Bialystok Legal Studies
Białostockie Studia Prawnicze
2026 vol. 31 no. 1

sciendo

DOI: 10.15290/bsp.2026.31.01.05

Received: 30.06.2025
Accepted: 08.01.2026

Ermanno Petrocchi
University of Macerata, Italy
e.petrocchi1@unimc.it
ORCID ID: 0009–0001–8298–3076

Letizia Conte
University of Macerata, Italy
l.conte1@unimc.it
ORCID ID: 0009–0006–8932–5515

Civic Engagement in the AI Age: The Role of Digital Activism in Fostering Democratic Technologies

Abstract: This paper examines the crucial role of activism and civil society in strengthening democratic values, particularly in the context of the contemporary challenges posed by artificial intelligence (AI). People's involvement through different forms of associative participation fosters accountability and social capital, essential for democratic resilience. However, the rise of AI poses risks, including manipulation through disinformation, deepfakes, and privacy violations, which threaten democratic processes and fundamental rights. In response to such threats, digital activism emerges as an essential tool, not only for mobilising people but also for steering technological development towards more inclusive and responsible forms. The article argues that digital activism should not be limited to awareness-raising or AI literacy initiatives, but rather should be recognised as a proactive force in the definition of an ethical AI consistent with democratic values. In this framework, participatory design represents a fundamental methodological tool; indeed, it aims at involving different stakeholders in AI design processes to create more responsible technologies. Therefore, through the active involvement of digital activism associations in AI decision-making processes, people can contribute to the democratic control of technologies and the safeguarding of fundamental freedoms. Just as civic engagement has historically strengthened democratic institutions, digital activism therefore has the potential to guide AI development towards fairer and more transparent and accountable technological futures.

© 2026 Letizia Conte, Ermanno Petrocchi published by Sciendo. This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.

Keywords: civil society, digital activism, democracy, AI ethics

Introduction

Civic activism and participation are among the key driving forces in the strengthening of democratic systems. Civil society, through non-governmental organisations, social movements, and other forms of associative participation, is a major player in ensuring the accountability of institutions and their capacity to respond effectively to people's needs. In particular, civic participation plays an important role in the formation of social capital, strengthening cohesion among community members and promoting skills essential for political engagement, such as information skills, tolerance, and a sense of belonging (Putnam, 1995).

In the current era, marked by technological acceleration and in particular the spread of artificial intelligence (AI), new and complex challenges for democracy have emerged. AI-based technologies can indeed be used to manipulate public opinion through disinformation, to compromise the integrity of electoral processes, and to jeopardise fundamental rights such as privacy and individual autonomy (Coeckelbergh, 2024). This paper argues that despite the challenges posed by the widespread use of AI technologies, it is precisely in this context that digital activism can emerge as a pivotal tool not only for the defence of democratic freedoms but also for the construction of a more just and inclusive technological ecosystem.

Among the various approaches to the responsible development of digital technologies, participatory design represents a particularly relevant perspective. This approach is based on the active involvement of stakeholders, such as experts, technologists, end users, and members of society in a broader sense, in the design and development processes of technologies.[1] The aim of participatory design is to overcome technocratic and top-down models, instead promoting collaborative practices that strive to co-create more inclusive, contextualised, and socially just solutions (Delgado et al., 2023). Participatory design is not limited to collecting post-hoc feedback, but rather requires continuous and direct involvement, paying particular attention to power inequalities and the technology's ethical and social impacts. Participatory design therefore emerges as a promising strategy to mitigate the risks of automated systems, improve transparency, and ensure the accountability of the actors involved (Delgado et al., 2023; Zytko et al., 2022).

In this context, this paper argues that digital activism, framed within the logic of participatory design, can constitute a privileged channel for people's involvement in AI systems' development processes. In this sense, activism represents not only a form of political mobilisation but a concrete way through which individuals can take an

1 We have decided to use alternative terms to 'citizens' in order to avoid excluding individuals who do not hold formal citizenship status.

active role in the co-creation of AI technologies. Similarly to how associative participation has historically strengthened traditional democracy, the involvement of civil society in the development of AI can help orient this technology towards democratic values (Brynjolfsson et al., 2025). The direct involvement of activists and experts in decision-making processes can enable them to shape AI technologies that are more inclusive and sensitive to their social impacts.

For these reasons, this paper proposes interpreting digital activism not simply as a form of mobilisation, protest, or denunciation, but as an active source of technological design, capable of contributing to the realisation of reliable, accountable, and democratically aligned AI systems. Our goal is to show how civic engagement through digital activism can be a concrete way to democratise AI and protect fundamental rights in a rapidly evolving technological context.

