Macrothink Institute™

Journal of Sociological Research
ISSN 1948-5468
2026, Vol. 17, No. 1

Algorithms and Consensus: A Critical Perspective on Artificial Intelligence and Political Communication

Daniele Battista[1,*] & Arthur Gugliucciello[1]

[1]University of Salerno, Italy

*Corresponding author: University of Salerno, Italy

Received: March 24, 2025 Accepted: August 11, 2025 Published: January 20, 2026

doi:10.5296/jsr.v17i1.23859 URL: https://doi.org/10.5296/jsr.v17i1.23859

Abstract

In contemporary society, the technological landscape has been profoundly transformed by the rapid and continuous evolution of artificial intelligence (AI), which is attracting increasing attention in a variety of contexts. For these reasons, this phenomenon has stimulated considerable reflection on various fronts, making a multidisciplinary approach to its study necessary. This article offers a critical-theoretical reflection, analysing the use of AI in international politics in recent years. This provides a privileged vantage point from which to question how this situation, and the nascent "synthetic propaganda," can influence and shape social dynamics. Political communication itself has taken on a new dimension in the digital society, changing its principles and influencing the way citizens are addressed. The analysis examines the risks of a drift towards forms of manipulation that can have a negative influence on society and on the relationship with political institutions themselves. The aim of the document is to provide an original contribution to the study of the dynamics that arise from the interaction between AI and political power, analysing both the opportunities offered and the risks arising from the use of this technology.

Keywords: AI, political communication, manipulation, democracy, digital society

1. Introduction

Today's scenario sees exponential growth in Artificial Intelligence (AI) in various fields, marking a profound restructuring of the media ecosystem in which democratic consensus is built and, at times, guided. If we have moved from a 'party' democracy to a 'public' democracy (Manin, 1997), today we are faced with what could be called an 'algorithmic democracy' (García-Marzá & Calvo, 2024). In this new paradigm, computer codes take on an almost legislative role, establishing the unwritten rules of inclusion and exclusion in public debate. Digital infrastructures do not merely host discussions, but become their invisible architects, often programmed to reward polarizing content at the expense of calm and in-depth reflection. The visibility and relevance of political issues are no longer determined exclusively by the agenda of traditional media or physical public spaces but follow the "opaque" and automated logic of digital platforms (Rieder & Hofmann, 2020). This transition has generated a profound crisis of authority for traditional information intermediaries, such as journalists and historical opinion leaders, who now find themselves having to chase the engagement metrics imposed by platforms. The result is an extremely fragmented information ecosystem, in which shared truth is gradually giving way to micro-narratives tailored to the niches of individual users, fuelling cognitive bubbles that are difficult to break. This evolution is not without positive potential for the public sphere. The integration of intelligent systems allows us to exploit unprecedented opportunities to breathe new life into the relationship between the electorate and institutions. For example, the use of advanced chatbots or virtual assistants could streamline bureaucracy, bringing public administration closer to segments of the population that have historically been less involved. Furthermore, the ability to analyse sentiments and needs expressed on social networks in real time offers policymakers a formidable tool for calibrating public policies more quickly and accurately, overcoming the inherent slowness of traditional consultations. In recent years, intelligent systems have moved beyond the confines of research laboratories and infiltrated the very heart of political processes: from electoral data management to predictive policy analysis to mediation between citizens and institutions (Battista, 2025). In this perspective, AI can act as a catalyst for more accessible and personalized participation, reducing information barriers and enabling more direct communication. Political communication, based on principles such as rhetoric, dialectical confrontation, and stage presence, is inevitably moving towards forms of computational propaganda (Bradshaw & Howard, 2019). Such forms of automated manipulation often operate below the voter's conscious attention threshold, exploiting cognitive biases through the use of inauthentic accounts, bots, and coordinated networks. The goal is no longer just to convince rationally, but to saturate the discursive space in order to confuse, disorient, and ultimately dissuade opponents from political participation, polluting the debate with incessant background noise. This is a scenario in which the interaction between candidate and voter is increasingly mediated by non-human technological artifacts capable of processing enormous amounts of data to profile, target, and influence users with unprecedented precision and effectiveness. This change in the human experience risks transforming it into free raw material for predictive policies, thus making citizens passive objects of behavioural calculation (Zuboff, 2019). The continuous extraction of psychographic data allows for the creation of increasingly accurate predictive models,

