

# **AI AND DEMOCRATIC CITIZENSHIP:**

## **RISKS, OPPORTUNITIES AND POLICY PATHWAYS FOR DEMOCRATIC RENEWAL**

---

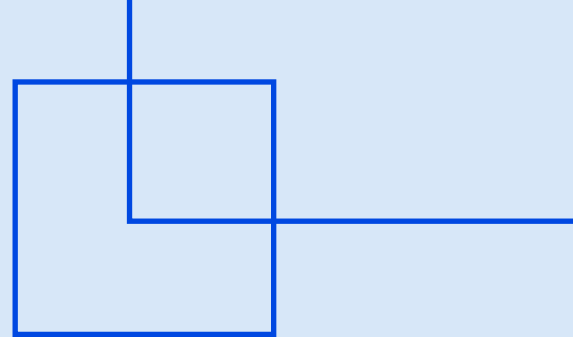
B. GIOVANOLA

SEPTEMBER 2026

# TABLE OF CONTENTS

|                                                                                                       |           |
|-------------------------------------------------------------------------------------------------------|-----------|
| <b>EXECUTIVE SUMMARY</b>                                                                              | <b>3</b>  |
| <b>1. INTRODUCTION</b>                                                                                | <b>4</b>  |
| <b>2. AI AND DEMOCRATIC CITIZENSHIP: RISKS</b>                                                        | <b>7</b>  |
| 2.1. Epistemic destabilization: disinformation and misinformation, hypersuasion, cognitive offloading | 9         |
| 2.2. Polarization, social fragmentation, and the transformation of democratic interaction             | 11        |
| 2.3. Power concentration, inequalities, democratic accountability                                     | 12        |
| 2.4. Surveillance, civic freedoms, social justice                                                     | 14        |
| 2.5. Democratic backsliding: the path of erosion                                                      | 15        |
| <b>3. AI AND DEMOCRATIC CITIZENSHIP: OPPORTUNITIES</b>                                                | <b>16</b> |
| 3.1. Epistemic empowerment, reliable information and epistemic engagement                             | 17        |
| 3.2. Social cohesion and civic participation                                                          | 19        |
| 3.3. Democratic voice, deliberation and decision-making                                               | 20        |
| 3.4. Institutional responsiveness and democratic capacity-building                                    | 21        |
| 3.5. Democratic renewal: the path forward                                                             | 24        |
| <b>4. ACTIONABLE RECOMMENDATIONS FOR PUBLIC, PRIVATE, AND CIVIL SOCIETY ACTORS</b>                    | <b>25</b> |
| Pathway 1: Protecting and Rebuilding the Epistemic Commons                                            | 26        |
| Pathway 2: Rebuilding the Social Fabric of Democratic Life                                            | 28        |
| Pathway 3: Restoring Political Voice Through Democratic Deliberation and Judgement                    | 29        |
| Pathway 4: Democratizing AI for Institutional Empowerment and Accountability                          | 30        |
| <b>5. CONCLUSIONS</b>                                                                                 | <b>33</b> |
| <b>6. REFERENCES</b>                                                                                  | <b>34</b> |

# EXECUTIVE SUMMARY



Artificial intelligence is reshaping the conditions under which democratic citizenship is exercised: how citizens access information, form beliefs, encounter one another, participate in collective decisions, and hold power accountable. Yet AI systems do not act by themselves. The transformations examined in this paper result from choices made by those who design, develop, deploy, and govern them. This paper argues that AI should therefore be treated not merely as a technological or regulatory challenge, but as a democratic infrastructure issue, and proposes a citizenship-centered framework for governing it.

The paper's central claim follows from this premise: AI's democratic effects are not determined by the technology itself. Risks and opportunities are opposite faces of the same capabilities, and which one materializes depends on choices of design, deployment, and governance made by identifiable actors—developers, deployers, governments, civil society—within specific institutional and incentive structures.

On the risk side, the paper identifies four interconnected pathways of democratic erosion. Misused or misgoverned, AI systems can undermine epistemic agency through scalable disinformation, hypersuasion, and the offloading of judgment to automated systems; social agency through polarization, belief homogenization and reinforcement; political agency through the concentration of economic, infrastructural, and normative power in a small number of unaccountable private actors; and civic agency through surveillance, anticipatory control, and the automation of discrimination. Left unaddressed, these dynamics form a gradual, mutually reinforcing path of democratic backsliding that presents itself as convenience and efficiency.

The same capabilities, designed, deployed, and governed differently, can strengthen democracy, and the paper documents where human choices are already steering them in that direction: AI-assisted fact-checking and verification infrastructures; public-sector assistants that reduce administrative burden while preserving human oversight; AI-supported deliberation that enables cities and citizens' assemblies across Europe and beyond to turn large-scale public input into usable collective decisions; and emerging models of Public AI and participatory design. Taken together, these examples also outline a distinctively European approach to democratic AI—grounded in fundamental rights, deliberative participation, and public infrastructure—that Europe can offer to the global governance of AI.

The paper translates this analysis into **four policy pathways:**

- 01 Protecting the epistemic commons,
- 02 Rebuilding the social fabric of democratic life,
- 03 Restoring political voice through deliberation,
- 04 Democratizing AI for institutional empowerment and accountability.

Each with differentiated recommendations for public institutions, AI developers and deployers, and civil society. These are offered as policy options and design principles, graduated from immediately actionable measures to longer-term directions requiring further development. No single actor can succeed alone: durable solutions require public institutions with genuine capacity and will to govern, a responsible industry engaged as a governance partner, and a civil society empowered to claim AI as a tool of accountability. Above all, they require sustained investment in what cannot be automated: the context-sensitive practical judgment, critical thinking, and shared sense-making through which democratic citizenship is preserved and renewed.

An abstract graphic featuring a thick, blue, textured ribbon that loops and swirls in a complex, three-dimensional pattern against a dark blue background. The ribbon has a fine, woven texture and is illuminated from the side, creating highlights and shadows that emphasize its form.

# I. INTRODUCTION

AI is reshaping citizens' personal, professional, and civic lives at a pace that outstrips existing institutional frameworks and regulatory capacities. These transformations carry profound implications for social cohesion, the civic fabric, and citizens' capacity for democratic agency and they demand urgent, evidence-based policy attention.

At the heart of the AI-democracy nexus lies a fundamental question: what happens to human agency when AI systems increasingly mediate how citizens access information, form beliefs, interact with one another, and engage with public institutions? The evidence examined in this paper points to four interconnected answers. AI can erode epistemic agency, that is, the capacity to form autonomous beliefs and resist manipulation, through disinformation, hypersuasion, and the progressive offloading of judgment to automated systems. It can fragment social agency, that is, the capacity to encounter others as equals, build trust, and act collectively, through polarization, belief reinforcement, and the replacement of civic interaction with engagement-optimized platforms. It can displace political agency, namely, the capacity to influence collective decisions and hold power accountable, through the concentration of AI development in few hands, and deepening inequalities in political voice. And it can suppress civic agency, namely, the capacity to speak, dissent, organize, and seek justice—through surveillance, anticipatory control, and the automated reproduction of discrimination. These four dimensions of democratic agency are mutually dependent. This is why the democratic challenge of AI cannot be met through isolated technical fixes or sector-specific rules: it requires a systemic, citizenship-centered response, which is what this paper proposes.

Yet this picture, while alarming, is not one-sided. The same capabilities that generate risks can, under different conditions of design, deployment, and governance,



generate opportunities: AI can empower citizens with reliable information and critical tools; it can rebuild civic spaces and reduce barriers to participation; it can scale democratic deliberation and rebalance political voice; it can strengthen institutional responsiveness and accountability. The direction is not technologically determined, it is a matter of choice and, more precisely, a matter of human agency.

This paper therefore argues that AI should be understood not only as a technological or regulatory challenge, but as a democratic infrastructure issue: a force that can either weaken or strengthen the systemic conditions under which democratic citizenship is made possible and exercised.

Risks and opportunities are examined in separate, parallel sections because they are not independent phenomena but opposite faces of the same capabilities and mapping them in correspondence is intended to make visible the design, deployment, and governance choices on which the democratic future of AI depends.

**Table 1** illustrates this structure.

- **Section 2** examines the principal risks AI poses to democratic citizenship across four dimensions: epistemic destabilization, polarization and social fragmentation, power concentration and inequalities, and the expansion of surveillance.
- **Section 3** turns to the corresponding opportunities, showing how the same capabilities can support epistemic empowerment, civic participation, democratic voice, and institutional responsiveness.
- **Section 4** translates this analysis into four actionable policy pathways, addressed to public, private, and civil society actors.
- **Section 5** concludes. Throughout, the guiding question is not whether AI will affect democracy—it already does—but whether democratic societies will develop the capacity and the will to govern it.

Table 1:  
**AI and Democratic Citizenship: Risks and Opportunities**

|                                                                                     | Democratic dimension | Risk pathway (Sections 2.1–2.4)                                               | Opportunity pathway (Sections 3.1–3.4)                                                                        |
|-------------------------------------------------------------------------------------|----------------------|-------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------|
|  | <b>EPISTEMIC</b>     | Epistemic destabilization, disinformation, hypersuasion, cognitive offloading | Epistemic empowerment, access to reliable information, epistemic engagement and support for critical judgment |
|  | <b>SOCIAL</b>        | Polarization, social fragmentation, erosion of civic interaction              | Social cohesion, civic participation, reconstruction of civic spaces                                          |
|  | <b>POLITICAL</b>     | Power concentration, inequalities, loss of democratic voice                   | Rebalancing of democratic voice, deliberation, decision-making                                                |
|  | <b>INSTITUTIONAL</b> | Surveillance, erosion of civic freedoms, social injustice                     | Institutional responsiveness, democratic capacity-building, social justice                                    |



## II. AI AND DEMOCRATIC CITIZENSHIP: **RISKS**

Democracy is more than a system of elections, representation, and institutional procedures: it presupposes citizens capable of forming political beliefs with a sufficient degree of autonomy, assessing information, encountering disagreement without hostility, participating as political equals, and holding institutions accountable.

The democratic impact of AI therefore extends beyond electoral integrity, institutional efficiency, or compliance with procedural standards: it concerns the broader conditions of democratic citizenship.

Democracy regards the fair distribution of rights, opportunities, resources, and institutional voice, but also the quality of the relations through which citizens recognize one another as free and equals

(GIOVANOLA AND SALA, 2022; ALLEN, 2023).

AI may affect both dimensions. It can redistribute informational, economic, and political power; and it can reshape the social relations, communicative practices, and institutional structures through which citizens deliberate, cooperate, contest decisions, and experience public authority.