To this end, the paper is structured as follows: Section 1 analyses the role of civic participation in democratic societies, highlighting how activism has also evolved in response to technological changes and introducing the concept of digital activism. Section 2 discusses the main risks posed by AI to democratic systems. Section 3 delves into the participatory design approach, emphasising how digital activism can be integrated within this approach to empower citizens in shaping technological development. It also underscores the responsibility of democratic institutions in fostering such activism as a means of supporting participatory technology design. Section 4 offers some concluding remarks.

1. The role of active participation in shaping democratic societies

The transformative role of civil society associations and social movements has long attracted scholarly interest due to their ongoing adaptation to political and social change. Particular attention has been paid to the contribution that the participation of activists offers to the creation of conditions for authentic democracy, by fostering social solidarity, reinforcing civic bonds, and enhancing the effectiveness of public policy through multiple activities (Maloney & Van Deth, 2010).

Tocqueville already highlighted the importance of associationism for the life of American democracy, highlighting some of its characteristics and effects that are still considered valid (Edwards et al., 2001). He reflected in particular on how the American associations he studied guaranteed a fundamental pre-political socialisation, spreading the spirit of cooperation, solidarity, and civic engagement among the participants (Biorcio & Vitale, 2016). Many scholars have re-proposed his idea of associations as 'schools of democracy', valuing especially the role of political socialisation that associative participation can play.

However, in contemporary democracies, the relationship between associations and politics is much more complex and has changed significantly in recent decades.

The socialisation of members is no longer the sole function that associations serve within a participatory democracy; today, they function as collective actors in the public sphere to directly influence policies with various forms of initiatives and mobilisations, to put pressure on governments, but also to direct the production of public goods (Biorcio & Vitale, 2016).

In this sense, Putnam (1995), introducing the concept of 'social capital', highlighted the importance of associative participation as an essential condition for the proper functioning of democracy. This term refers to a set of elements such as relationships based on trust, the rules that govern collective life, and networks of civic associations. These factors contribute to making social organisation more efficient, favouring shared initiatives and decisions made collaboratively. The existence of a high level of social capital in a community should therefore trigger a virtuous circle, thanks to which it can grow further, guaranteeing the best conditions for democratic life.

Associative participation also increases the resources available for political participation: civic competence, information, membership in social networks, a sense of personal effectiveness, and the ability to act politically (Binder, 2020). Active participation fosters the acquisition of democratic decision-making methods, directing individuals towards dealing with issues of collective interest and promoting the construction of bonds based on trust and reciprocity. Engaging in socially relevant activities, with tangible and recognisable outcomes, helps to strengthen the self-efficacy and self-esteem of the participants. This process reflects a reworking of the value systems of activists, who tend to develop a greater awareness of and attention towards the political sphere, public policies, and institutional decisions (Maloney & Van Deth, 2010).

Since the 1980s, the rise of neoliberalism and globalisation has contributed to an erosion of faith in state intervention, leading to a reconfiguration of civil society's role. Indeed, civil society organisations have increasingly assumed functions traditionally carried out by political parties and institutions, such as mediating social demands and fostering political socialisation, especially among young people (Biorcio & Vitale, 2016).

In addition, the technological revolution – from the emergence of social media to the development of AI – has deeply transformed the landscape of activism, reshaping its tools, strategic approaches, and associated risks. This constantly evolving context even makes it difficult to precisely delineate the conceptual boundaries of emerging digital activism, as the rapid transformations taking place make its definition continuously provisional and subject to redefinition (Özkula, 2021a).

In this scenario, activist associations do not limit themselves to using digital technologies as operational tools, but rather take on a crucial role in monitoring, reporting, raising awareness of, and disseminating information about the social and political effects of technological innovations. These organisations, however, are themselves exposed to critical issues deriving from the use of these tools, including 'invisibilisa-

tion' phenomena, i.e. the reduced visibility of politically sensitive or dissenting content within digital platforms, and surveillance practices, namely the systematic monitoring, data collection, and profiling of activists by the state or private companies (Dencik et al., 2016; Etter & Albu, 2021; Lane et al., 2017; Özkula, 2021a). In the following section, the distinctive features of contemporary digital activism and the specific functions performed by modern social capital in this new digital ecosystem will be explored.

1.1. Activism in the digital age: Opportunities and limits

The rise of social media and digital technologies over the past decades has significantly reshaped the landscape of civic engagement and activism (Bennett & Segerberg, 2012). Civil society actors are increasingly using digital functionalities to articulate political claims, reach broader audiences, and challenge dominant narratives. Indeed, these tools have expanded the communicative and organisational capacities of social movements, enabling wider participation and more agile mobilisation strategies, as well as the diversification of voices within the public sphere (Castillo Esparcia et al., 2023). However, this transformation has also prompted critical reflections on the efficacy and nature of this new form of digital activism in a lot of scholarly literature in recent decades (Helmond, 2015; Mora, 2014; Özkula, 2021a, 2021b, 2021c).