capable of intercepting not only politically expressed preferences, but also the latent emotional vulnerabilities of individuals. This gives rise to a new form of structural power, in which large technology companies effectively hold the keys to democratic infrastructure, operating in many cases outside the traditional checks and balances that characterize states governed by the rule of law. It is therefore essential to adopt a critical perspective that allows us to question the social implications of this change, enabling us to avoid algorithmic drift (Vreese, 2023). The urgency of this investigation is now empirically confirmed in many contexts of international politics, which are taken as paradigmatic case studies in this work. The analysis of these scenarios will allow us to observe how computational propaganda has gone beyond factual disinformation to arrive at "synthetic emotional propaganda," in which AI does not merely manipulate data, but fabricates imaginaries and alternative realities capable of redefining the very foundations of democratic consensus. Methodologically, the analysis adopts a qualitative approach based on the study of these cases, selected for their relevance in illustrating the primary ways in which artificial intelligence impacts democratic processes. This examination makes it possible to link empirical evidence with the theoretical framework, highlighting both the risks associated with algorithmic political disinformation and the potential of AI as a tool for participation.

2. The Impact of Artificial Intelligence on Politics and Democracy

In the current political landscape, the link between digital technologies and power has become so symbiotic that it has radically transformed the way politics is conceived and consumed (Papacharissi, 2002). To adequately explore the many facets of this symbiosis, this paper adopts a precise methodology based on a critical-theoretical approach. The goal is to reveal how technology is not a simple neutral tool, but a biopolitical and socio-technical device capable of actively shaping the subjectivity of citizens and the very ontology of public space, thus requiring rigorous ethical scrutiny. This transformation is not limited to a simple digitization of existing processes but reconfigures the entire media ecosystem into a hybrid environment, where the vertical transmission logic of traditional mass media is inextricably intertwined with the horizontal dynamics of online sharing and participation (Bentivegna & Boccia Artieri, 2019). This evolution has introduced unprecedented dynamics between institutional actors, media, and citizens. In this hybrid ecosystem, political leaders strategically disintermediate their communication, bypassing the critical and fact-checking filter of traditional journalism to address 'digital squares' directly, where the message is constantly adapted based on instant feedback from the network. Algorithmic modelling plays a central role, tending to fragment the space for discussion: the personalization of content mediated by recommendation algorithms often ends up creating "echo chambers" that limit democratic debate (Kaye, 2018). This personalized filtering mechanism isolates users in an information universe that exclusively reflects their previous preferences (Pariser, 2011), rendering opposing arguments invisible and drastically reducing opportunities for shared, pluralistic public debate (Sunstein, 2017). This cognitive isolation is not an accidental flaw in the system, but the intentional result of a business model that maximizes profits by keeping users on the platform for as long as possible. The immediate democratic consequence is

unprecedented emotional polarization, in which political opponents cease to be perceived as bearers of different ideas with whom to engage in dialogue and instead become threats to the values of one's own group. Furthermore, the technical architecture of platforms is never neutral, but exercises real infrastructural power that shapes users' possibilities for action, often hiding discriminatory biases behind a facade of mathematical objectivity (O'Neil, 2016). In this scenario, artificial intelligence does not merely optimize messages but imposes new parameters of success based on emotional engagement and virality rather than factual accuracy, transforming political competition into a race for attention governed by an opaque logic (Pasquale, 2015). Further complicating this relationship is the massive use of AI, which introduces new challenges for information integrity. AI's ability to produce artificial multimedia content has paved the way for a season of automated disinformation, which erodes trust in institutions and consolidates a ‘post-truth’ regime, in which objective facts lose relevance compared to emotions and personal beliefs in the formation of public opinion (McIntyre, 2018). The proliferation of deepfakes, voice cloning, and the large-scale generation of persuasive synthetic texts undermines the very foundations of trust upon which civil society is built. When citizens lose the ability to distinguish genuine events from hyper-realistic fabrications, a climate of generalized epistemic cynicism emerges, in which the a priori rejection of any information becomes a defensive mechanism against constant manipulation. The combination of algorithms and artificial intelligence reshapes citizens' cognitive structures, exposing them to manipulative narratives that often exceed their capacity for critical reflection. Nevertheless, it is essential to avoid an absolutist form of technological determinism: there remain spaces for political action that escape the logic of digital platforms and enable forms of civic mobilization capable of breaking through algorithmic "filter bubbles" (Vaccari & Valeriani, 2021). These dynamics can be better understood through several paradigmatic case studies that illustrate how artificial intelligence can directly affect democratic processes (Battista, 2024). A first example is provided by the 2024 New Hampshire Democratic primary in the United States, during which thousands of voters received automated phone calls featuring an AI-generated imitation of President Joe Biden's voice. The message urged recipients not to participate in the primary election, with the apparent aim of suppressing voter turnout. Although the authorities later identified those responsible and imposed sanctions, the episode demonstrated how voice-cloning technologies can be exploited to undermine public trust and interfere directly with the integrity of electoral processes. More broadly, the case illustrates that political manipulation is no longer confined to the dissemination of false information but increasingly relies on the credible simulation of institutional identities, making it progressively more difficult to distinguish authentic communication from artificially generated content. A second emblematic case concerns Pakistan's 2024 general election. Following the imprisonment of former Prime Minister Imran Khan, his political party employed artificial intelligence to recreate his voice from previous recordings, enabling him to communicate with supporters despite the restrictions imposed by his detention. In this instance, the same technology commonly associated with misinformation was instead used to facilitate political participation and circumvent censorship. The episode highlights that artificial intelligence is not inherently a source of democratic erosion but rather a technological infrastructure whose consequences depend on