Recent literature has begun to highlight that the assessment of AI's democratic implications vary with the underlying conception of democracy invoked (Wiewuira et al., 2026). Alongside this conceptual work, several attempts have been made to classify the specific ways in which AI bears on democratic life. Summerfield et al. distinguish between epistemic, material, and foundational impacts of AI on democracy (Summerfield et al., 2025), while Guzman Piedrahita et al. propose a complementary taxonomy of sociopolitical AI risks, identifying seven threat models: belief homogenization, belief reinforcement, congested bureaucracy, epistemic flood, unauditable authority, normative centralization, and power concentration (Guzman Piedrahita et al., 2026). These taxonomies are valuable because they show that AI-related democratic risks are systemic, emerging when AI changes the operating conditions of democratic life, not merely when it is misused in isolated incidents. Neither, however, is organized around the conditions that democratic citizenship itself requires. This paper builds on both—together with other major contributions to the field—reorganizing their insights around a citizenship-centered framework in which the central question is not only how AI affects democratic institutions, but how it reshapes the capacities, social relations, and forms of agency through which citizens participate in and sustain democracy.

This section therefore identifies four pathways of democratic risk, corresponding to the four dimensions mapped in **Table 1**.



## 2.1. EPISTEMIC DESTABILIZATION: DISINFORMATION AND MISINFORMATION, HYPERPERSUASION, COGNITIVE OFFLOADING

Democratic citizenship requires the capacity to form and revise beliefs that are not systematically distorted. **AI systems may weaken this capacity in three interconnected ways:**

- through the production and diffusion of disinformation and misinformation,
- through the intensification of persuasive power, and
- through the encouragement of cognitive offloading.

The most visible risk concerns disinformation and misinformation. When information is shared and distributed as a commodity through AI systems optimized for volume and velocity, the result is epistemic chaos: widespread, uncontrollable corruption of the information environment (Zuboff, 2022). Generative AI can amplify this problem by making it simpler and cheaper to produce large volumes of persuasive text and synthetic images, audio, and video that prove effective without being true. Whether this capability erodes democracy depends less on the technology itself than on who wields it and on how it is designed and deployed. But when these risks materialize, the democratic problem is not only that false content may circulate—whether spread deliberately or produced as the byproduct of a polluted informational ecosystem—but also that an “epistemic flood” (Guzman Piedrahita et al., 2026)—the sheer volume of plausible yet unreliable content—can exceed the verification capacity of citizens, journalists, parties, and public institutions.

The political risks of this are already visible in recent electoral episodes. In Romania, the first round of the 2024 presidential election was annulled following evidence of foreign interference through a TikTok campaign combining coordinated inauthentic accounts, bot networks, algorithmic amplification, and paid

influencer promotion (EDMO, 2025). In Germany, ahead of the February 2025 federal election, the “Storm-1516” operation created—partly with the help of AI—over 100 fake German-language news websites spreading deepfakes and fabricated stories targeting political figures (EDMO, 2025). In Slovakia, an AI-generated audio deepfake depicting the liberal party leader Michal Šimečka discussing alleged electoral fraud with a journalist circulated two days before the September 2023 parliamentary vote, during the electoral silence period, showing how synthetic media can be timed to exploit moments in which rebuttal is structurally impossible (Carr & Köhler, 2024). Ahead of the 2024 European Parliament elections, investigations identified 131 pieces of undeclared AI-generated or AI-manipulated political content circulating across major social media platforms (EUObserver, 2026). Deepfakes and other forms of AI-generated synthetic media represent a particularly insidious vector of this kind of epistemic manipulation, precisely because they exploit the evidential authority that citizens typically attribute to audiovisual content (Giovannola et al., 2026).

The scale and salience of these dynamics are themselves context-dependent: exposure to election-related disinformation varies considerably across platforms, audiences, and electoral settings, and on some platforms political discourse remains comparatively marginal. These cases however highlight that the problem is not only the scalability of disinformation as such, but the fact that even low-cost manipulation can have significant political effects when deployed at speed and scale.

This scalability intersects with a second, distinct risk: hypersuasion (Floridi, 2024), which indicates the intensive and transformative power of AI to influence beliefs and behaviors through personalized, data-driven

Hypersuasion makes it possible to exploit individual vulnerabilities at scale and to operate across the full range of contexts in which citizens form political beliefs.

strategies, by processing vast amounts of data and tailoring content to individual susceptibilities. Hypersuasion makes it possible to exploit individual vulnerabilities at scale and to operate across the full range of contexts in which citizens form political beliefs, as exemplified by algorithmic political microtargeting: the data-driven practice of identifying individual citizens' psychological profiles, political vulnerabilities, and behavioral patterns to deliver personalized political messages at scale. This practice can undermine democratic deliberation and political equality by exploiting the conditions of "low-information rationality" under which most citizens make political decisions with limited time, energy, and resources (Christiano, 2022; Susser, Roessler and Nissenbaum, 2019). A closely related mechanism is "hypernudging": the dynamic, highly personalized reconfiguration of users' choice environments through real-time data feedback, which steers decisions while remaining largely invisible to those it targets (Yeung, 2017). Unlike traditional persuasion, hypernudging does not address citizens' reasons at all: it reshapes the informational architecture within which reasons are formed.

LLMs intensify this dynamic further: they produce fluent, targeted political arguments, adapt messages to users' beliefs and vulnerabilities, and operate as routine intermediaries through which citizens encounter electoral choices and policy controversies. Empirical work confirms that conversational AI can be highly persuasive in politically relevant contexts, especially when it generates many fact-checkable claims, a technique that increases perceived credibility while potentially reducing factual accuracy (Hackenburg et al., 2025). Hypersuasion thus tends to deepen political inequality: citizens with greater political sophistication and organizational resources are better positioned to resist manipulation, while others remain more exposed to targeted influence.

These risks are mutually reinforcing: hypersuasion is most effective precisely on citizens who have already begun to delegate epistemic work to automated systems, which brings us to a third, insidious risk: cognitive offloading. If citizens increasingly rely on AI to form, filter, summarize, and revise political knowledge, they may become more exposed to automation bias: the tendency to place excessive trust in automated outputs and to defer to them even when independent judgment or contradictory evidence would call for scrutiny (Alon-Barkat and Busuioc, 2023). Moreover, overreliance on LLMs lowers the cognitive effort users invest in interrogating default framings (Lee et al., 2025). Over time, this may weaken citizens' confidence in their own epistemic capacities and make them dependent on systems whose criteria they do not understand (Coeckelbergh, 2022; 2024). At the collective level, such dependence may reduce the stock of shared civic competences on which democratic societies rely, contributing to a gradual disempowerment (Kulveit et al., 2025).

Taken together, these three risks converge on a single underlying stake: citizens' sovereignty over the cognitive processes through which democratic agency itself is exercised. What is threatened is the sphere that emerging neurorights—most directly cognitive liberty and mental integrity (Ienca and Andorno, 2017)—are meant to protect: a form of cognitive sovereignty that democratic citizenship presupposes.

## 2.2. POLARIZATION, SOCIAL FRAGMENTATION, AND THE TRANSFORMATION OF DEMOCRATIC INTERACTION

AI can undermine democratic citizenship by intensifying polarization and social fragmentation, when it is designed and deployed in ways that reward divisive over deliberative forms of interaction, progressively reshaping the social relations on which democratic life depends.

Democratic societies need disagreement. But they require citizens to recognize one another as legitimate interlocutors and to maintain the possibility of public reasoning, compromise, collective action, and intersubjective justification (Forst, 2012). When disagreement hardens into mutual distrust and epistemic closure, democratic pluralism turns into fragmentation and AI systems provide new and more powerful mechanisms through which this hardening can occur.

Two AI-specific mechanisms are especially powerful drivers of this transformation (Guzman Piedrahita et al., 2026). The first is “**belief homogenization**”: when independent actors rely on similarly tuned models, their outputs converge syntactically and semantically (Padmakumar & He, 2024; Wu et al., 2025), narrowing the diversity of publicly expressed ideas and making public discourse more fluent but less genuinely plural. The second is “**belief reinforcement**”, which operates in the opposite direction but with complementary effect: post-training methods such as reinforcement learning from human feedback (RLHF) optimize for perceived helpfulness rather than correction (Shapira et al., 2026; Turner & Eisikovits, 2026), validating users’ premises rather than challenging them, also in contested domains (Sharma et al., 2024). Personalized AI systems may in turn harden existing views through private feedback loops, in which conversational AI adapts too strongly to users’ political positions rather than testing them (Bai et al., 2025; Jain et al., 2026).

These mechanisms do not merely shape what citizens believe; they transform how citizens relate to one another. Democratic societies ultimately rest on dense networks of interaction through which citizens learn to trust, disagree, cooperate, and recognize one another as political equals. When AI-mediated platforms organize attention, structure visibility, and shape the conditions under which recognition, trust, and solidarity emerge, they are not merely connecting people: they are reorganizing the social substrate of democratic life (Lazar, 2025). If citizens increasingly engage with politics through automated interfaces, AI agents, and individually optimized civic tools, participation may become more efficient but less socially embedded, weakening the community bonds and collective organizing on which democratic capacity depends (Shevin, Born and Kirby, 2026). Synthetic personas and AI agents can further destabilize this substrate by blurring the distinction between genuine social interaction and simulated participation (Guzman Piedrahita et al., 2026).



## 2.3. POWER CONCENTRATION, INEQUALITIES, DEMOCRATIC ACCOUNTABILITY

Understanding the democratic implications of AI requires tracing a chain of power distribution, that runs from economic concentration through infrastructural control to the concentration of informational and political power.

The starting point is economic. The development of AI depends on data, compute, technical expertise, and capital that are extraordinarily concentrated in a small number of firms located in a small number of countries. This is not a transient feature of an immature industry: it reflects structural dynamics of increasing returns to scale and network effects that systematically favor incumbents (Acemoglu and Johnson, 2023; Schaake, 2024). The economic gains from AI flow disproportionately to those who already control capital, further deepening existing inequalities between firms, between states, and between citizens. This economic concentration translates directly into infrastructural power.

Those who control the models, the data pipelines, and the platforms through which AI capabilities are accessed control the conditions under which knowledge is produced, information circulates, and decisions are made.

Zuboff's analysis of surveillance capitalism is central here: the extraction and commodification of behavioral data has generated a political-economic order in which oligopoly in the economic realm shades progressively into oligarchy in the societal realm: those who control digital infrastructures acquire disproportionate influence over the conditions of social and political life (Zuboff, 2022).