In general terms, digital activism can be defined as a form of political activism that develops through the internet (McCaughey & Ayers, 2003). Its defining actions mirror traditional practices of activism (such as petitions and protests) while also employing digital technologies to support or organise offline activities, such as the promotion of events through social media. Over the past two decades, this type of activism has fostered high levels of interaction and interconnection, for instance through tweets, posts, chats, and sharing of content, especially user-generated content, across geographical and institutional boundaries (Özkula, 2021b).

However, the terminology associated with this phenomenon is ambiguous, reflecting the constant evolution of digital technologies. Indeed, there are many alternative expressions according to the tools used: among these are 'online activism', 'social media activism', 'internet activism', and 'hashtag activism' (Mora, 2014). In this context, the term 'digital activism' is configured as an umbrella term, as it encompasses these different forms and aligns with recent linguistic trends that favour the use of the term 'digital' over 'online', in recognition of the increasingly pervasive processes of advanced digitisation (Özkula, 2021b, 2021c).

Özkula (2021b) distinguishes practices of digital activism in five main categories: (i) advocacy and political commentary, (ii) recruitment and movement building, (iii) organisation and coordination, (iv) online direct action, hacktivism, and civil disobedience, and (v) research and documentation. The latter category in particular can be fundamental strategic tools for activist action, as they are functional to the dissemination of information relating, for example, to human rights violations. These activities can be carried out directly, e.g. through the use of mobile phones, or conveyed through

traditional media (McCaughey & Ayers, 2003). This also includes practices such as the disclosure of confidential information, election observation, countering disinformation and electoral fraud, as well as forms of 'sousveillance', i.e. bottom-up surveillance practices, in which citizens document and publicise abusive behaviour by state institutions, such as police forces or government authorities. Although it is a matter of debate whether such documentation activities can be considered forms of activism in their own right, they are often carried out in support of activist action or with the explicit intention of expressing dissent, as in the emblematic case of Wikileaks (Özkula, 2021b).

As we will see in Section 4, this is one of the most recurrent strategies in the cases of civil society organisations that constantly carry out studies, monitoring, denunciation, and awareness work regarding the political and social risks of the use of AI tools.

Although digital activism offers numerous opportunities for mobilisation and participation, it is not immune to the risk of incurring the same anti-democratic logic of control and limitation against which it often stands in opposition. In particular, the activism conveyed through social media platforms is strongly influenced by the dynamics of visibility imposed by the techno-commercial infrastructures that regulate their operation. As highlighted by Helmond (2015), the increasing platformisation of digital public spheres has concentrated decision-making power over content circulation in the hands of private actors, whose moderation mechanisms are frequently opaque and lack effective accountability.

As a result, activists are forced to operate within digital spaces increasingly determined by non-transparent algorithmic logics, exposing themselves to the risk of the invisibility of their messages and communicative precariousness due to the automatic classification of content (Etter & Albu, 2021). The surveillance of activists by platforms represents, in this context, a crucial dimension to be analysed, given the centrality assumed by these tools in the configuration of contemporary protests. Facebook, for instance, is configured as a primary hub for the communication, mobilisation, and organisation of activist networks. Nevertheless, activist engagement on Facebook generates digital traces, and the platform's expansive operational reach intensifies the vulnerabilities linked to digital protest, amplifying the harmful consequences of online surveillance. This form of surveillance can inhibit free expression and lead to the unequal restriction of civil rights (Nurik, 2022). This case highlights how the evolution of technologies, especially those that are AI-based, has significant implications that require careful critical analysis.

2. How AI undermines democracy

The massive adoption of AI, and more recently generative AI, in the public and political space entails several risks that go far beyond purely technical concerns. These systems are not mere technical tools but devices that mediate access to knowl-

edge and help structure the conditions of democratic debate. As Coeckelbergh (2024) and Mentxaka et al. (2025) have pointed out, the political risks of large-scale adoption of AI systems are multiple and profound.

One of the most immediate dangers lies in the ability of generative AI systems to produce fake content. Such content can take the form of deepfakes, i.e. artificial videos capable of simulating faces, voices, and movements in an extremely realistic manner, or hallucinations, i.e. errors typical of large language models (LLMs) that generate seemingly plausible statements that lack any real substance (Huang et al., 2025; Westerlund, 2019). These phenomena, which may be accidental or deliberately orchestrated, fuel the production of fake news, targeted attacks, and political disinformation campaigns (Schick, 2020). The speed and scale with which such content is generated and disseminated pose a direct threat to the integrity of public discourse and citizens' capacity to form opinions based on verifiable facts (Coeckelbergh, 2025a).