the objectives pursued by political actors and the institutional context in which it is deployed. The Russia–Ukraine conflict provides another significant example of AI's growing role in information warfare. In March 2022, a deepfake video depicting Ukrainian President Volodymyr Zelensky apparently urging Ukrainian troops to surrender was disseminated online. Although the video was technically unsophisticated and was quickly debunked by Ukrainian authorities, it represented one of the earliest explicit uses of deepfake technology as an instrument of psychological warfare. Its strategic significance lies less in its immediate effectiveness than in its demonstration that audiovisual manipulation can become an integral component of disinformation campaigns during international conflicts. Taken together, these cases demonstrate that the central challenge posed by artificial intelligence lies not merely in its capacity to generate false content, but in its transformation of the epistemic conditions under which citizens form political judgments. The increasing difficulty of distinguishing truth from falsehood contributes to what many scholars describe as a "crisis of trust," in which the declining credibility of information sources progressively weakens democratic deliberation. Understanding this phenomenon therefore requires moving beyond its technical dimension to address its broader ethical and political implications. The rise of artificial intelligence represents a profound historical transformation, as it increasingly functions as both an epistemic and political actor capable of processing complex data and adapting in real time to individual preferences (Floridi, 2022). Although AI remains a form of "simulated intelligence" focused on replicating cognitive functions (Vicente, 2020), its growing influence redefines the very conditions of democratic participation (Battista, 2023), making an urgent reflection on transparency, accountability, and equitable access to information indispensable. From a critical theoretical perspective, artificial intelligence should therefore not be understood as an inevitable technological destiny but as a governable socio-technical construct. In this regard, initiatives such as the European Union's AI Act, which establishes transparency obligations for certain AI systems and specific requirements for AI-generated content, together with the strengthening of independent fact-checking mechanisms and digital literacy programmes, represent important steps toward democratic oversight. Restoring citizens' capacity for autonomous political agency ultimately requires the promotion of genuine algorithmic literacy and the redesign of regulatory frameworks capable of subjecting these algorithmic "black boxes" to continuous and independent public scrutiny. In conclusion, the joint examination of the analysed cases makes it possible to move beyond their merely descriptive dimension and to identify several recurring structural patterns. In all three episodes, what emerges first and foremost is a transformation of the very notion of “evidence” in contemporary political discourse, in line with what Lippmann (2017) already highlighted regarding the crisis of opinion formation in complex democracies. Credibility no longer depends on the verification of content, but on its perceptual plausibility and the prior trust placed in the source. The synthetic voice, the deepfake image, or the artificially generated message do not need to be believed to produce political effects; they only need to introduce a stable margin of epistemic uncertainty. Secondly, the cases highlight a growing temporal compression between the production, dissemination, and debunking of information, which results in a structural crisis of the deliberative public sphere. In the U.S. electoral context, the time window of disinformation is sufficient to affect voter turnout before institutional, or

media corrective mechanisms can intervene. This shifts attention away from the truthfulness of content towards its capacity for circulation, making speed a properly political variable. The Pakistani case instead introduces a different logic, in which artificial intelligence does not produce informational exclusion but rather communicative continuity under conditions of political restriction. Here, technology operates as an infrastructure of resilience, while at the same time confirming Floridi's notion (2015) of the growing "onlife condition," in which the distinction between real and mediated presence becomes increasingly blurred. This dynamic entails a transformation in the relationship between leadership and citizenship, increasingly mediated by systems of synthetic generation of public presence. In the Ukrainian context, finally, the cognitive dimension intersects with the strategic and systemic one: the deepfake does not merely aim to deceive, but to saturate the informational environment, generating generalized uncertainty and progressively reducing the epistemic stability of the political system. This phenomenon fits within the logic of the so-called information disorder (Gallo et al. 2022), where the objective is not necessarily immediate persuasion, but the cumulative erosion of trust in information sources. Taken together, these cases indicate that the impact of artificial intelligence cannot be understood through the binary of true/false, but rather through its capacity to alter the conditions under which political credibility is produced. This leads to a structural shift from a verification-based regime of information to one based on the strategic management of uncertainty, in which trust becomes a scarce and politically contested resource.

