination of this practical argument from a philosophical perspective, shows that authority in a democracy cannot come only from efficiency or consent; it must also be based on a process that ensures decisions are fair and reasoned. This is where the models of Madison, Habermas, and Young become important. James Madison defended representative authority as a “filter” that would “improve and broaden the public's views.”⁴⁵ Representatives were meant to use their judgment, not just pass on preferences. A digital twin, working on pre-set algorithms, does not have this capacity for judgment. It cannot think about the common good; it can only act on individual interest. More importantly, there is the danger of “instrumental reason,” an idea criticized by Habermas. This means reducing all political questions to technical problems to be efficiently solved by experts and their algorithms. If digital twins are given decision-making authority, politics stops being a “human conversation to be had” and becomes a cold, computational process. This risks creating a new, hidden form of rule by experts where authority is given not to public reason but to the secret code of the algorithms. The paper raises the concern that a small group of tech experts could secretly control the discussion and influence the outcomes. Using the postulation of Iris Marion Young that authoritative decisions must come from a variety of voices and ways of communicating, and not just the calm, logical debate that an AI might copy, but also through “personal stories, passionate speeches, and protests.”⁴⁶ It follows that a digital twin's “authority” would be sterile, unable to process or make legitimate

the full range of human political expression. Therefore, giving full decision-making authority to digital twins would fundamentally damage the thoughtful and representative foundations of democratic authority, replacing public justification with private calculation.

The final part of democratic legitimacy is accountability. This is the system that checks authority and provides a way to correct things when power is misused. The paper identifies accountability as the most complicated and dangerous issue within augmented democracy, describing a possible “accountability vacuum” caused by the digital twin’s unclear nature. In a traditional representative model, accountability is, at least in theory, clear. The elected official is responsible for their decisions and can be held responsible by the voters through elections, public examination, and legal processes. The line of responsibility is relatively simple. However, using digital twins, there is a complex network of actors, thereby, hiding who is finally responsible for a decision or action. The network of actors includes but not limited to the citizen, the algorithm, the software company, the government. Thus, in situation of breach, who should be held accountable?

1. The Citizen: Did they carelessly set up their twin or fail to monitor it?
2. The Algorithm: Can the AI itself be held responsible?
3. The Software Company: Are the private companies that create the algorithms, train the models, and possibly include their own biases responsible?
4. The Government: Is the institution that adopts and approves the use of these systems ultimately responsible for what happens?

This lack of clarity leads to a serious loss of trust. Thus, as rightly captured by Bernard Williams, that “democracy is more than just a

machine for making decisions; it is fundamentally a system that runs on trust.”⁴⁷ This trust is built on practices of “honesty, transparency, and, most importantly, holding people and institutions accountable.” When no single actor can be clearly held responsible for a decision made by a “black box” algorithm, public trust in the whole democratic system falls apart. If a citizen’s digital twin votes for a policy that hurts them, who can they hold responsible? Who can they vote out of office? Thus, without clear answers, augmented democracy becomes an unaccountable system, which is the opposite of a legitimate democracy.

CONCLUSION

The paper does not just list these problems; it suggests a solution by redefining the role of the digital twin. The response to the three-part problem of agency, authority, and accountability is not to reject the technology, but to strictly limit its purpose and design. The paper suggests that the only way to keep democracy legitimate while using digital twins is to strictly limit their power. They should act as helpful tools, not as replacements for people. Their job should be, first, to support human decision-making by doing things like summarizing complex news, managing calendars, and providing personalized updates. This helps citizens participate more effectively when they want to, strengthening their own ability to act rather than replacing it. Second, their role should be to inform, not to decide. They could handle small, agreed-upon tasks, like voting on a minor schedule change. But for important decisions that involve values or morals, the final choice must always be made by a real, thinking person. Finally, to solve the problem of accountability, the technology must be completely transparent. This means the computer code that runs the digital twins should be open for everyone to see and check. This transparency allows people to examine how a decision was made, which helps figure out who is responsible if something goes wrong; whether it was due to a mistake in the code, carelessness by the company that made it,

or biased information. In short, digital twins should be designed to serve and empower citizens, never to control or replace them.

In conclusion, the paper responds to the deep challenges of agency, authority, and accountability by stating that the legitimacy of augmented democracy depends on one central idea: human primacy. Digital twins can be legitimate only as tools that serve and increase human will and public reasoning. They become illegitimate the moment they try to replace them. The goal is not to create a smooth, automated democracy, but to use technology to support a stronger, more inclusive, and more thoughtful human political conversation. The digital twin should be a bridge back to real engagement, not a one-way path to a passive political future.