For democracy to function effectively, it presupposes that citizens can exercise a minimum of cognitive autonomy, the ability to critically evaluate information and make independent judgements. However, increasing exposure to misleading content and uncritical interaction with AI systems undermines this capacity. As pointed out by Rini (2020), Fallis (2021), and Coeckelbergh (2025b), one of the main risks of the use of generative AI systems is that of the progressive weakening of individual epistemic agency, with direct consequences on the quality of civic participation and democratic deliberation.

In parallel, the use of AI in social media and search engines encourages the formation of epistemic bubbles and echo chambers, closed digital environments in which users are exposed almost exclusively to content that confirms their pre-existing beliefs (Nguyen, 2020). This hinders encounters with divergent opinions and inhibits critical confrontation, fuelling a polarised worldview (Croce & Piazza, 2022). Under such conditions, dissent turns into hostility, dialogue gives way to monologue, and democratic cohesion disintegrates.

A further risk, less obvious but equally dangerous, is the self-referential repetition of knowledge. When LLMs are trained on content generated by other models, a closed circuit is created in which information reproduces itself without innovating. This phenomenon, referred to as 'epistemic incest', undermines society's ability to receive new signals, elaborate answers to emerging problems, or include tacit or uncodified forms of knowledge (Coeckelbergh, 2025a).

In brief, the uncontrolled and opaque use of AI systems not only jeopardises the quality of information but also profoundly affects the very structure of democratic deliberation. In the absence of regulation, critical education, and transparent digital infrastructures, AI risks becoming an instrument of cognitive fragmentation and the erosion of civic bonding, rather than being an ally of democratic progress.

3. Reframing participation: Digital activism, democratic values, and the role of participatory design in the AI era

The democratisation of AI cannot be delegated solely to experts but requires inclusive participatory processes capable of redistributing technological power and fostering systems that are fairer, more transparent, and responsive to collective needs. Individuals should be empowered with the capacity to engage in informed deliberation and exert influence over AI development and use in ways that are more responsive to collective needs.

In contemporary digitised society, participation has emerged as a widespread promise – manifested in practices ranging from crowdfunding to user-generated content on social media – often seen as a driver of democratic innovation. Yet this promise coexists with a contradictory trend: the increasing centralisation of capital, infrastructure, and data through large-scale platforms, which undermines core democratic ideals such as civic participation, data ownership, and privacy (Smith et al., 2017).

In a context where digital infrastructures, in particular those underpinned by AI, are increasingly governed by opaque logics, it becomes essential to provide users with critical literacy and awareness of the associated risks. Several civil society organisations have begun to play a crucial role in this regard, offering tools for public education, conducting oversight, and fostering democratic scrutiny. These initiatives reflect a broader effort to embed values such as equity, accountability, and transparency in the development and deployment of digital technologies (Milan, 2015).

Multiple perspectives on participation show its evolving relevance across contemporary contexts, highlighting how the democratic ideals underlying participatory design continue to serve as effective tools for critical engagement.

Participatory design, with its emphasis on values, aspirations, and the sociopolitical contexts in which technologies are developed, provides depth and direction often overlooked in the dominant discourse, which is focused primarily on innovation or efficiency (Smith et al., 2017).

This paper contends that activist networks – already engaged in monitoring, raising awareness, and pressuring for ethical technological practices – should be included as key actors within participatory processes (Tsai & Pentland, 2025). These networks provide valuable perspectives both in influencing public policy and in guiding users towards more informed and responsible use of AI-based tools. Examples like Algorithm Watch, which will be analysed in the following section, demonstrate how activist organisations are strategically mobilising various tools to challenge the unchecked expansion of biased algorithms and to contest algorithmic injustices.

A key strategy for achieving this is participatory design, which promotes the inclusion of different actors in AI design and implementation (Delgado et al., 2023; Zytko et al., 2022). Originating from Scandinavian labour movements in the 1970s, it has

evolved into a critical methodology for ensuring that technological systems reflect the values, needs, and experiences of different user communities (Bødker et al., 2004).[2]

In the context of AI, the involvement of activist organisations, particularly those addressing the concerns of historically marginalised or underrepresented communities, in co-design processes can mitigate algorithmic harms and promote socially just outcomes, particularly by tackling epistemic injustices embedded in data practices (D'Ignazio & Klein, 2020; Overton, 2025). As such, participatory design constitutes not only a technique but also an epistemic and political intervention, grounded in subaltern knowledge, local worldviews, and situated experiences, as emphasised also by decolonial and feminist theories of AI ethics (Mohamed et al., 2020; Suchman, 2002). Through these efforts, participatory design contributes to building more pluralistic, just, and democratically aligned technological futures.