3. The Risks of Algorithmic Mediation: Manipulation, Opacity, and Crisis of Debate

The pervasive integration of artificial intelligence systems into political communication processes not only raises technical questions but also profoundly reconfigures the very architecture of the public sphere. Although the promises of efficiency and personalization are evident, critical analysis must necessarily focus on the systemic distortions that this infrastructure imposes on democracy. The main risk lies not in a single act of disinformation, but in the structural transformation of the environment in which citizens form their opinions, an environment increasingly governed by invisible commercial logic and mechanisms of information segregation. A first level of criticality concerns the very nature of the platforms that host the debate. Contrary to the ideal image of an open and neutral "digital square," these spaces are techno-social constructs governed by precise rules, where public values are systematically subordinated to the business models of large technology companies. The platform society is not a simple container, but an actor that actively shapes social practices: the mechanisms of datafication and commodification transform citizens into users and political debate into monetizable engagement flows, often at the expense of the quality of deliberation (Van Dijck, Poell & de Waal, 2018). This commodification of attention is based on the systematic exploitation of human cognitive vulnerabilities: algorithms favour sensationalist, divisive, or emotionally charged content because it generates more immediate reactions, transforming political debate into a spectacle of perpetual outrage. In this hybrid ecosystem, the logic of traditional media transmission merges with that of social sharing, creating a chaotic environment in which the distinction between verified information and

viral content gradually becomes blurred (Bentivegna & Boccia Artieri, 2019). This architecture favours the emergence of cognitive isolation phenomena. The extreme personalization of content, managed by predictive algorithms, tends to lock users into a 'tailor-made' information universe, invisibly curated to confirm their biases and maximize the time spent on the platform. This results in the 'filter bubble' effect: the algorithm acts as an invisible editor who, without the user's consent or awareness, removes from the feed the dissenting opinions and intellectual challenges necessary for healthy democratic debate (Pariser, 2011). Citizens are thus deprived of the cognitive friction that is fundamental to the development of critical thinking, finding themselves immersed in a reassuring but epistemologically poor stream, in which their beliefs are never questioned. Unlike traditional media, where the editorial line is evident, here the selection is hidden, creating the illusion of a comprehensive view of the world that is a narcissistic reflection of one's own previous preferences. The consequences of such fragmentation are devastating for social cohesion. When citizens lose access to a shared body of facts and arguments, the possibility of constructive debate vanishes. Segregation into homogeneous groups, or "echo chambers," triggers group polarization dynamics in which opinions tend to become radicalized simply because they are repeated and validated within a small circle. Without unscheduled exposure to different ideas, the connective tissue of the republic wears away, making political compromise impossible (Sunstein, 2017). Within these closed enclaves, democratic dialectic turns into tribal conflict: political opponents undergo a process of dehumanization or delegitimization, being described not as interlocutors with legitimate interests, but as an existential threat to the group's core values. This shift from disagreement to affective polarization erodes the very foundations of democratic tolerance. Democracy requires common ground; algorithms, on the contrary, are designed to segment and divide. A second, perhaps even more insidious level of risk concerns the opacity of decision-making systems. The algorithms that determine the visibility of a political leader or the dissemination of an election campaign issue function as veritable "black boxes." Not only is the code protected by trade secrets, but the very complexity of machine learning often makes it impossible to understand why some information takes priority over others. This lack of intelligibility creates a radical asymmetry of power: while platforms "see" and analyse every aspect of citizens' lives, citizens (and regulators) cannot see or understand the logic that governs them (Pasquale, 2015). This information asymmetry renders the democratic principle of transparency of power meaningless. When decisions about what is visible or censurable in public debate are delegated to computational architectures that are inaccessible even to state legislative bodies, a veritable privatization of civic space occurs. The risk is that behind the facade of mathematical objectivity lie discriminatory biases, systematic errors, or intentional manipulations that influence the outcome of electoral competitions without any possibility of public control or accountability. Blind reliance on big data can also turn these tools into "weapons of mathematical destruction." Algorithmic models, trained on historical data that reflects existing inequalities in society, tend to project these injustices into the future, automating discrimination. In the political arena, this can translate into micro-targeting campaigns that exclude specific demographic groups from the debate or exploit the psychological vulnerabilities of marginalized groups, amplifying disparities in influence and