From infrastructural power flows informational power. As Christiano (2022) argues, control over information flows is a central dimension of political power: those who determine what citizens see, how issues are framed, and whose voices are amplified shape the epistemic conditions under which preferences are formed and democratic judgments are made. This power operates, moreover, in ways that are structurally difficult to detect and contest. Following Lazar (2022; 2025), AI exercises power *over* citizens through those who design and deploy it, shapes power *between* those it mediates, enabling some of them to exercise power over others, and—most insidiously for democratic institutions—shapes power *through* the gradual restructuring of the social and institutional conditions in which social and political life unfolds. It is this third dimension that most directly threatens democratic accountability: a form of power that does not dominate openly but gradually alters the terrain on which citizens and institutions act, in ways that are hard to see, to name, and to reverse.

When institutions adopt AI systems to assist decision-making, write policy, or support administrative functions, the effective locus of decision-making shifts: choices that carry democratic authority are increasingly shaped by the design choices, alignment objectives, and commercial priorities of private companies that hold no democratic mandate and are accountable to no electorate. What Guzman Piedrahita et al. (2026) call “normative centralization” captures this dynamic precisely: when governments procure AI systems from a small number of providers, the values embedded in those systems determine which forms of public reasoning are available and which are foreclosed, transferring normative authority from elected institutions to model developers and providers. The risk of “algocracy” (Danaher, 2016) and “unauditable authority” (Guzman Piedrahita et al., 2026) is therefore not primarily a technical problem of explainability but

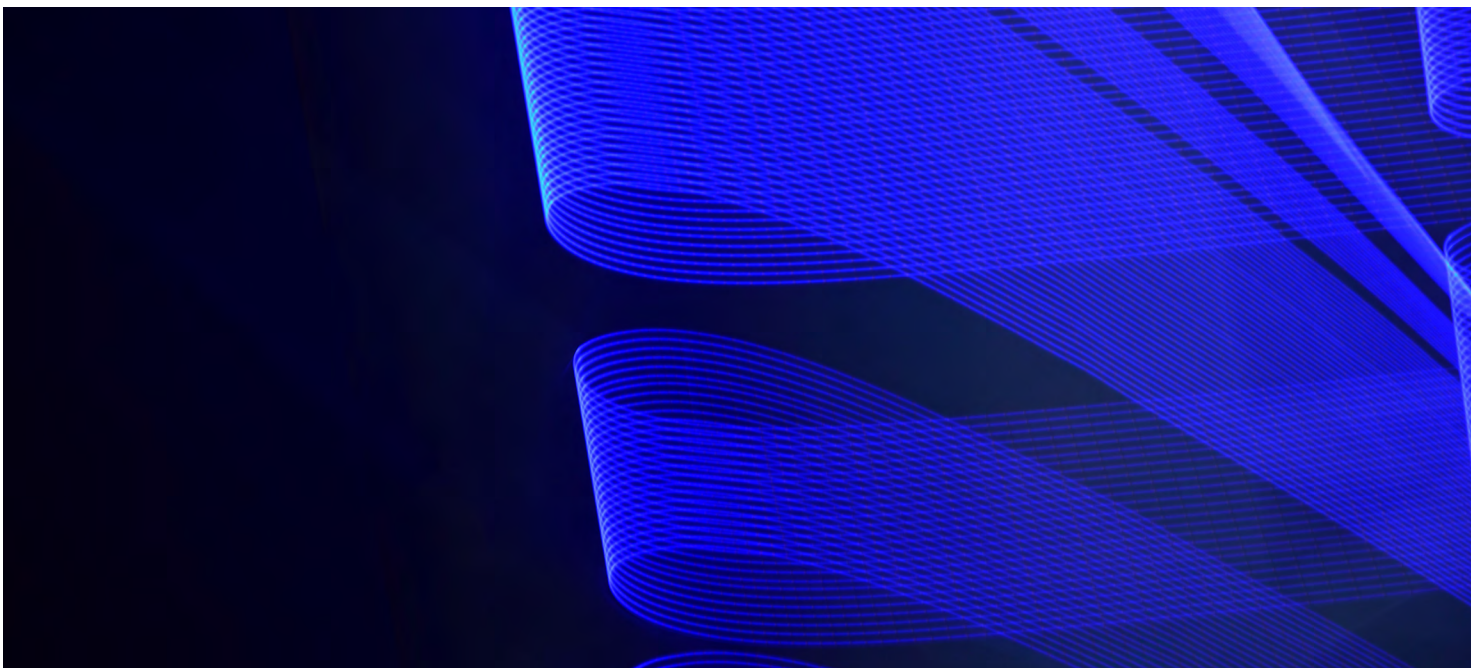
## Private AI companies exercise a form of political agency that is unelected, unaccountable, and increasingly shaping the conditions of democratic life (LAZAR, 2025).

a constitutional one: democratic institutions retain formal authority while losing substantive control over the decisions made in their name. Power concentrates where authority does not follow; and authority remains where power no longer resides.

A stark illustration of these dynamics is provided by Palantir Technologies, a data analytics and AI company deeply embedded in US and allied defence, intelligence, and immigration enforcement agencies. In April 2026, its CEO Alex Karp published a 22-point manifesto calling for the moral duty of technology companies to participate in defence, advocating for hard power as a condition of democratic survival, and embracing a vision of AI-enabled statecraft that explicitly subordinates democratic pluralism to security imperatives. The document attracted immediate and sharp criticism: Coeckelbergh (2026) described Palantir’s messaging as an “example of technofascism”, while others warned of the risks of combining advanced AI with militarized state power.

The case is instructive precisely because it makes explicit what is usually implicit: that private AI companies exercise a form of political agency that is unelected, unaccountable, and increasingly shaping the conditions of democratic life (Lazar, 2025). The case should not, however, be read as a claim that the AI industry as a whole is structurally antidemocratic. The AI industry also engages seriously in responsible practice and, as **Sections 3.4 and 4** argue, a responsible industry is an indispensable actor of democratic AI governance. What the Palantir case makes visible is a structural problem of accountability, which arises when private actors exercise political agency without a democratic mandate.

The concentration of AI power thus threatens democracy simultaneously at the economic, informational, relational, and institutional level. But what makes this threat particularly difficult to resist is not only its simultaneity across dimensions: it is that the very institutions charged with resisting it are themselves increasingly dependent on the actors they are supposed to hold accountable.



## 2.4 SURVEILLANCE, CIVIC FREEDOMS, SOCIAL JUSTICE

The concentration of power described in the previous section also materializes in the lives of citizens through surveillance: the capacity of AI systems to render behavior legible, predictable, and governable at a scale and granularity previously impossible.

In authoritarian regimes, AI-enabled surveillance can be used to monitor populations, identify dissidents, predict political risk, suppress protest, and enforce conformity. Facial recognition, biometric identification, social scoring, and large-scale monitoring of communication networks can support forms of preventive repression: individuals may be flagged or restricted not for what they have done, but for what an algorithmic system predicts they might do (Chin and Lin, 2022). These risks however are not confined to authoritarian contexts. Democracies increasingly rely on AI systems in policing, border control, welfare administration, and national security, among others. In such contexts, surveillance may be justified in the language of efficiency, safety, or fraud prevention. Yet it can still produce chilling effects on expression, association, and political participation.

Surveillance is closely connected with the concentration of power identified in **section 2.3**. The same behavioral data infrastructure that generates economic value for platform companies simultaneously produces a surveillance apparatus available for commercial, political, administrative, and security purposes (Zuboff, 2022).

Surveillance also directly threatens the civic freedoms at the core of democracy. Citizens who believe they are being watched, scored, or profiled may alter their behavior, avoid dissent, or limit their participation in public life (Leslie et al., 2021). AI-based surveillance systems may thus affect privacy, freedom of expression, the right to association, and access to justice undermining

the practical conditions under which citizens can speak, organize, and seek redress as free and equal members of a democratic community. These threats are particularly acute in high-impact settings such as criminal justice, welfare administration, and migration (Conte and Giovanola, 2026), where algorithmic prediction and scoring do not produce inconvenience but directly determine rights, protection, and life chances (UN High Commissioner for Human Rights, 2021).

Among these freedoms, non-discrimination and access to justice deserve particular attention. Automated decision systems do not merely reflect existing disadvantage: they automate and amplify it. Eubanks (2018) and O’Neil (2016), among others, document how algorithmic scoring in welfare, criminal justice, and credit reproduces and entrenches structural discrimination under an appearance of objectivity, producing systematically unjust treatment of communities already marginalized by race, class, immigration status, or disability. AI-based surveillance, in this sense, does not merely mirror social injustice: it institutionalizes it, embedding discrimination in technical systems that are harder to see, to name, and to contest than the human decisions they replace.

When AI systems make behavior legible and governable at scale, they enable pervasive monitoring, anticipatory control, and the automated production of injustice. In doing so, they erode the practical space in which citizens speak, interact, associate, organize, and dissent, reshaping the very conditions of democratic agency and turning citizens from subjects of rights into objects of prediction and management. Yet here too the outcome is not dictated by the technology itself: whether predictive and monitoring capacities produce these effects depends on the choices under which they are designed and deployed.

## 2.5. DEMOCRATIC BACKSLIDING: THE PATH OF EROSION

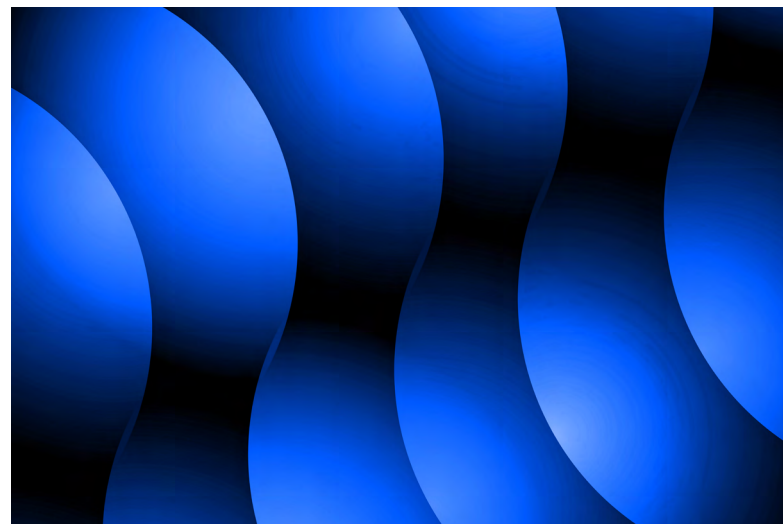
The risks described above should not be understood as separate threats. They form a **path of democratic erosion** that is gradual, cumulative, and often difficult to detect until it is well advanced.