Drawing on these opportunities, this article proposes that digital activism should not be limited to raising awareness about the risks of AI or enhancing public literacy; rather, it should be recognised as a central mechanism for fostering inclusive and justice-oriented AI development (Castillo Esparcia et al., 2023; Özkula, 2021c). By integrating activist participation into AI decision-making processes, democratic engagement can be enhanced and technological systems can be aligned more closely with fundamental rights and collective values. In this light, digital activism, when reformulated through the lens of participatory design, emerges as a transformative force in the democratic governance of algorithmic systems.

Nonetheless, incorporating justice-oriented participation into AI development presents significant challenges, predominantly within capitalist frameworks that prioritise scalability and efficiency (Sloane et al., 2020). As machine-learning systems expand, they often lose the contextual sensitivity achieved through previous participatory efforts (Boyd & Crawford, 2012; Selbst et al., 2019). Despite these constraints, participatory design, anchored in principles of justice, transparency, and accountability, remains a vital tool aligned with the aims of digital activism, offering a pathway towards more equitable and democratic technological ecosystems.

Indeed, this type of activism represents a privileged channel through which members of society can be included in technology co-design processes. After all, without a supportive association or network, it is extremely difficult for the individual to find a way to speak out about technology, bring forward their demands, and thus be considered from the early stages of AI development processes. Digital activism in this sense takes the form of an accessible and collective way of participating, which allows even those who do not work in the tech sector to shape the technological tools they use in their everyday life.

2 The reference to the Scandinavian labour model lends substantial support to our argument that democratic activists and affiliated organisations can mitigate the societal risks and abuses associated with AI.

4. Institutions' responsibilities

In the design of AI technologies that increasingly permeate everyday life, public institutions can no longer confine themselves to ex-post regulation or the passive adoption of technical tools. Instead, they must assume an active role in promoting participatory practices that engage members of society in the definition, oversight, and evaluation of digital infrastructures (Coeugnet et al., 2023).

As previously discussed, digital activism emerges as a crucial actor in this regard. Recognising the civic value of digital activism, however, is not sufficient in itself (Stamboliev, 2023); what is needed is a deliberate institutional effort to create the conditions under which this form of participation can flourish and exert real influence on design and governance processes. In this regard, two fundamental responsibilities for institutions follow from these considerations: the first is to build an institutional ecosystem in which digital activism associations are recognised as legitimate interlocutors and included in the design, implementation, and monitoring processes of AI systems. The second, more general, is to actively encourage the spread of digital activism. Digital activism in fact constitutes not only a channel for citizens to participate in the co-creation of technologies but also assumes a more educative role, i.e. that of a driver for the development of the skills needed to critically understand and manage digital technologies. Indeed, in light of the risks to the democratic system highlighted in the previous sections, there is a clear need to promote greater public awareness of how AI works, the skills required to consciously interact with such systems, and the tools available to defend against asymmetries of digital power. This is the only way to strengthen democratic resilience and ensure informed, inclusive, and active participation in today's algorithmic society.

Institutions can therefore adopt some concrete strategies to support this dual role of digital activism and promote a more participatory and democratic technological design:

Involve civic actors in decision-making processes on AI, algorithms, and data use (Borchers, 2024). This means including associations of digital activism in committees, public forums, and regulatory bodies where guidelines, standards, and regulations related to the use of AI technologies are defined. This type of involvement favours pluralist deliberation, in which technical knowledge is confronted with the ethical, social, and territorial needs expressed by the community.

Open co-design bodies to digital activists (Feltrero & Osuna-Acedo, 2023). Public technologies, such as digital platforms, surveillance systems, and predictive intelligence tools, should be designed through structured participatory processes that include the different voices of the community. These bodies can function as hybrid spaces where user experiences are shared, concrete needs are defined, ethical implications are assessed, and sustainable, accessible, and inclusive solutions are co-designed.

Fund educational projects and awareness-raising campaigns (Hsu et al., 2022; Stamboliev, 2023). Critical digital literacy is a prerequisite for effective participation; institutions can support digital activism associations to promote educational programmes, public workshops, information toolkits, and social campaigns to help people understand how algorithms work, what data is used, and what the risks and opportunities of AI are, not just providing technical skills but educating aware individuals capable of influencing the public debate.

Favour bottom-up initiatives to carry out algorithmic audits and ethical and social impact assessments of public technologies. Administrations can enter into partnerships with independent organisations and digital activists to conduct transparent audits of the functioning of algorithmic systems in use by public bodies or outsourced services. Such collaborations can make it possible to identify bias, discriminatory effects, privacy risks, and indirect impacts on social cohesion. Participatory auditing can turn into a tool of collective accountability, capable of giving citizens back an active role in the surveillance of technological power.