voice (O'Neil, 2016). Psychographic profiling practices can, for example, be used not only to persuade, but also to strategically discourage specific minorities from voting, altering election results through invisible and targeted forms of manipulation that affect the most vulnerable segments of the population. Mathematical efficiency, therefore, is no guarantee of fairness; on the contrary, without ethical oversight, it risks crystallizing existing power hierarchies. Finally, the interaction between these mechanisms and the production of synthetic content threatens the epistemic integrity of society. In a context where virality rewards emotionality and indignation, factual truth loses its primacy as the currency of public discourse. We thus slip into the era of 'post-truth', where objective facts have less influence in shaping public opinion than appeals to emotions and personal beliefs. If AI can generate false evidence indistinguishable from reality (deepfakes) and spreading it in a targeted manner to those most likely to believe it, the very basis of institutional trust is undermined (McIntyre, 2018). Furthermore, the mere existence of technologies capable of generating perfect fakes introduces the phenomenon of the so-called 'liar's dividend': any politician caught red-handed can easily defend themselves by dismissing real and incriminating audio or video evidence as mere artificial fabrications, irreparably polluting the process of ascertaining responsibility and historical truth. Despite the gravity of these scenarios, it is necessary to avoid a paralyzing technological determinism. Users are not passive automatons and retain margins of action and resistance. There are spaces for political participation "outside the bubble," where digital social networks are used for civic mobilization that challenges algorithmic logic (Vaccari and Valeriani, 2021). Practices of 'data activism', advanced media literacy and the tactical use of the same digital tools to organise counter-narratives show that citizens can still reclaim public space by imposing a bottom-up use of technologies. However, awareness of these risks is a prerequisite for imagining possible corrective measures and opportunities, as we will see in the next section.

4. Towards an Enhanced Democracy: Artificial Sociality, Digital Citizenship, and New Frontiers of Institutional Responsibility

The integration of artificial intelligence (AI) into democratic processes and contemporary political communication represents an opportunity for transformation that goes far beyond simple automation, outlining a new "artificial sociality" capable of mediating the relationship between institutions and citizens in ways that were previously unimaginable. Far from being limited to mere algorithmic calculation, these technologies construct an appearance of social behaviour that stimulates humans to project social contexts and meanings onto machines, facilitating a normalization of digital interaction that can make political participation less alienating and more integrated into everyday life (Depounti and Natale, 2025). The shift from a purely administrative approach, typical of traditional e-government, to one of interactive algorithmic governance allows the state to abandon its stern and distant image. Citizens no longer experience public interaction as a tedious bureaucratic task, but as an ongoing dialogue, supported by interfaces that simulate empathetic closeness and uninterrupted availability. In this perspective, AI acts not only as a cold transmission tool, but as a means capable of deeply integrating into social life, bridging the communication gap that often

separates the complexity of the bureaucratic machine from citizens' need for more human and empathetic interaction (Natale, 2021). The adoption of instrumental bot identities, when placed within a framework of transparency, offers the possibility of designing strategically relevant and familiar communication flows that resonate with the social identities and values of specific groups that often feel excluded from traditional political discourse, thus promoting a more inclusive, widespread, and diverse public debate (Letourneau, 2025). In this way, linguistic minorities, segments of the population that are less digitally literate, or marginalized groups can find in conversational systems a channel for listening that is free from human physiological cognitive biases. The artificial interface, if programmed according to strict principles of fairness, does not judge or discriminate, offering a neutral and reassuring point of access to the democratic infrastructure. The ability to create systems that not only transmit data but actively participate in the "production of meaning" transforms artificial agents from simple message repeaters into true facilitators of public debate, enabling the dissemination of service information that adapts with surgical precision to the nuances of natural language and the needs of users (Veale and Cook, 2018). This evolution is strongly reflected in the adoption of advanced chatbots in the public sphere; borrowing the logic of efficiency and responsiveness typical of digital customer service, these tools can respond instantly to the informational and administrative needs of the population, drastically reducing waiting times and significantly improving the perception of an institution that is finally agile, responsive, and oriented toward serving citizens (Xu *et al*. 2017). This metamorphosis of the "face" of public administration is essential to stem rampant abstentionism and systemic mistrust: citizens who receive clear and timely responses to their daily needs are inherently more likely to renew their pact of loyalty with institutions. Another fundamental pillar of the opportunities offered by AI concerns privacy management and the building of a new bond of trust between rulers and ruled; analysis of institutional discourse suggests that, where technological innovation is accompanied by a political narrative that places individual protection and data security as cardinal values rather than bureaucratic obstacles, the way is paved for a more secure, aware, and willing digital citizenry to actively collaborate with the state (Zhang *et al*. 2024). Advanced analysis of semantic networks on social media also highlights how a deep understanding of opinion flows can be used to map community needs in real time, enabling policymakers to calibrate legislative interventions based on authentic linguistic and sentimental evidence, overcoming the limitations of traditional surveys (Yuan, Feng, and Danowski, 2013). This outlines the horizon of predictive and proactive policy: institutions no longer need to limit themselves to reacting to social emergencies, but can anticipate critical issues, acting preventively to defuse discontent before it radicalizes in the streets or at the polls. At the same time, the opportunity to combat disinformation lies not only in technical refutation after the fact, but in understanding the ontological narratives through which citizens construct their social distinction and sense of belonging. The use of AI to intercept these patterns allows for the structuring of much more effective information and awareness campaigns that enhance the cultural capital of the recipient instead of taking a confrontational or paternalistic attitude (Hall, Chadwick, and Vaccari, 2024). Understanding the deep mechanisms through which users respond emotionally and cognitively to distorted news allows us to design decision support systems that do not merely flag factual errors, but