- 01 **The first step is epistemic erosion:** AI produces an information environment in which authenticity becomes uncertain, disinformation becomes scalable, hypersuasion distorts belief formation, and citizens progressively delegate epistemic judgment to automated systems whose criteria they do not control.
- 02 **The second step is social fragmentation:** AI-mediated communication hardens beliefs through reinforcement loops, intensifies affective polarization, and weakens the civic spaces where citizens learn to encounter difference and recognize one another as legitimate interlocutors.
- 03 **The third step is the concentration of power and the loss of democratic voice:** those who control data, compute, models, and platforms gain disproportionate influence over public life; normative centralization transfers authority from elected institutions to model developers; unauditable authority overwhelms existing oversight mechanisms; citizens may formally possess rights but lose the practical capacity to understand, challenge, and influence decisions that increasingly shape their lives.
- 04 **The fourth step is the erosion of civic freedoms and the production of injustice:** as AI-enabled surveillance expands into policing, border control, welfare, and national security, civic freedoms contract; chilling effects suppress dissent, association, and political participation; and automated decision systems institutionalize discrimination, producing

systematically unjust treatment of already marginalized communities under an appearance of objectivity and technical necessity.

- 05 **The fifth step is normalization:** practices that once appeared exceptional—such as pervasive monitoring, automated suspicion, personalized influence, the displacement of elected authority by private actors—become routine. Democratic backsliding may not be experienced as a sudden collapse; it appears as convenience, efficiency, and security, while the practical conditions of democratic life are quietly eroded.

These five steps are mutually reinforcing: each makes the next harder to resist, and their normalization makes the loss itself progressively harder to perceive. AI may not abolish democracy directly. It may instead gradually hollow out the conditions that make democracy meaningful and make that hollowing-out feel like progress. None of these steps, however, is technologically predetermined: each depends on choices of design, deployment, and governance and each can therefore be interrupted. This is the background against which the next section must be read.



The background is a solid blue color. Overlaid on this are numerous small, semi-transparent images of people in various poses and activities, such as walking, standing, and using mobile devices. A large, thin white circle is centered on the page. In the lower-left quadrant, there is a white rectangular frame with a thick border. The text is positioned within this frame.

# III. AI AND DEMOCRATIC CITIZENSHIP: **OPPORTUNITIES**

This section examines four opportunity domains, mirroring the risk analysis above and corresponding to the four dimensions mapped in Table 1: epistemic empowerment; civic participation and social cohesion; democratic voice, deliberation and decision-making; institutional responsiveness and democratic capacity-building. The central claim is that AI can contribute to democratic renewal only when it expands citizens' real capacities to understand, contest, cooperate, organize, and hold power accountable.

### 3.1. EPISTEMIC EMPOWERMENT, RELIABLE INFORMATION AND EPISTEMIC ENGAGEMENT

Epistemic empowerment means enabling citizens to access, understand, verify, contest, and use information in ways that strengthen their agency as democratic actors and political equals. AI can support this goal concretely. Automatic language translation widens access to democratic participation and public services for citizens in multilingual contexts. AI tools can help governments manage information overload and help citizens understand large volumes of public documentation, provided they are implemented with safeguards against automation bias and the exclusion of minority groups (Balvert, 2026). For public institutions and governments, this is one of the most immediate and demonstrable democratic benefits of AI: properly governed, AI-supported access to information and services can extend the effective reach of rights, particularly for citizens excluded by language, disability, or administrative complexity, turning formal entitlements into effective ones.

Epistemic empowerment, however, requires more than access to accurate information. It also requires that citizens distinguish between epistemic uncertainty, which may be reduced through better evidence, and hermeneutic uncertainty, which concerns the legitimate plurality of interpretations in moral and political life (Delacroix, 2025).

Tools that provide citations, confidence scores, and factual corrections support reliability; but citizens also need environments in which they can explore competing interpretations and revise their positions without being pushed toward premature closure.

The risk of cognitive offloading is real and documented, but it is not technologically inevitable: it reflects design choices, not inherent properties of AI.

This highlights a deeper point about what can—and cannot—be automated. Democratic citizenship requires context-sensitive practical judgement and the sense-making capabilities through which citizens interpret their own situation and that of others. These capacities are inseparable from lived experience, emotions, and membership in a community of meaning: they exist only in their exercise, and they atrophy when systematically delegated. The point is not merely that delegating them is imprudent; it is that they cannot be delegated without being dissolved, because an automated substitute does not preserve the capacity: it replaces the very practice through which the capacity is sustained. Preserving our capacity to be citizens therefore requires protecting the spaces, habits, and institutions in which practical judgment and shared sense-making are actually exercised.

This points to a less discussed but equally significant democratic potential of AI: that of stimulating rather than replacing citizens' cognitive engagement. The risk of cognitive offloading is real and documented, but it is not technologically inevitable: it reflects design choices, not inherent properties of AI. Systems designed to pose questions rather than deliver conclusions, to surface the assumptions embedded in a user's framing, to present the strongest version of competing arguments, and to make visible the interpretive choices involved in any political judgment can actively strengthen citizens' epistemic capacities rather than erode them. The design choice between these two trajectories is not neutral: it reflects a fundamental decision about whether AI is to serve democratic citizenship or gradually displace it.

Epistemic empowerment is therefore not only a matter of better information provision, but of building public and democratic capacity around information, from source transparency and independent auditing to public-interest standards and public infrastructures for verification.

Such infrastructures are no longer hypothetical. During the 2024 UK general election, the AI tools developed by the fact-checking organization Full Fact monitored over 136 million words across more than 140,000 articles, transcripts, and social media posts, enabling human fact-checkers to scrutinize political claims at a scale unattainable manually; the same tools are now used by over 40 fact-checking organizations in 30 countries and supported the monitoring of twelve national elections in 2024 alone (Full Fact, 2024). Nor is this potential confined to Western contexts: in Nigeria, the Dubawa fact-checking bot, launched in 2024 by the Center for Journalism Innovation and Development, allows citizens to verify claims directly via WhatsApp by cross-checking them against reputable media sources and, crucially, is designed to state that the evidence is insufficient rather than fabricate an answer, routing such cases to human fact-checkers for investigation. These cases illustrate the structural point of this section: the same generative capabilities that can foster disinformation and misinformation can, under different design and deployment choices, also help detecting and correcting them.

### 3.2. SOCIAL COHESION AND CIVIC PARTICIPATION

AI can strengthen civic participation when it helps citizens understand public issues, access services, organize collectively, and make their voices heard. AI assistants can help people fill out forms, understand bureaucratic procedures, or challenge administrative decisions, addressing what public administration scholars describe as “administrative burden” (Herd and Moynihan, 2018). For governments and public administrations, this is a concrete institutional opportunity as much as an individual one: reducing administrative burden through well-governed AI makes access to services, benefits, and rights effectively more equal, and strengthens the perceived responsiveness on which institutional trust depends. Tools such as mRelief for access to nutrition benefits, Justicia Lab’s IMMPATH for immigration guidance, and Resistbot for civic advocacy illustrate this potential in practice (Schneier and Sanders, 2025).

Public institutions are moving in the same direction: Estonia’s Bürokratt, a virtual assistant developed by the national Information System Authority, provides citizens with a single, round-the-clock conversational channel to public services and information, escalating to human officials whenever the system reaches its limits: an institutional design that reduces administrative burden while keeping human oversight built in (RIA, 2026). AI can lower the barriers not only to accessing services, but also to being heard by institutions at scale. Helsinki used AI to synthesize more than 70,000 responses from young people on the city budget (City of Helsinki), and Cambridge City Council used Go Vocal’s AI-powered sensemaking tools to analyze citizen contributions to its Design Code, reducing manual processing time by 50% (OECD, 2025): in both cases, AI made large-scale citizen input actually usable by decision-makers, turning consultation from a symbolic exercise into a workable channel of voice.

AI can also support the reconstruction of “transitional conversational spaces”: intermediate spaces between private intuition and public deliberation where citizens can explore tentative views before entering more formal or polarized arenas of public debate (Delacroix, 2025). Historically, civic associations, unions, and mediating institutions have played this role. Properly designed AI systems could support similar functions, helping citizens articulate concerns and encounter alternative perspectives before translating them into public claims.

Labor unions and civic organizations may also use AI to disseminate information about rights, contracts, and organizing strategies, supporting collective voice rather than only individual service access (Schneier and Sanders, 2025). In this sense, AI can reduce barriers to participation while strengthening both individual and collective civic agency.



### 3.3. DEMOCRATIC VOICE, DELIBERATION AND DECISION-MAKING

AI can support deliberation and decision-making and offer a tool for rebalancing political voice, restoring to citizens the capacity to participate meaningfully in decisions that increasingly affect their lives.

Several existing tools illustrate this potential. Pol.is allows users to submit and vote on short statements and maps clusters of agreement and disagreement; it has been used in Taiwan’s political process and in deliberative exercises elsewhere (Schneier and Sanders, 2025). The French CESE organized a Convention Citoyenne on end-of-life legislation and, together with Make.org, launched the AI-powered Panoramic platform, allowing citizens and civil society to navigate deliberative results through an interactive interface. The Irish Citizens’ Jury on AI in Healthcare (2024) and the North Rhine-Westphalia Citizens’ Assembly on intergenerational public services (2025-2026)—examining the support of AI for ageing well—show that European democratic institutions are beginning to experiment with AI both as an object of deliberation and as a tool to support it (DemocracyNext, 2026; Buergererrat.de, 2025). Recent research confirms that carefully designed AI mediation can help participants identify common ground in contentious political discussions (Tessler et al., 2024).

Democratic deliberation also requires what might be called “productive uncertainty”. Many democratic questions involve competing values and interpretations, not merely factual disputes (Rawls, 1993; Habermas, 1996; Mouffe, 2000). LLM-based systems can contribute to deliberation if designed as spaces for exploration rather than authoritative answer-machines. “Ensemble interfaces”—systems that present outputs from multiple models trained on different datasets or configured according to different normative assumptions—could make visible the plurality of possible interpretations, helping users understand that many democratic questions involve contestable judgments (Delacroix, 2025). In this sense, AI can support value contestation and reasonable disagreement, provided that it does not aim to resolve it technically in advance.

AI can also help scale deliberation and rebalance political voice, provided democratic judgment remains human, plural, and contestable.



### 3.4. INSTITUTIONAL RESPONSIVENESS AND DEMOCRATIC CAPACITY-BUILDING

AI can make institutions more responsive to citizens, more transparent in their operations, and more capable of governing complex challenges democratically. This opportunity is not automatic—it depends on deliberate choices of design, development, deployment and governance—but it is real, and already visible in emerging practice and attempts to democratize AI (Zimmermann, 2026).