These actions, if implemented consistently and continuously, can help transform digital activism from a marginal phenomenon to a structural component of technological democracy, ensuring in this way that decisions on the digital future of societies are addressed in an open, accountable, and inclusive manner (Spielkamp, 2017).

Examples already exist of the educative function of digital activism. One is Algorithm Watch, an independent non-profit organisation engaged in the monitoring, analysis, and public communication of automated decision-making systems. Algorithm Watch exposes the opacity and discriminatory risks inherent in algorithms used in the public and private spheres, promotes the adoption of ethical, transparent, and fair practices by institutions, and offers information, training, and support tools for citizens and decision-makers (Stark et al., 2020). By collaborating with academic bodies, media, and policymakers, this organisation demonstrates how digital activism can act as a mediator between technology, society, and politics, facilitating a more open and knowledgeable dialogue on issues of automation and algorithmic governance (Cornils, 2020). Algorithm Watch thus embodies the role of activism as an essential civic infrastructure, capable of providing critical expertise and building civic surveillance networks (Loi, 2020).

However, this paper argues that the function of digital activism, well represented by experiences such as that of Algorithm Watch, should not be exhausted in the role of critical observer or facilitator of public awareness. On the contrary, this function can and must be extended to processes of technological co-creation, actively involving civic associations and activist networks in the design phase of the technologies themselves. This marks a paradigm shift, from viewing activism as a downstream watchdog to recognising it as an upstream co-designer. Involving activist organisations from the earliest stages of the technology life cycle means redistributing decision-making power and ensuring that automated systems are not designed solely by

technical experts or industry stakeholders but also by those who represent the public interest, digital rights, and social justice.

From this perspective, digital activism emerges as an indispensable actor in the construction of a truly democratic technological ecosystem, capable not only of denouncing distortions but also of proposing design alternatives oriented towards the common good.

Conclusions

This article has explored how digital activism and participatory design can serve as crucial frameworks to guide the development of AI technologies towards democratic values, social justice, and civic empowerment. Indeed, AI architectures, datasets, and results are deeply rooted in the existing power structures and social biases that arise from historical systems of oppression (Noble, 2018). As these systems become increasingly integrated into governance, the risks they pose to democratic accountability and transparency are amplified significantly. In this context, the defence of democratic principles requires not only institutional reform but also encompasses the technological infrastructures that shape decision-making itself.

Digital activism has emerged as a vital force in this challenge, committing to exposing algorithmic injustice and calling for participatory governance structures that prioritise the voices of marginalised communities (Benjamin, 2019). In this regard, participatory design offers a path to institutionalise democratic engagement in AI design and implementation (Costanza-Chock, 2020).

The paper has argued that if institutions engaged with activist organisations involved in monitoring and raising awareness about the risks of AI, such as Algorithm Watch, from the earliest stages of the technology life cycle, there would be a stronger assurance that automated systems reflect the public interest. This is due to the enhanced capacity of digital activists to connect directly with individuals and represent diverse societal concerns.

By advancing values such as epistemic justice through the recognition of different forms of knowledge and the redistribution of power, these approaches challenge the technocratic rationalities that often dominate AI discourse and practice (D'Ignazio & Klein, 2020).

However, some limitations exist: firstly, participatory practices, while promising, require significant resources, time, and institutional commitment. These conditions are often at odds with the speed and priorities of the tech industry (Boyd & Crawford, 2012). In addition, authentic participation also requires a revisiting of skills and authority, challenging entrenched hierarchies within both the technology and governance sectors.

Despite these limitations, the approaches described stimulate an important reflection on what posture to take with respect to the challenges that AI poses to democratic values. Claiming AI as a democratic space is not just about mitigating damage; it is a matter of expanding the realm of possibilities, of imagining and building technological futures that are just, pluralistic, and respectful of human dignity. Digital activism thus offers a critique of the current trajectory of AI and stands as a model for its democratic transformation, which not only takes into account multiple voices but also works to ensure that individuals can participate equally, with equal respect, in the fight against injustice.

REFERENCES

Benjamin, R. (2019). *Race after technology: Abolitionist tools for the New Jim Code*. Polity.

Bennett, W. L., & Segerberg, A. (2012). The logic of connective action. *Information Communication & Society*, *15*(5), 739–768.

Binder, M. (2020). Enhancing democracy: Can civic engagement foster political participation? *Social Science Quarterly*, *102*(1), 47–68.

Biorcio, R., & Vitale, T. (2016). *Italia Civile: Associazionismo, partecipazione e politica*. Donzelli Editore.