guide users towards developing autonomous and proactive critical vigilance (Tandoc, Lim, and Ling, 2020). This approach is particularly valuable for younger generations who, while often expressing a lack of trust in traditional media and institutional hierarchies, show a marked ability to critically evaluate news when supported by digital tools that adapt to their rapid and hyperconnected information consumption tactics (Swart and Broersma, 2022). However, for this dynamic to be truly emancipatory, it is essential that technological implementation goes hand in hand with massive investments in algorithmic civic literacy. The goal of socialization agencies must be to train individuals to actively question AI outputs, transforming interfaces from tools of passive consumption to platforms for the defence of democracy. Furthermore, intelligent automation of communication processes offers the opportunity to manage consensus building transparently and at scale, enabling democratic institutions to overcome the physical limitations of popular consultation and include millions of voices in fluid, secure, and anonymous deliberative processes, while ensuring the integrity of the democratic process (Woolley, 2023). Pioneering experiences in civic technology demonstrate how machine learning can be used to map and synthesize hundreds of thousands of qualitative contributions in actual digital assemblies, extracting areas of cross-cutting consensus and neutralizing extreme polarization. In this way, the "wisdom of the crowd" becomes a formal and structured legislative contribution, no longer just chaotic background noise. In short, artificial intelligence acts as a formidable catalyst for civic empowerment which, if properly directed by a forward-looking political vision, can reinvigorate democracy by transforming the noise of big data into shared knowledge and common action. Through ethical personalization of the message, citizens are no longer perceived as anonymous numbers in an electoral database, but as active participants whose needs and desires are understood, respected, and anticipated by systems that operate strictly within the framework of "privacy by design." This symbiosis between human and artificial intelligence breaks down linguistic, cultural, and historical cognitive barriers, finally democratizing access to complex decisions and transforming algorithmic transparency from a cold technical obligation into a new and vital social contract of responsibility and participation. Ultimately, this promotes the transition from intermittent representative democracy to a "continuous" and relational democracy, in which the feedback loop between elected officials and voters does not end the day after the vote, but is fed daily through responsive digital ecosystems. The challenge of the future will be to use artificial sociality not to isolate individuals in their digital solipsism, but to connect them more deeply, consciously, and knowledgeably to the community, creating a hybrid public sphere in which computational efficiency is constantly at the service of human dignity, social equity, and common deliberation for the good of the polis. Only by embracing these technologies as strategic allies rather than threats to be contained can democracy evolve towards more agile, transparent, and inclusive models capable of resisting the pressures of post-truth and regenerating a solid and lasting bond of trust between political power and society.

5. Conclusions

Looking ahead, enthusiasm for the potential of so-called "augmented democracy" must