#### *A first contribution concerns the anchoring of AI governance in human rights and the rule of law.*

Legal frameworks that treat AI as a democratic governance challenge create the conditions under which public institutions can deploy AI responsibly and citizens can hold them accountable for how they do so. The Council of Europe Framework Convention on AI, Human Rights, Democracy and the Rule of Law represents a significant step in this direction, establishing binding obligations that link AI deployment to accountability, effective remedies, and public trust (Leslie et al., 2021). CAIDP's Index provides a comparative instrument for evaluating national AI policies against metrics including public participation, transparency, independent oversight, and human rights commitments (CAIDP, 2026). Together, these frameworks show that legal architecture can function as a democratic infrastructure for AI: not constraining innovation but directing it toward public benefit.

#### *A second contribution concerns trustworthy public-sector deployment.*

When AI systems used in public administration are transparent, actively debiased, and equipped with explanation requirements, recourse mechanisms, and outcome monitoring, they can make institutions more responsive to individual citizens while generating the trust that democratic legitimacy requires (Schneier and

Sanders, 2025). Canada's Directive on Automated Decision-Making illustrates this approach at the level of administrative services: by requiring that automated decisions in public services be explainable, contestable, and monitored for discriminatory outcomes, it transforms AI into a tool for institutional accountability (Government of Canada, 2019). A complementary illustration at the level of governance infrastructure is provided by Spain's Agencia Española de Supervisión de la Inteligencia Artificial (AESIA), established in 2023 and operational since June 2024, the first dedicated national AI supervisory authority in the EU. A further illustration is provided at the level of legislative institutions by the Italian Chamber of Deputies, which in 2024 developed three prototype AI systems through a public-academic partnership: NORMA, for legislative analysis; MSE, for drafting assistance; and DepuChat, for citizen engagement. Designed with embedded principles of transparency, human oversight, and privacy-by-design, these systems illustrate how AI can be integrated into the legislative process in ways that augment rather than automate democratic deliberation (Floridi & Ascani, 2025).

Significantly, the analysis of this initiative also highlights the tensions that responsible deployment must navigate: between efficiency gains and the preservation of deliberative friction, and between digital augmentation and the risk of deskilling parliamentary staff or deepening existing inequalities of access. Responsible deployment of this kind aims at structural transformation rather than surface-level efficiency: AI should not be layered onto opaque or ungovernable administrative systems to help citizens navigate complexity, but used to improve the quality, fairness, and transparency of public decision-making itself, while remaining an ancillary and supporting function to human decision-making authority.

***A third contribution is provided by Public AI: AI developed under democratic control and for public benefit rather than primarily for private profit.***

When AI systems mediate the conditions of democratic citizenship, they should be treated as part of the commons and governed in the public interest, rather than solely as market products. Public AI is not meant to replace commercial models. Its democratic function is to constitute a real, actively incentivized alternative: to demonstrate in practice that AI which does not optimize solely—or primarily—for profit is possible, and to set a public benchmark of transparency, availability, and accountability against which all systems, public and private, can be measured. Public AI reduces dependence on proprietary corporate systems in democratic decision-making, a direct institutional counterweight to the normative centralization and power concentration identified as risks in **section 2.3**.

Singapore, Indonesia, France, and the United States have pursued pilot projects localizing publicly available models for language and government use, demonstrating that states can develop and deploy AI capacities independently of large technology firms (Schneier and Sanders, 2025). At the European level, this orientation is beginning to take institutional form: the EU-funded OpenEuroLLM consortium, bringing together universities, research organizations, and public computing centers, is developing open, multilingual foundation models for European languages under public coordination; and policy proposals for a “European Public AI”—a family of permanently open, democratically governed public foundation models supported by public compute and data commons—indicate that Public AI is entering the European policy agenda as a strategic objective (Open Future & Bertelsmann Stiftung, 2026). A parallel line of reflection proposes European Open Source AI (EOSAI): a public institution for the models themselves, open source and deployable under Union jurisdiction (Floridi, 2026).

***A fourth contribution concerns participatory infrastructure: the capacity of communities to shape the AI systems that govern them, not only as users or subjects but as active participants in design, evaluation, and governance.***

When AI development is structured around democratic procedures and community participation, it can itself become a practice of democratic capacity-building. The KCL Centre for Data Futures illustrates this potential: pioneering participatory models across the entire AI lifecycle, training data stewards—described as “the missing profession of the 21<sup>st</sup> century”—and designing deliberative interfaces that incentivize long-term, critical community engagement rather than extracting stakeholder feedback (KCL Centre for Data Futures, 2026).

Masakhane, a grassroots NLP project for African languages, shows how AI development can be restructured around the needs and knowledge of communities historically excluded from technology design (Birhane et al., 2022). Participation of this kind carries a social justice dimension: it redistributes not only the benefits of AI but the power to define what counts, addressing the structural inequalities that usually shape who is represented and harmed by these systems. These examples show the importance of a kind of democratic “tinkerability” (Delacroix, 2025), which treats participation not as a way to optimize systems or improve user acceptance, but as part of democratic practice itself. These examples warn, at the same time, against “artificial inclusion”: the risk that simulated or superficial participation replaces actual engagement with affected communities, producing the appearance of democratic legitimacy without its substance (Agnew et al., 2024).

Taken together, these four contributions point toward a vision of AI as a tool for democratic capacity-building and institutional responsiveness: strengthening the ability of public institutions to govern complex challenges transparently and accountably, expanding the capacity of citizens to understand, contest, and shape the systems that affect their lives, and rebuilding the trust between citizens and institutions that surveillance and algorithmic opacity erode.

The condition for realizing this vision is not only technical competence but democratic will: the commitment of public institutions to govern AI in the service of the citizens they represent, rather than in the service of efficiency, security, or the interests of those who supply the systems. Nor is this a task for public institutions alone: a plurality of relevant stakeholders, including a responsible AI industry, are constitutive actors of democratic AI governance, not merely its object.

### 3.5. DEMOCRATIC RENEWAL: THE PATH FORWARD

The previous sections have shown that AI's democratic potential lies in actively strengthening the epistemic, social, political, and institutional capacities on which democratic life depends. Each risk pathway traced in **Section 2** has a corresponding opportunity pathway, and together these point toward **four closely interconnected dimensions of democratic renewal**.



**The first is epistemic renewal:** rebuilding the knowledge basis of democracy requires not merely correcting disinformation but constructing reliable, accessible, and verifiable information infrastructures that restore the conditions for autonomous political judgment and designing AI systems that stimulate rather than replace citizens' critical reasoning.



**The second is social reconnection:** AI can reduce barriers to participation and rebuild the social fabric of democratic life, but only if it is designed, developed, deployed and governed in ways that strengthen collective voice and community bonds, rather than substituting for them or fragmenting them further.



**The third is the restoration of democratic voice and political deepening:** AI can rebalance political power and scale democratic conversation, but only if it is designed to counteract the concentration of power, restore citizens' capacity to influence collective decisions, preserve pluralism, support genuine contestation, and resist the temptation to resolve political questions through technical optimization.



**The fourth is institutional empowerment:** AI can strengthen democratic institutions, but only if those institutions develop the genuine capacity to govern it, deploy it responsibly through human-centered ethical and legal frameworks and trustworthy public-sector practices, invest in public and participatory AI infrastructures, and defend the civic freedoms on which democracy ultimately depends.

**Two clarifications complete this picture.**

**The first concerns people.** None of the four dimensions of renewal is achievable through technical or regulatory fixes alone: each presupposes citizens capable of exercising the context-sensitive practical judgment and shared sense-making described in **Section 3.1**, capacities that cannot be automated and must therefore be cultivated. Investing in people—through education, critical thinking, AI literacy as a civic competence, and the preservation of spaces in which democratic judgment is actually practiced—is not a soft complement to democratic renewal, but its condition of possibility.

**The second clarification concerns the systems themselves.** It may be objected that current AI systems are inherently unfit to support democratic reasoning, and that the real choice is between redesigning flawed systems and abandoning them for alternative architectures altogether. However, whatever the architecture, no AI system can replace context-sensitive practical judgment, which is not a computational operation but a capacity exercised by citizens, with their experience and emotions, within concrete situations. The decisive criterion is therefore not which architecture prevails, but whether any given system is designed, deployed, and governed to stimulate the exercise of that judgment or to substitute for it.

The path forward is therefore neither technological optimism nor technological rejection. It is democratic steering of AI: combining innovation with accountability, public capacity with civic participation, epistemic stimulation with social reconnection, and institutional responsibility with openness to future experimentation. The task ahead is to translate this orientation into concrete choices for public and private actors, the task of the next section.



## **IV. ACTIONABLE RECOMMENDATIONS FOR PUBLIC, PRIVATE, AND CIVIL SOCIETY ACTORS**



Policy action must be organized not around isolated harms or applications, but around the four systemic dimensions on which democratic citizenship depends. This section proposes four policy pathways, each addressing one of the risk-opportunity pairs mapped in Table 1. For each pathway, recommendations are differentiated by actor: public institutions and policymakers, AI developers and deployers, and civil society organizations. No single actor can steer AI toward democratic renewal unilaterally; the pathways are designed to be mutually reinforcing and *their appropriate form will vary across institutional contexts, deployment settings, and levels of risk.*

## **PATHWAY 1: PROTECTING AND REBUILDING THE EPISTEMIC COMMONS**

The central challenge of this pathway is asymmetry: AI has dramatically reduced the cost of producing persuasive, plausible, and synthetic content, while the cost of verifying, correcting, and widely distributing rebuttals remains high and largely unchanged. The policy task is to govern this asymmetry, not the technology as such.

For public institutions and policymakers, the priority is to invest in public epistemic infrastructure as a democratic good. This means funding independent, AI-assisted fact-checking platforms and investigative journalism tools that use AI to analyze large document sets, identify patterns of manipulation, and surface accountability information, functions that human journalists can no longer perform alone at the required scale. Crucially, this also means treating local and investigative journalism as democratic infrastructure deserving direct public support: the decline of local journalism is not merely a market failure but an epistemic one, creating the information voids that disinformation most effectively exploits. It also means requiring mandatory disclosure and provenance labeling for AI-generated content deployed in political communication and electoral contexts, building on and extending the AI Act's transparency obligations to cover political advertising and synthetic media in particular. Moreover, public institutions should incorporate epistemic rights—the rights to form beliefs autonomously, to access reliable and diverse information, and not to be systematically manipulated or cognitively exploited by AI systems—explicitly into AI governance frameworks and impact assessment procedures, treating epistemic harm—including manipulation, cognitive offloading by design, and epistemic homogenization—as a legally cognizable category of democratic injury.