Bødker, K., Kensing, F., & Simonsen, J. (2004). *Participatory IT design*. The MIT Press.

Borchers, M., Gierlich-Joas, M., Tavanapour, N., & Bittner, E. (2024, May). Let citizens speak up: Designing intelligent online participation for urban planning. In Mandviwalla, M., Söllner, M., & Tuunanen, T., (Eds.), *International conference on design science research in information systems and technology* (pp. 18–32). Springer Nature Switzerland.

Boyd, D., & Crawford, K. (2012). Critical questions for big data. *Information Communication & Society*, *15*(5), 662–679.

Brynjolfsson, E., Pentland, A., Persily, N., Condoleezza, R., & Aristidou, A. (2025). Introduction: Artificial intelligence and democracy in America, *The Digitalist Papers: Artificial Intelligence and Democracy in America*, https://www.digitalistpapers.com/essays/introduction.

Castillo Esparcia, A., Smolak-Lozano, E., & Llorent Vázquez, V. (2023). Artificial intelligence and citizen communication: Risk awareness, digital literacy, and civic participation. *Communication & Society*, *36*(1), 1–17.

Coeckelbergh, M. (2024). *Why AI undermines democracy and what to do about it*. John Wiley & Sons.

Coeckelbergh, M. (2025a). LLMs, truth, and democracy: An overview of risks. *Science and Engineering Ethics*, *31*(1), 1–13.

Coeckelbergh, M. (2025b). AI and epistemic agency: How AI influences belief revision and its normative implications. *Social Epistemology*, 40(1), 1–13.

Coeugnet, P., Labatut, J., Duval, J., & Vourc'h, G. (2023). Including citizens through co-design in a participatory research project to explore innovative agro-food systems: The case of future dairy livestock systems. *Frontiers in Sustainable Food Systems*, *7*, 1098295.

Cornils, M. (2020). Designing platform governance: A normative perspective on needs, strategies, and tools to regulate intermediaries. *Algorithm Watch*, *26, 1–88*.

Costanza-Chock, S. (2020). *Design justice: Community-led practices to build the worlds we need*. MIT Press eBooks.

Croce, M., & Piazza, T. (2022). *Che cosa sono le fake news*. Carocci Editore.

Delgado, F., Yang, S., Madaio, M., & Yang, Q. (2023). The Participatory Turn in AI Design: Theoretical Foundations and the Current State of Practice. In Association for Computing Machinery (Ed.), *Proceedings of the 3rd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization*, 37, 1–23. Association for Computing Machinery. https://doi.org/10.1145/3617694.3623261.

D'Ignazio, C., & Klein, L. F. (2020). *Data feminism*. MIT Press eBooks.

Dencik, L., Hintz, A., & Cable, J. (2016). Towards data justice? The ambiguity of antisurveillance resistance in political activism. *Big Data & Society*, *3*(2), 1–12.

Edwards, B., Foley, M., & Diani, M. (2001). Civil society and political context. In B. Edwards, M. Foley, & M. Diani (Eds.), *Beyond Tocqueville: Civil society and the social capital debate in comparative perspective* (pp. 17–31). University Press of New England.

Etter, M., & Albu, O. B. (2021). Activists in the dark: Social media algorithms and collective action in two social movement organizations. *Organization*, *28*(1), 68–91.

Fallis, D. (2021). The epistemic threat of deepfakes. *Philosophy & Technology*, *34*(4), 623–643.

Feltrero, R., & Osuna-Acedo, S. (2023). Social innovation on educational AI developments: A case study on social participation on designing AI generative models for diversity. In Kubincová, Z., Hao, T., Capuano, N., Temperini, M., Ge, S., Mu, Y., Fantozzi, P., & Yang, J. (Eds.), *International symposium on emerging technologies for education* (pp. 16–26). Singapore: Springer Nature Singapore.

Helmond, A. (2015). The platformization of the web: Making web data platform ready. *Social Media + Society*, *1*(2), 1–11.

Hsu, Y. C., Verma, H., Mauri, A., Nourbakhsh, I., & Bozzon, A. (2022). Empowering local communities using artificial intelligence. *Patterns*, *3*(3), 1–7.

Huang, L., Yu, W., Ma, W., Zhong, W., Feng, Z., Wang, H., Chen, Q., Peng, W., Feng, X., Qin, B., & Liu, T. (2025). A survey on hallucination in large language models: Principles, taxonomy, challenges, and open questions. *ACM Transactions on Information Systems*, *43*(2), 1–55.

Lane, D. S., Kim, D. H., Lee, S. S., Weeks, B. E., & Kwak, N. (2017). From online disagreement to offline action: How diverse motivations for using social media can increase political information sharing and catalyze offline political participation. *Social Media + Society*, *3*(3), 1–14.