necessarily be balanced against a more structural assessment of the systemic risks that emerge when algorithmic mediation becomes not only widespread but infrastructural. The critical issue no longer concerns solely the intentional production of disinformation, but rather the transformation of the entire epistemic environment within which political judgment is formed. In this context, one of the most insidious phenomena is not merely the spread of false content, but the emergence of a condition of "reality apathy", in which an overabundance of synthetic, contradictory, and indistinguishable information leads to a gradual weakening of individuals' capacity for discernment (Vincent and Gismondi, 2021). This condition does not necessarily imply naive credulity, but rather a more radical form of cognitive withdrawal: citizens do not believe everything but progressively cease to believe in a selective and meaningful way. We thus move from a crisis of truth to a crisis of the relevance of truth itself, in which the cognitive cost required to verify information exceeds the perceived benefit of knowledge. These dynamic fosters a form of epistemic exhaustion that does not strengthen critical resilience, but instead tends to produce cynicism, disengagement, and, in extreme cases, withdrawal from public participation. The result is a weakening of the social contract not only at the informational level, but also at the motivational and normative levels, as trust in the very possibility of a shared common ground for democratic deliberation is eroded. In parallel, the technological infrastructure underpinning the new digital public sphere reveals increasing normative and cognitive opacity. Large language models (LLMs), far from being neutral tools of informational mediation, inevitably incorporate the asymmetries, distortions, and hierarchies embedded in their training data. As recent scholarship has shown, such systems do not merely reflect existing social biases, but may also amplify and stabilise them, contributing to the automated reproduction of forms of symbolic exclusion (Wyer and Black, 2025). The most significant risk therefore lies not only in explicit error, but in the invisible normalisation of discursive patterns that define what can be said, thought, or recognised as legitimate within the public sphere. If such systems are integrated into institutional communication processes without adequate critical safeguards, a form of cognitive delegation emerges that transforms the interaction between citizens and institutions into a relationship mediated by opaque systems that appear objective but are in fact normatively loaded. In this sense, automation does not concern only procedures, but also the production of political meaning itself, with the risk of entrenching a structurally invisible form of algorithmic exclusion that is therefore even more difficult to contest. A further dimension of concern emerges from the observation of digital platform dynamics as environments of symbolic conflict. Far from fostering rational deliberation, these ecosystems tend to incentivise high-emotion, identity-driven content, reinforcing logics of polarisation and antagonism. Online culture wars do not simply represent a degradation of public discourse, but rather the structural outcome of algorithmic architectures that reward confrontation over dialogue (Vowles and Hultman, 2021). In this scenario, the public sphere does not merely fragment into informational bubbles; it increasingly polarises into closed epistemic communities that no longer share common criteria for validating reality. Against this backdrop, the central challenge is no longer purely technological, but deeply political and philosophical. Future governance cannot be limited to innovation understood as the development of increasingly efficient tools; it must also include a sustained practice of ethical and institutional

maintenance of digital infrastructures of democracy. In this sense, the notion of ongoing ethical "care" for technology becomes central: the aim is not to perfect computational systems, but to preserve the conditions of possibility for democratic dialogue within technically mediated environments (Young and Coeckelbergh, 2024). This implies a fundamental rethinking of the relationship between citizenship and technology, in which political participation is not treated as an optimisable function or an engagement variable, but as a fragile, situated, and inherently non-automatable practice. From this perspective, democratic resilience does not consist in the elimination of uncertainty or error, but in the capacity to keep spaces of shared interpretation open, recognising the inevitably imperfect nature of human judgment as a constitutive element of political life. Ultimately, the survival of a democratic public sphere will not depend on the construction of perfectly transparent or fully verifiable information systems, but on the collective capacity to govern opacity without being overwhelmed by it. The task of institutions will not be to eliminate complexity, but to make it inhabitable; not to replace human judgment, but to protect the conditions under which it can be exercised. In this sense, the democracy of the future cannot be understood as an optimised decision-making system, but rather as a continuously unstable equilibrium between technology, trust, and shared responsibility.

References

Battista, D. (2023). Play, participate, decide: The emergence of video games as tools for political participation. *Journal of sociological research*, *15*, 1-14. https://doi.org/10.5296/jsr.v15i1.21360

Battista, D. (2024). Political communication in the age of artificial intelligence: an overview of deepfakes and their implications. *Society Register*, *8*(2), 7-24. https://doi.org/10.14746/sr.2024.8.2.01

Battista, D. (2025). From electronic voting to AI: lessons from the Estonian Case on the transformation of digital democracy. *Em Questão*, *31*, e-145465. https://doi.org/10.1590/1808-5245.31.145465

Bentivegna, S., & Boccia Artieri, G. (2019). *Le teorie delle comunicazioni di massa e la sfida digitale*. Laterza.

Bradshaw, S., & Howard, P. N. (2019). *The Global Disinformation Order: 2019 Global Inventory of Organised Social Media Manipulation*. Oxford Internet Institute.

Depounti, I., & Natale, S. (2025). Decoding Artificial Sociality: Technologies, Dynamics, and Implications. *Sage Journals*. https://doi.org/10.1177/14614448251359217

Dwivedi, Y. K., Hughes, L., Ismagilova, E., et al. (2021). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, *57*, 101994. https://doi.org/10.1016/j.ijinfomgt.2019.08.002

Floridi, L. (2015). *The Onlife Manifesto: Being Human in a Hyperconnected Era*. Springer. https://doi.org/10.1007/978-3-319-04093-6

Gallo, M., Fenza, G., & Battista, D. (2022). Information Disorder: What about global security implications?. *Rivista di Digital Politics*, *2*(3), 523-538.