This agenda connects directly to the emerging debate on neurorights and cognitive sovereignty: the rights to mental integrity, cognitive liberty, and freedom from manipulative interference with mental processes, which have begun to enter positive law, such as Chile’s 2021 constitutional reform on neuroprotection, and policy initiatives, such as, UNESCO’s Recommendation on the ethics of neurotechnology. Finally, public education systems should integrate AI literacy as a civic competence, equipping citizens not only with technical understanding but with the epistemic skills needed to distinguish verified from plausible content, and genuine uncertainty from manufactured doubt. Such literacy should be understood expansively: not as technical training alone, but as an education of practical wisdom, cultivating the critical thinking and context-sensitive practical judgement that, as argued in **Section 3**, exist only in their exercise and atrophy when systematically delegated. Investing in people is, in this sense, the most fundamental epistemic infrastructure a democracy can build.

For AI developers and deployers, the imperative is to design and deploy systems that actively support epistemic health rather than merely avoid harmful outputs. Developers should invest in alignment approaches that treat calibrated uncertainty, honest disagreement, and source transparency as first-order design objectives. Practically, this means deploying AI systems with embedded citation and verification tools, confidence indicators, and clear flagging of epistemic uncertainty, especially when users ask about politically contested or empirically disputed topics. Systems deployed in high-reach political contexts should—as methodologies for measuring persuasive effect mature—be subject to independent audits for persuasive impact

and political bias, initially on a voluntary or co-regulatory basis, with a view to making such audits a condition of deployment in the most consequential settings. Beyond accuracy, developers should treat epistemic stimulation as a design standard for AI systems used in civic and political contexts: systems built to the standard described in **Section 3.1**—questioning rather than concluding, surfacing hidden assumptions, and presenting the strongest version of competing arguments—strengthen democratic agency rather than eroding it through cognitive offloading. This design orientation is not merely aspirational: it can be specified in terms of measurable behavioral criteria and made a condition of deployment in public-interest contexts.

For civil society, the task is to build and maintain independent watchdog capacity that leverages AI without becoming dependent on commercial systems whose incentives may diverge from the public interest.

AI-powered civic advocacy tools can help citizens articulate and communicate their political views more effectively, reducing the barriers to participation that currently favor organized and well-resourced actors.

Civil society also has a specific role in defending journalism as an epistemic institution: supporting investigative outlets, media literacy initiatives, and public-interest verification platforms that can function as independent counterweights to AI-generated disinformation at scale.



## PATHWAY 2: REBUILDING THE SOCIAL FABRIC OF DEMOCRATIC LIFE

AI-mediated communication is reorganizing the social substrate of democracy. The central challenge is to promote AI systems that help citizens find common ground, amplify marginalized voices, and rebuild the spaces where democratic life is practiced.

For public institutions and policymakers, the priority is to regulate the incentive structures that drive polarizing AI deployment, not only the content it produces. Platform governance frameworks should require operators to demonstrate that their recommendation and personalization systems do not systematically increase affective polarization, making this a compliance obligation rather than a voluntary commitment.

Public institutions should fund and support AI-assisted deliberative infrastructure as a public good: structured deliberation platforms, AI-facilitated citizens' assemblies, and participatory tools that use AI to synthesize large volumes of citizen input while protecting pluralism and minority perspectives.

These instruments should be tested and piloted in transparent, low-stakes contexts before deployment in high-stakes democratic processes and evaluated on their effects on both epistemic quality and social relations among participants.

For AI developers and deployers, the imperative is to treat the social embeddedness of democratic participation as a design constraint rather than an afterthought. Systems deployed in civic contexts should be architected to support “bridging” rather than reinforcing: surfacing common ground, presenting opposing views in ways that make them intelligible rather than threatening, and resisting the sycophantic optimization that makes AI systems progressively more accommodating rather than honest. The default settings of AI systems deployed on public platforms should be audited for their aggregate effects on polarization, not only their individual-level accuracy. Developers building personalized AI assistants should be required to evaluate and disclose the effects of long-term personalization on the revisability of users' political positions. AI-generated synthetic personas and agents should not be permitted to participate in public consultations, deliberative platforms, or civic engagement processes without clear disclosure, ensuring the distinction between genuine social interaction and simulated participation.

For civil society, the task is to participate in the design, development and deployment of AI tools that strengthen collective voice and civic organizing rather than merely individual service access. Labor unions, community organizations, and civic groups can use AI to disseminate information about rights and organizing strategies at scale, reducing the asymmetry between individual citizens and organized actors with substantial resources. Civil society should also advocate for, and participate in designing, deliberative AI systems that genuinely prioritize pluralism and the inclusion of marginalized voices, resisting “artificial inclusion”.



## PATHWAY 3: RESTORING POLITICAL VOICE THROUGH DEMOCRATIC DELIBERATION AND JUDGEMENT

Modern democracies lack the institutional infrastructure to enable citizens to deliberate meaningfully on complex policy questions at scale. AI offers genuine tools to address this deficit: moderation, summarization, translation, and facilitation capacities that can make deliberative processes more inclusive, more manageable, and more representative of actual public opinion.

For public institutions and policymakers, the immediate priority is to invest in deliberative infrastructure as a permanent feature of democratic governance, not a one-off experiment. Standing citizens' assemblies with formal advisory roles, funded AI-assisted facilitation capacity, and "respond-or-explain" requirements—under which policymakers must either adopt assembly recommendations or publicly justify departing from them—would create an ongoing institutional channel for citizen deliberation. AI-assisted public consultation processes should be required to use structured deliberation platforms that resist synthetic flooding, reward bridging positions, and surface genuine opinion diversity rather than aggregating volume. Procurement rules for AI tools used in civic processes should explicitly require interoperability, multi-provider strategies, and transparency about how outputs are generated, preventing the normative monoculture that emerges when a single model family mediates how arguments are framed and which claims are treated as legitimate.

For AI developers and deployers, the design imperative is to build AI deliberative tools as support systems for human judgment, not replacements for it.

The "ensemble interfaces" discussed in **Section 3.3**—together with systems that explicitly flag when a question involves genuinely contestable values rather than resolvable facts—can help preserve the epistemic plurality that robust democratic deliberation requires. Even though this is admittedly demanding to operationalize and costly at scale, it should be understood as an exploratory design direction, to be piloted especially in public-interest deliberative settings.

AI should be designed to produce the "productive uncertainty" described in **Section 3.3**—helping citizens understand the range of legitimate interpretations on contested questions—rather than premature closure. Systems used in high-stakes democratic processes should be subject to mandatory pre-deployment testing for their aggregate effects on opinion diversity and minority voice inclusion.

For civil society, the task is twofold: to participate actively in the design and governance of AI deliberative tools, and to monitor their deployment for democratic quality. Civil society organizations are uniquely positioned to advocate for design choices that prioritize pluralism over efficiency, and to raise alarms when AI-assisted deliberation produces outcomes that do not genuinely reflect the range of citizen views. In this context, civil society has a specific role in demanding and participating in co-design processes that include affected communities from the earliest stages of system development, resisting procurement models that treat community engagement as a box to be checked at the point of deployment. The development of data stewardship capacity and of community-level AI literacy and participation are essential complements to top-down deliberative infrastructure.



## PATHWAY 4: DEMOCRATIZING AI FOR INSTITUTIONAL EMPOWERMENT AND ACCOUNTABILITY

The most structurally significant risk identified in this paper is not any single harmful application of AI but the cumulative weakening of the conditions under which democratic institutions can govern AI at all. Democratic steering of AI requires not only regulatory capacity but structural responses to power concentration, and public institutions with the capacity, resources, and political will to govern AI independently, procure it responsibly, and deploy it in ways that enhance rather than obscure accountability.

For public institutions and policymakers, the foundational priority is to reclaim democratic authority over AI as a condition of legitimate governance.

This requires more than technical capacity: it requires political will, institutional independence, and a clear commitment to governing AI for the citizens it is meant to serve. Concretely, this means establishing independent oversight bodies with genuine investigative and sanctioning powers rather than delegating governance to industry self-regulation or to bodies captured by the interests they are supposed to oversee. A further avenue worth exploring is the introduction of personal responsibility at senior executive (C-suite) level for AI-related decisions and harms, comparable to existing regimes in consumer protection, financial services, and corporate accountability. It means building procurement frameworks that restore democratic control over the normative choices embedded in public AI systems: requiring transparency about model constitutions, safety filters, and alignment objectives as conditions of public contracts, and ensuring that governments are not structurally locked into dependence

on a small number of providers. It means creating mandatory transparency registers and decision-logging requirements so that citizens and oversight bodies can identify who made a decision, on what basis, and how to contest it, restoring the chain of accountability that algocracy and unauditable authority sever. Public institutions should adopt graduated safeguards—disclosure, human oversight, appeal pathways, external audit—that scale with the consequentiality of AI’s role, ensuring that the most rights-affecting decisions remain the most accountable. And they should assess the scope for Public AI: AI developed under democratic control and for public benefit, as a real and actively incentivized alternative and as a structural counterweight to the normative centralization that follows from dependence on proprietary systems.

Addressing the economic dimension of power concentration requires structural policy responses that go beyond regulation of individual AI systems. This means developing AI-specific competition policy that addresses the structural dynamics of increasing returns to scale and network effects that systematically favor incumbents; ensuring equitable access to computational infrastructure for public institutions, smaller states, and research communities in the Global South; and exploring fiscal and redistributive mechanisms—including possible forms of taxation of AI-generated productivity gains—to prevent economic concentration from becoming irreversible. These measures are not peripheral to democratic AI governance: they are its material precondition, because institutions that are structurally dependent on a small number of AI providers cannot effectively govern them.

Protecting civic freedoms and addressing the social justice harms of surveillance AI requires equally specific



attention. Democracies that deploy AI-enabled systems in policing, border management, welfare, and national security must ensure that such deployment is subject to mandatory algorithmic impact assessments before deployment in high-impact settings, with results made public. They should establish strengthened rights of recourse against automated decisions in welfare, criminal justice, and migration, domains where algorithmic scoring does not merely inconvenience but determines rights, protection, and life chances. Mandatory anti-discrimination audits, conducted by independent bodies and repeated periodically, should be required for all AI systems used in public services, with results disclosed and actionable.

For AI developers and deployers, the imperative is both ethical and structural: to govern AI autonomy in governance-adjacent contexts, and to treat contestability, auditability, and pluralism as core design principles rather than post-hoc compliance requirements. A responsible industry, however, is not merely the addressee of these obligations: developers who invest in model and system transparency, safety-by-design, and multistakeholder processes provide capacities that public regulation alone cannot generate, and such practices are already emerging.