Endnotes

1. Jürgen Habermas, *Moral Consciousness and Communicative Action*, trans. Christian Lenhardt and Shierry Weber Nicholsen (Cambridge, MA: MIT Press, 1990), 66.
2. Jürgen Habermas, "Discourse Ethics: Notes on a Program of Philosophical Justification," in *Moral Consciousness and Communicative Action* (Cambridge, MA: MIT Press, 1990), 86
3. John Dewey, *Democracy and Education* (New York: Macmillan, 1916), 101.
4. John Dewey, *The Public and Its Problems* (New York: Henry Holt and Company, 1927), 146.
5. Amy Gutmann and Dennis Thompson, *Why Deliberative Democracy?* (Princeton: Princeton University Press, 2004), 3.
6. Amy Gutmann and Dennis Thompson, *Why Deliberative Democracy?*, 7.
7. Bernard Manin, *The Principles of Representative Government* (Cambridge: Cambridge University Press, 1997), 4
8. James Madison, "The Federalist No. 10," in *The Federalist Papers*, ed. Clinton Rossiter (New York: Signet Classics, 2003), 76.
9. Iris Marion Young, "Communication and the Other: Beyond Deliberative Democracy," in *Democracy and Difference: Contesting the Boundaries of the Political*, ed. Seyla Benhabib (Princeton: Princeton University Press, 1996), 124.
10. Jane Mansbridge et al., "A Systemic Approach to Deliberative Democracy," in *Deliberative Systems: Deliberative Democracy at the Large Scale*, ed. John Parkinson and Jane Mansbridge (Cambridge: Cambridge University Press, 2012), 4.
11. Jürgen Habermas, *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy*, trans. William Rehg (Cambridge, MA: MIT Press, 1996), 308
12. César Hidalgo, "A Bold Idea to Replace Politicians," *TED Talk*, April 2018.
13. Hanna Pitkin, *The Concept of Representation* (Berkeley: University of California Press, 1967), 38
14. Hannah Arendt, *The Human Condition* (Chicago: University of Chicago Press, 1958), 175.
15. Evangelos Pournaras, "Proof of Witness Presence: Blockchain Consensus for Augmented Democracy in Smart Cities," *arXiv*, June 30, 2019, 5.
16. Evangelos Pournaras, "Proof of Witness Presence," 7.

17. Jacques Derrida, *Archive Fever* (Chicago: University of Chicago Press, 1996), 17.
18. Will Stedden, “Augmented Democracy — Intentional Capitalism,” December 2020, <https://opus.stedden.org>.
19. Jürgen Habermas, *Between Facts and Norms* (Cambridge, MA: MIT Press, 1996), 287.
20. Paul Burgess, “Algorithmic Augmentation of Democracy,” *AI & Society* (2021): 4–6.
21. Joseph Raz, *The Morality of Freedom* (Oxford: Oxford University Press, 1986), 53–57.
22. Daniela Piana, “Augmented Democracy: More Open, Less Free?” *Knowledge Cultures* 10, no. 2 (2022): 94–96.
23. Isaiah Berlin, “Two Concepts of Liberty,” in *Four Essays on Liberty* (Oxford: Oxford University Press, 1969), 131–32.
24. Jairo Gudiño-Rosero, Umberto Grandi, and César Hidalgo, “Large Language Models as Agents for Augmented Democracy,” *arXiv*, May 6, 2024, 8.
25. Gilbert Ryle, *The Concept of Mind* (London: Hutchinson, 1949), 16.
26. Miranda Fricker, *Epistemic Injustice* (Oxford: Oxford University Press, 2007), 12.
27. Christopher Summerfield et al., “How Will Advanced AI Systems Impact Democracy?” *arXiv*, August 27, 2024, 15.
28. Bernard Williams, *Truth and Truthfulness: An Essay in Genealogy* (Princeton: Princeton University Press, 2002), 36.
29. Ash Center for Democratic Governance and Innovation, “How AI Can Support Democracy Movements” (Harvard University, February 2025), 3.
30. “AI Systems for Legislative Innovation in the Italian Parliament,” *SSRN*, September 2025, 4.
31. Martin Heidegger, “The Question Concerning Technology,” in *The Question Concerning Technology and Other Essays* (New York: Harper, 1977), 23.
32. Jean-Jacques Rousseau, *On the Social Contract* (Indianapolis: Hackett Publishing Company, 1987), 25.
33. Jean-Jacques Rousseau, *On the Social Contract*, 29, 55.
34. Jürgen Habermas, *Between Facts and Norms* (Cambridge, MA: MIT Press, 1996), 110.
35. Jürgen Habermas, *The Theory of Communicative Action, Volume 1* (Boston: Beacon Press, 1984), 94.
36. Jürgen Habermas, *The Theory of Communicative Action, Volume 2* (Boston: Beacon Press, 1987), 320.

37. Hannah Arendt, *The Human Condition*, 2nd ed. (Chicago: University of Chicago Press, 1998), 7
38. Hannah Arendt, *The Human Condition*, 177.
39. Hannah Arendt, "What Is Freedom?," in *Between Past and Future* (New York: Penguin Books, 2006), 163.
40. César A. Hidalgo, "A Bold Idea to Replace Politicians," filmed April 2018 in Vancouver, Canada, *TED video*, 13:01, https://www.ted.com/talks/cesar_hidalgo_a_bold_idea_to_replace_politicians.
41. Hanna Fenichel Pitkin, *The Concept of Representation* (Berkeley: University of California Press, 1967), 209.
42. Jürgen Habermas, *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy*, trans. William Rehg (Cambridge, MA: MIT Press, 1996), 110.
43. Bernard Manin, *The Principles of Representative Government*, 6.
44. Will Stedden, "Augmented Democracy and Intentional Capitalism"
45. James Madison, "The Federalist No. 10," in *The Federalist Papers*, ed. Clinton Rossiter 76.
46. Iris Marion Young, "Communication and the Other: Beyond Deliberative Democracy," 124.
47. Bernard Manin, *The Principles of Representative Government* 6.