García-Marzá, D., & Calvo, P. (2024). Algorithmic democracy. *A critical perspective based on deliberative democracy.* https://doi.org/10.1007/978-3-031-53015-9_12

Hall, N. A., Chadwick, A., & Vaccari, C. (2024). Online misinformation and everyday ontological narratives of social distinction. *Media, Culture & Society*, *46*(3), 572-590. https://doi.org/10.1177/01634437231211678

Letourneau, A. M. (2025). Instrumental bot identities and the politics of online discourse. *Sage Journals*. https://doi.org/10.1177/20563051251378233

Lippmann, W. (2017). *Public opinion*. Routledge. https://doi.org/10.4324/9781315127736

Manin, B. (1997). *The Principles of Representative Government*. Cambridge University Press. https://doi.org/10.1017/CBO9780511659935

McIntyre, L. (2018). *Post-Truth*. MIT Press. https://doi.org/10.7551/mitpress/11483.001.0001

Natale, S. (2021). *Deceitful Media: Artificial Intelligence and Social Life after the Turing Test*. Oxford University Press. https://doi.org/10.1093/oso/9780190080365.001.0001

O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown.

Pariser, E. (2011). *The Filter Bubble: What the Internet Is Hiding from You*. Penguin Press. https://doi.org/10.3139/9783446431164

Pasquale, F. (2015). *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard University Press. https://doi.org/10.4159/harvard.9780674736061

Rieder, B., & Hofmann, J. (2020). Towards platform observability. *Internet Policy Review*, *9*(4), 1-28. https://doi.org/10.14763/2020.4.1535

Sunstein, C. R. (2017). *Republic: Divided Democracy in the Age of Social Media*. Princeton University Press. https://doi.org/10.1515/9781400884711

Swart, J., & Broersma, M. (2022). The Trust Gap: Young People's Tactics for Assessing the Reliability of Political News. *The International Journal of Press/Politics*, *27*(2), 396-416. https://doi.org/10.1177/19401612211006696

Tandoc, E., Lim, D., & Ling, R. (2020). Diffusion of disinformation: How social media users respond to fake news and why. *Journalism*, *21*(3), 381-398. https://doi.org/10.1177/1464884919868325

Vaccari, C., & Valeriani, A. (2021). *Outside the Bubble: Social Media and Political Participation in Western Democracies*. Oxford University Press.

https://doi.org/10.1093/oso/9780190858476.001.0001

Van Dijck, J., Poell, T., & de Waal, M. (2018). *The Platform Society: Public Values in a Connective World*. Oxford University Press. https://doi.org/10.1093/oso/9780190889760.001.0001

Veale, T., & Cook, M. (2018). *Twitterbots: Making Machines that Make Meaning*. MIT Press. https://doi.org/10.7551/mitpress/10859.001.0001

Vincent, C. S., & Gismondi, A. (2021). Fake news, reality apathy, and the erosion of trust and authenticity in American politics. In *Democracy in the Disinformation Age*. Routledge. https://doi.org/10.4324/9781003008828-7

Vowles, K., & Hultman, M. (2021). Dead White men vs. Greta Thunberg: Nationalism, misogyny, and climate change denial in Swedish far-right digital media. *Australian Feminist Studies*, *36*(110). https://doi.org/10.1080/08164649.2022.2062669

Vreese, C. H. de (2023). AI and Political Communication. In *The Routledge Handbook of Political Communication*. Routledge.

Woolley, S. C. (2023). *Manufacturing Consent: Understanding Propaganda in the Era of Automation and Anonymity*. Yale University Press. https://doi.org/10.12987/yale/9780300251234.001.0001

Wyer, S., & Black, S. (2025). Algorithmic bias: sexual violence against women in GPT-3. *AI Ethics*, *5*. https://doi.org/10.1007/s43681-024-00641-0

Xu, A., Liu, Z., Guo, Y., Sinha, V., & Akkiraju, R. (2017). A new chatbot for customer service on social media. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3025453.3025496

Young, M. T., & Coeckelbergh, M. (2024). Keeping things running: maintenance and philosophy of technology. In *Maintenance and Philosophy of Technology*. Routledge. https://doi.org/10.4324/9781003316213

Yuan, E. J., Feng, M., & Danowski, J. A. (2013). "Privacy" in semantic networks on Chinese social media: The case of Sina Weibo. *Journal of Communication*, *63*(6), 1011-1031. https://doi.org/10.1111/jcom.12058

Zhang, R., Li, H., Chen, A., Liu, Z., & Lee, Y.-C. (2024). AI Privacy in Context: A Comparative Study of Public and Institutional Discourse on Conversational AI Privacy. *Sage Journals*. https://doi.org/10.31219/osf.io/qw7gx

Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. PublicAffairs.

Acknowledgments

Not applicable.

Authors contributions

Not applicable.

Funding

Not applicable.

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Informed consent

Obtained.

Ethics approval

The Publication Ethics Committee of the Macrothink Institute.

The journal's policies adhere to the Core Practices established by the Committee on Publication Ethics (COPE).

Provenance and peer review

Not commissioned; externally double-blind peer reviewed.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Data sharing statement

No additional data are available.

Open access

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.