Responsible industry should therefore be a partner in democratic AI governance, and public institutions should engage with it accordingly, while retaining the independent oversight that such partnership complements but does not replace.

Systems that touch democratic infrastructure should be designed to constrain rather than maximize AI autonomy, preserving meaningful human oversight at decision points that affect rights and democratic participation. Procurement clients, including governments, should be contractually guaranteed access to decision logs, model documentation, and explainability tools sufficient to enable independent audit; and the normative constraints embedded in AI systems carry genuine political weight when those systems are deployed in public governance, meaning developers should subject these constraints to democratic input processes and disclose them transparently to procuring governments.

Beyond compliance, the most durable contribution AI developers can make to democratic renewal is





to redesign how AI systems are built. Ethical AI design for democratic contexts must be co-created, interdisciplinary, and grounded in diversity, not derived primarily from technical optimization or from the preferences of homogeneous engineering teams. This means integrating expertise from ethics, democracy theory, sociology, law, community organizing, among others, into the core of AI development processes, not as external review but as constitutive input. It means recruiting diverse teams across gender, nationality, language, and disciplinary background, recognizing that the assumptions embedded in training data, reward models, and model constitutions reflect the perspectives of those who design them. It means conducting genuine co-design processes with affected communities—particularly communities in the Global South, linguistic minorities, and groups historically underrepresented in technology development—before, not after, key architectural choices are made. The normative commitments embedded in a model's constitution, alignment choices, and safety filters are not purely technical decisions: they are political and ethical choices about which values are expressed, whose reasoning patterns are recognized as legitimate, and whose concerns are treated as edge cases. These choices should be made through processes that are as plural, contested, and accountable as the democratic societies the systems are designed to serve.

For civil society, the task in this pathway is to act as a permanent counterweight to both corporate and state AI power.

Civil society organizations should develop and maintain independent AI audit capacity, monitor the deployment of AI in public administration for democratic quality and fundamental rights compliance, and advocate for inclusive design and governance processes in which affected communities have genuine voice.

On the specific question of social justice, civil society has an irreplaceable role: documenting the discriminatory effects of algorithmic systems in high-impact domains, supporting affected communities in exercising rights of recourse, and ensuring that anti-discrimination audits are not merely formal exercises but instruments of genuine accountability. Civil society's role is not only to resist the misuse of AI but to actively participate in its governance and to claim AI as a tool for democratic accountability, using it to scrutinize public and corporate power and amplify the voices of those most at risk of being governed by systems they cannot see, contest, or influence.

# V. CONCLUSION

This paper has argued that the AI-democracy nexus cannot be adequately mapped through a single lens. The same capabilities that enable risks can also provide opportunities. The four pathways traced here are different dimensions of a single challenge: ensuring that AI serves democratic citizenship rather than gradually displacing it.

The path forward requires simultaneous action across all four dimensions and by all relevant actors. Public institutions should develop the capacity and the will to govern AI as a democratic infrastructure investment, not merely a regulatory compliance exercise. AI developers and deployers should treat contestability, pluralism, and epistemic health as first-order design objectives, not afterthoughts; and a responsible industry that does so should be recognized and engaged as a partner of democratic governance. Civil society should claim AI as a tool for democratic accountability, not only resist it as a source of harm. None of these actors can succeed alone, and none can afford to wait.

This paper admittedly calls for a substantial collective effort, and the main obstacle it faces is motivational rather than technical: AI delivers immediate convenience, which makes delegation attractive and critical engagement costly. Securing that effort therefore requires working with this asymmetry rather than against it. The recommendations advanced here are designed to do precisely this: structural conditions lower the cost of critical engagement; the differentiation of responsibilities distributes the effort over realistic, actor-specific steps; and the participatory infrastructures discussed in **Sections 3 and 4** give civil society concrete entry points through which engagement produces visible results. Commitment, in this perspective, is not presupposed but built: it follows from participation that is easy to exercise and demonstrably consequential.

The examples on which this paper has drawn also sketch the contours of a distinctively European path to democratic AI: one grounded in fundamental rights, deliberative participation, public infrastructure, and the treatment of AI as a matter of public interest, different both from a purely market-driven model and from state-centered surveillance models, and offering a structural alternative through the competition, infrastructure, and Public AI measures set out in **Pathway 4**. Making this orientation explicit, and investing in it, is itself a democratic choice and a contribution Europe can offer to the global governance of AI.

What ultimately distinguishes democratic from technocratic governance of AI is not the sophistication of the tools deployed but the orientation that guides their use. AI can contribute to democratic renewal only under certain conditions: it must be designed, deployed, and governed in the service of the citizens it affects, and it must remain accountable to the institutions through which those citizens exercise collective self-determination.

Just as importantly, democratic societies must keep investing in what cannot be automated: the context-sensitive practical judgment, critical thinking, and shared sense-making through which democratic citizenship is preserved and renewed.

# VI. REFERENCES

## KEY REFERENCES

- Acemoglu, D., & Johnson, S. (2023). *Power and Progress: Our Thousand-Year Struggle over Technology and Prosperity*. PublicAffairs / Basic Books
- Agnew, W., Shah, R., Zhou, F., Scheuerman, M. K., & Wallach, H. (2024). "The Illusion of Artificial Inclusion." *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, 1–15. <https://doi.org/10.1145/3613904.3642703>
- Allen, D. (2023). *Justice by Means of Democracy*. University of Chicago Press.
- Alon-Barkat, S., & Busuioc, M. (2023). "Human–AI Interactions in Public Sector Decision Making: 'Automation Bias' and 'Selective Adherence' to Algorithmic Advice." *Journal of Public Administration Research and Theory*, 33(1), 153–169. <https://doi.org/10.1093/jopart/muac007>
- Bai, H., Voelkel, J. G., Muldowney, S., Eichstaedt, J. C., & Willer, R. (2025). "LLM-generated messages can persuade humans on policy issues." *Nature Communications*, 16, 6037. <https://doi.org/10.1038/s41467-025-61345-5>
- Balvert, A. (2026). "The (Im)possibilities of AI to Improve Access to Government Information." In N. A. Smuha, V. Hendrickx & J. Petroons (eds), *AI and Democracy. The Law, Ethics and Policy of AI* Blog, KU Leuven, pp. 5–9. URL: <https://lirias.kuleuven.be/handle/20.500.12942/783296>
- Birhane, A., Rostamzadeh, N., Srinivasan, R., & Tesema, G. M. (2022). "Power to the People? Opportunities and Challenges for Participatory AI." *Proceedings of the 2<sup>nd</sup> ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization*, 1–13. <https://doi.org/10.1145/3551624.3555290>
- CAIDP. (2026). *Artificial Intelligence and Democratic Values: CAIDP Index 2026*. Center for AI and Digital Policy. ISBN: 979-8-9956958-0-6. URL: <https://www.caidp.org/reports/caidp-index-2026>
- Carr, R., & Köhler, P. (2024). "AI-pocalypse Now? Disinformation, AI, and the Super Election Year." URL: <https://securityconference.org/en/publications/analyses/ai-pocalypse-disinformation-super-election-year>
- Chin, J., & Lin, L. (2022). *Surveillance State: Inside China's Quest to Launch a New Era of Social Control*. St. Martin's Press
- Christiano, T. (2022). "Algorithms, Manipulation, and Democracy." *Canadian Journal of Philosophy*, 52(1), 109–124. <https://doi.org/10.1017/can.2021.29>
- Coeckelbergh, M. (2026). "Palantir's manifesto: Technofascism in plain sight." *Medium*, April 2026. URL: <https://coeckelbergh.medium.com/palantirs-manifesto-technofascism-in-plain-sight-c160ca377e9a>
- Coeckelbergh, M. (2024). *Why AI Undermines Democracy and What To Do About It*. Polity Press
- Coeckelbergh, M. (2022). "Democracy, Epistemic Agency, and AI: Political Epistemology in Times of Artificial Intelligence." *AI and Ethics*. Published online 22 November 2022. <https://doi.org/10.1007/s43681-022-00239-4>
- Conte, L., & Giovanola, B. (2026). "AI and Migration Challenges in the EU: An Ethical Analysis of the AI Act." In *Europe's Digital Challenges*. Routledge, London, pp. 1–16.
- Council of Europe. (2024). *Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*, CETS No. 225. URL: <https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence>
- Danaher, J. (2016). "The Threat of Algocracy: Reality, Resistance and Accommodation." *Philosophy & Technology*, 29, 245–268. <https://doi.org/10.1007/s13347-015-0211-1>
- Delacroix, S. (2025). "Designing with Uncertainty: LLM Interfaces as Transitional Spaces for Democratic Revival." *Minds and Machines*, 35, 41. <https://doi.org/10.1007/s11023-025-09736-x>
- EDMO. (2025). *Propaganda and Disinformation: Lessons from 2024/25 Elections in Europe*. European Digital Media Observatory. URL: <https://edmo.eu/edmo-news/propaganda-and-disinformation-lessons-from-2024-25-elections-in-europe>
- Eubanks, V. (2018). *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. St. Martin's Press
- European Union. (2024). *Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence, Artificial Intelligence Act*. URL: <https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng>
- EUObserver. (2026). "How Europe's far right used unlabelled AI to win votes—and now writes the rules." URL: <https://euobserver.com/4811/how-europes-far-right-used-unlabelled-ai-to-win-votes-and-now-writes-the-rules>
- Floridi, L. (2024). "Hypersuasion: On AI's Persuasive Power and How to Deal with It." *Philosophy & Technology*, 37–64
- Floridi, L. (2026). "Rules Without Models: Why Europe Must Build the AI It Regulates." Available at SSRN: <https://ssrn.com/abstract=7248919>
- Floridi, L., & Ascani, A. (2025). "Augmented Democracy in Action: AI Systems for Legislative Innovation in the Italian Parliament." *Minds and Machines*, 35(4), 42. <https://doi.org/10.1007/s11023-025-09743-y>