Loi, M. (2020). People analytics must benefit the people: An ethical analysis of data-driven algorithmic systems in human resources management. *AlgorithmWatch, 1–56*.

Maloney, W. A., & Van Deth, J. W. (2010). *Civil society and activism in Europe*, Routledge.

McCaughey, M., & Ayers, M. D. (2003). *Cyberactivism*. Routledge eBooks.

Mentxaka, O., Díaz-Rodríguez, N., Coeckelbergh, M., de Prado, M. L., Gómez, E., Llorca, D. F., Herrera-Viedma, E., & Herrera, F. (2025). Aligning trustworthy AI with democracy: A dual taxonomy of opportunities and risks. *arXiv preprint arXiv:2505.13565*.

Milan, S. (2015). From social movements to cloud protesting: the evolution of collective identity. *Information Communication & Society*, *18*(8), 887–900.

Mohamed, S., Png, M., & Isaac, W. (2020). Decolonial AI: Decolonial theory as sociotechnical foresight in artificial intelligence. *Philosophy & Technology*, *33*(4), 659–684.

Mora, F. A. (2014). Emergent digital activism: The generational/technological connection. *Journal of Community Informatics*, *10*(1), 1–13.

Nguyen, C. T. (2020). Echo chambers and epistemic bubbles. *Episteme*, *17*(2), 141–161.

Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. New York University Press.

Nurik, C. L. (2022). Facebook and the surveillance assemblage: Policing Black Lives Matter activists & suppressing dissent. *Surveillance & Society*, *20*(1), 30–46.

Overton, S. (2025). Overcoming racial harms to democracy from artificial intelligence. *Iowa Law Review*, 110, 805–866.

Özkula, S. M. (2021a). The unmaking of collective action: Changing organizing logics in civil society organizations through social media activism culture. *International Journal of Communication*, *15* (19), 1984–2002.

Özkula, S. M. (2021b). The problem of history in digital activism: Ideological narratives in digital activism literature. *Journal of Digital Social Research,* 3(3), 60–84.

Özkula, S. M. (2021c). What is digital activism anyway? Social constructions of the 'digital' in contemporary activism. *Journal of Digital Social Research*, *3*(3), 60–84.

Putnam, R. D. (1995). Bowling alone: America's declining social capital. *Journal of Democracy*, *6*(1), 65–78.

Rini, R. (2020). Deepfakes and the epistemic backstop. *Philosophers' Imprint*, *20*(24), 1–16.

Schick, N. (2020). *Deep fakes and the infocalypse*. Octopus Books.

Selbst, A. D., Andrew, L., & Narayanan, A. (2019). Fairness in machine learning: A survey. *Communications of the ACM*, *62*(2), 56–65.

Sloane, M., Moss, E., Awomolo, O., & Forlano, L. (2020). Participation is not a design fix for machine learning. *arXiv.org*, https://doi.org/10.1145/3551624.3555285.

Smith, R. C., Bossen, C., & Kanstrup, A. M. (2017). Participatory design in an era of participation. *CoDesign*, *13*(2), 65–69.

Spielkamp, M. (2017). AlgorithmWatch: What role can a watchdog organization play in ensuring algorithmic accountability? *Transparent Data Mining for Big and Small Data*, 32, 207–215.

Stamboliev, E. (2023). Proposing a postcritical AI literacy: Why we should worry less about algorithmic transparency and more about citizen empowerment. *Media Theory*, *7*(1), 202–232.

Stark, B., Stegmann, D., Magin, M., & Jürgens, P. (2020). Are algorithms a threat to democracy? The rise of intermediaries: A challenge for public discourse. *AlgorithmWatch*, *26*, https://algorithmwatch.org/en/wp-content/uploads/2020/05/Governing-Platforms-communications-study-Stark-May-2020-AlgorithmWatch.pdf.

Suchman, L. (2002). Located accountabilities in technology production. *Scandinavian Journal of Information Systems*, *14*(2), 91–105.

Tsai, L., & Pentland, A. (2025). Rediscovering the pleasures of pluralism: The potential of digitally mediated civic participation. *UCLA Journal of Law and Technology*, 30, 179–195.

Westerlund, M. (2019). The emergence of deepfake technology: A review. *Technology Innovation Management Review*, *9*(11), 40–53.

Zytko, D., Wisniewski, P., Guha, S., Baumer, E., & Lee, M. (2022). Participatory Design of AI Systems: Opportunities and Challenges Across Diverse Users, Relationships, and Application Domains. *Conference: CHI '22: CHI Conference on Human Factors in Computing Systems,* 154, 1–4. https://doi.org/10.1145/3491101.3516506.