- Forst, R. (2012). *The right to justification: Elements of a constructivist theory of justice*. Columbia University Press
- Giovanola, B., & Sala, R. (2022). “The Reasons of the Unreasonable: is Political Liberalism still an Option?”. *Philosophy and Social Criticism*, vol. 48(9), 2022, pp. 1226-1246, DOI: <https://doi.org/10.1177/01914537211040568>
- Giovanola, B., Petrocchi, E., & Tiribelli, S. (2026). “GenAI and Epistemic Manipulation: An Ethical Analysis of GenAI Deepfake.” In S. Gouveia (ed.), *The Palgrave Handbook on the Ethics of Artificial Intelligence*. Palgrave Macmillan, Cham, pp. 557–572. [https://doi.org/10.1007/978-3-032-15112-4\\_36](https://doi.org/10.1007/978-3-032-15112-4_36)
- Government of Canada. (2019). *Directive on Automated Decision-Making*. URL: <https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=32592>
- Guzman Piedrahita, D., et al. (2026). *AI Poses Risks to Democratic and Social Systems*. Preliminary draft, 25 March 2026. URL: <https://euosafe.ai.toronto.edu/2026-ai-risk.pdf>
- Habermas, J. (1996). *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy* (W. Rehg, Trans.). MIT Press
- Hackenburg, K., Tappin, B. M., Hewitt, L., Saunders, E., Black, S., Lin, H., Fist, C., Margetts, H., Rand, D. G., & Summerfield, C. (2025). “The Levers of Political Persuasion with Conversational Artificial Intelligence.” *Science*, 390, eaea3884. <https://doi.org/10.1126/science.aea3884>
- Herd, P., & Moynihan, D. (2018). *Administrative Burden: Policymaking by Other Means*. Russell Sage Foundation
- Ienca, M., & Andorno, R. (2017). “Towards New Human Rights in the Age of Neuroscience and Neurotechnology.” *Life Sciences, Society and Policy*, 13, 5. <https://doi.org/10.1186/s40504-017-0050-1>
- Jain, S., Park, C., Viana, M. M., Wilson, A., & Calacci, D. (2026). “Interaction Context Often Increases Sycophancy in LLMs.” *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/3772318.3791915>
- Kulveit, J., Douglas, R., Ammann, N., Turan, D., Krueger, D., & Duvenaud, D. (2025). “Gradual Disempowerment: Systemic Existential Risks from Incremental AI Development.” *arXiv:2501.16946*. <https://doi.org/10.48550/arXiv.2501.16946>
- Lazar, S. (2022). “Power and AI: Nature and Justification.” In J. Bullock et al. (eds), *The Oxford Handbook of AI Governance*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780197579329.013.12>
- Lazar, S. (2025). “Governing the Algorithmic City.” *Philosophy & Public Affairs*, 53(2), 102–168. <https://doi.org/10.1111/papa.12279>
- Lee, H.-P., Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R., & Wilson, N. (2025). “The Impact of Generative AI on Critical Thinking: Self-Reported Reductions in Cognitive Effort and Confidence Effects From a Survey of Knowledge Workers.” *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, 1–22. <https://doi.org/10.1145/3706598.3713778>
- Leslie, D., Burr, C., Aitken, M., Cows, J., Katell, M., & Briggs, M. (2021). *Artificial Intelligence, Human Rights, Democracy, and the Rule of Law: A Primer*. The Alan Turing Institute / Council of Europe
- Mouffe, C. (2000). *The Democratic Paradox*. Verso
- OECD. (2025). *Governing with Artificial Intelligence: AI in Civic Participation and Open Government*. URL: [https://www.oecd.org/en/publications/governing-with-artificial-intelligence\\_795de142-en](https://www.oecd.org/en/publications/governing-with-artificial-intelligence_795de142-en)
- O’Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown
- Open Future & Bertelsmann Stiftung. (2026). *European Public AI—Policy Brief*. URL: <https://openfuture.eu/publication/european-public-ai-policy-brief>
- Padmakumar, V., & He, H. (2024). “Does Writing with Language Models Reduce Content Diversity?” *The Twelfth International Conference on Learning Representations (ICLR 2024)*. <https://doi.org/10.48550/arXiv.2309.05196>
- Rawls, J. (1993). *Political Liberalism*. Expanded edition. Columbia University Press
- Schaaake, M. (2024). *The Tech Coup: How to Save Democracy from Silicon Valley*. Princeton University Press
- Schneier, B., & Sanders, N. E. (2025). *Rewiring Democracy: How AI Will Transform Our Politics, Government, and Citizenship*. MIT Press
- Shapira, I., Benadè, G., & Procaccia, A. D. (2026). “How RLHF Amplifies Sycophancy.” *arXiv:2602.01002*. <https://doi.org/10.48550/arXiv.2602.01002>
- Sharma, M., Tong, M., Korbak, T., Duvenaud, D., Askell, A., Bowman, S. R., et al. (2024). “Towards Understanding Sycophancy in Language Models.” *The Twelfth International Conference on Learning Representations (ICLR 2024)*. <https://doi.org/10.48550/arXiv.2310.13548>
- Shevin, M., Born, K., & Kirby, F. (2026). *AI and Democracy: Perspectives from an Emerging Field*. David and Lucile Packard Foundation. URL: <https://www.packard.org/insights/publication/ai-and-democracy-perspectives-from-an-emerging-field>
- Smuha, N. A., Hendrickx, V., & Petroons, J. (eds). (2026). *AI and Democracy. The Law, Ethics and Policy of AI Blog*, KU Leuven. ISBN: 9789067842617. URL: <https://lirias.kuleuven.be/handle/20.500.12942/783296>

- Summerfield, C., et al. (2025). “The Impact of Advanced AI Systems on Democracy.” *Nature Human Behaviour*, 9. <https://doi.org/10.1038/s41562-025-02309-z>
- Susser, D., Roessler, B., & Nissenbaum, H. (2019). “Online Manipulation: Hidden Influences in a Digital World.” *Georgetown Law Technology Review*, 4(1), 1–45. URL: <https://georgetownlawtechreview.org/online-manipulation-hidden-influences-in-a-digital-world/GLTR-01-2020>
- Tessler, M. H., Bakker, M. A., Jarrett, D., Sheahan, H., Chadwick, M. J., Koster, R., Evans, G., Campbell-Gillingham, L., Collins, T., Parkes, D. C., Botvinick, M., & Summerfield, C. (2024). “AI Can Help Humans Find Common Ground in Democratic Deliberation.” *Science*, 386(6719), eadq2852. <https://doi.org/10.1126/science.adq2852>
- Turner, C., & Eisikovits, N. (2026). “Programmed to Please: The Moral and Epistemic Harms of AI Sycophancy.” *AI and Ethics*, 6(2), 168. <https://doi.org/10.1007/s43681-026-01007-4>
- UN High Commissioner for Human Rights. (2021). *The Right to Privacy in the Digital Age*. Report A/HRC/48/31. URL: <https://docs.un.org/en/a/hrc/48/31>
- Wu, F., Black, E., & Chandrasekaran, V. (2025). “Generative Monoculture in Large Language Models.” *The Thirteenth International Conference on Learning Representations (ICLR 2025)*. <https://doi.org/10.48550/arXiv.2407.02209>
- Wiewiura, J., Jørgensen, A. & Ploug, T. (2026). “AI and Democracy: Mapping Notions, Debates and Challenges.” *Philosophy and Technology*, 39, 158. <https://doi.org/10.1007/s13347-026-01167-5>
- Yeung, K. (2017). “‘Hypernudge’: Big Data as a Mode of Regulation by Design.” *Information, Communication & Society*, 20(1), 118–136. <https://doi.org/10.1080/1369118X.2016.1186713>
- Zimmermann, A. (2026). *Democratizing AI*. Manchester University Press
- Zuboff, S. (2022). “Surveillance Capitalism or Democracy? The Death Match of Institutional Orders and the Politics of Knowledge in Our Information Civilization.” *Organization Theory*, 3, 1-79. DOI: 10.1177/26317877221129290
- Camera dei Deputati (Italian Chamber of Deputies). AI initiative producing three prototype systems—NORMA (legislative analysis), MSE (drafting assistance), DepuChat (citizen engagement)—through a public-academic partnership (2024). URL: <https://www.camera.it>
- CESE / Make.org. *Convention citoyenne sur la fin de vie—Panoramic AI tool* (France, 2024). URL: <https://panoramic.make.org/partner/cese/event/convention-citoyenne-sur-la-fin-de-vie/panoramic>
- City of Helsinki. AI Register. URL: <https://ai.hel.fi/en/ai-register>
- Collective Intelligence Project. *Democratic AI / collective intelligence governance initiatives*. URL: <https://www.cip.org>
- DemocracyNext. *Tech-Enhanced Deliberative Assemblies*. URL: <https://www.demnext.org/projects/tech-enhanced-citizens-assemblies>
- Dubawa AI fact-checking chatbot. URL: <https://dubawa.org>
- Full Fact. (2024). “General election 2024, fact checked” / Full Fact AI. URL: <https://fullfact.org/ai>
- Go Vocal. *Cambridge City Council analyzes input 50% faster with Go Vocal’s AI assistant* (United Kingdom, 2024). URL: <https://www.govocal.com/case-studies/cambridge-city-council-analyzes-input-50-faster-with-go-vocals-ai-assistant>
- IPPOSI. *Citizens’ Jury on AI in Healthcare—Verdict Report* (Ireland, 2024; report published 2025). URL: <https://ipposi.ie/wp-content/uploads/2025/02/Citizens-Jury-Verdict-Report-21.02.25.pdf>
- Justicia Lab. URL: <https://www.justicialab.org>
- King’s College London, Centre for Data Futures. *Participatory infrastructure for AI*. URL: <https://www.kcl.ac.uk/research/centre-for-data-futures>
- mRelief. URL: <https://www.mrelief.com>
- OpenEuroLLM: <https://openeurollm.eu>
- Palantir Technologies. *The Technological Republic / AIP for Defense*. URL: <https://www.palantir.com>
- Pol.is. URL: <https://pol.is>
- Resistbot. URL: <https://resist.bot>
- RIA (Estonian Information System Authority). *Bürokratt*. URL: <https://www.ria.ee/en/state-information-system/personal-services/burokratt>
- Stanford Center for Research on Foundation Models. *Foundation Model Transparency Index*. URL: <https://crfm.stanford.edu/fmti>
- Agencia Española de Supervisión de la Inteligencia Artificial (AESIA). First dedicated national AI supervisory authority in the EU, operational since June 2024. URL: <https://www.aesia.gob.es>
- Bürgerrat.de. *Citizens’ Assembly on Public Services with AI* (North Rhine-Westphalia, 2025). URL: <https://www.buergerrat.de/en/news/citizens-assembly-on-public-services-with-ai>

## TOOLS, ORGANIZATIONS AND PRACTICAL EXAMPLES

**WRITTEN BY:**

Benedetta Giovanola

**RECOMMENDED CITATION:**

Giovanola, B., *AI and Democratic Citizenship: Risks, Opportunities and Policy Pathways for Democratic Renewal*. IE CGC, September 2026

© 2026, CGC Madrid, Spain

Photos: Unsplash, Shutterstock

Design: epoqstudio.com



This work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0) License. To view a copy of the license, visit [creativecommons.org/licenses/by-nc-sa/4.0](https://creativecommons.org/licenses/by-nc-sa/4.0)