

TWON Policy Brief #3

On Shaping Online Social Networks for a Democratic Digital Public Sphere

March 2026

What does digital sovereignty mean in practice when a handful of global platforms structure Europe's public sphere? How can the European Union ensure that online social networks (OSNs) operate in line with democratic values, fundamental rights, and the protection of minors? And how can publicly funded research support policy-makers in shaping and enforcing a distinct European model of platform governance? These questions have become increasingly urgent in light of geopolitical tensions, systemic disinformation, and rapid advances in generative AI. Allegations of foreign interference in democratic processes illustrate how digital platforms can affect electoral integrity and public trust, such as in the 2024 presidential election in Romania, where actors linked to Russia were suspected of leveraging TikTok.

The EU has adopted an ambitious regulatory framework, notably the Digital Services Act (DSA) and the Digital Markets Act (DMA). Yet effective enforcement, meaningful transparency, and robust research access remain key challenges. A small number of predominantly non-European, profit-driven companies control the infrastructures that shape information flows and public debate. These platforms are not neutral intermediaries: their opaque algorithms and engagement-optimized business models influence which voices are amplified and which are marginalized. Algorithmic ranking systems often privilege emotional and sensational content, thereby reinforcing affective polarization. Meanwhile, the attention economy fosters information overload, addictive usage patterns, and psychological harm. The integration of AI tools has further lowered the barriers for producing deepfakes and coordinated propaganda. When infrastructures are designed primarily to maximize retention and advertising revenue, democratic considerations run the risk of being subordinated to commercial incentives.

At the same time, online social networks offer significant democratic opportunities. Discussions in the TWON Citizen Labs across Europe highlighted their capacity to coordinate civic initiatives, mobilize support, and connect marginalized communities. OSNs can increase political participation, foster transnational solidarity, and provide low-threshold access to knowledge. They can strengthen visibility for diverse perspectives and create new forms of accountability and collective memory. If appropriately governed and designed, digital platforms can expand democratic agency.

In response to these structural challenges and opportunities, this policy brief advances a coherent strategy for strengthening European digital sovereignty and democratic resilience. It calls for the effective enforcement of existing EU digital laws, robust and standardized access to platform data for research purposes, and greater transparency and accountability of algorithms. At the same time, it emphasizes the need for safety-by-design standards,

support for European public-interest-oriented platform alternatives, and stronger investment in media literacy, independent journalism, and evidence-based policymaking. Taken together, these recommendations aim to shift Europe toward a proactive model of digital governance, actively shaping digital infrastructures aligned with democratic values. The policy brief builds on the empirical findings of TWON and the insights of our participatory formats, including four Citizen Labs and a Policy Hackathon. The recommendations were created and refined in an iterative process between TWON researchers, policymakers, domain experts, and citizens to ensure their relevance and comprehensiveness. We sincerely thank all participants involved for their valuable contributions.

The TWON project examined how the design of online platforms influences the quality of online democratic discourse. At its core, the interdisciplinary team developed an innovative “digital twin” approach. Rather than experimenting on real users, simulations model how different platform architectures and ranking algorithms influence the quality of online debate and exposure to harmful content. By translating these findings into policy recommendations and discussing them in participatory Citizen Labs across Europe, TWON contributes to evidence-based policymaking and digital citizenship. The consortium includes leading European research institutions, among them the University of Amsterdam, the Karlsruhe Institute of Technology (KIT), the University of Belgrade, the University of Trier, the FZI Research Center for Information Technology, the Jožef Stefan Institute, the Slovenian Press Agency (STA), the Robert Koch Institute (RKI), and DialoguePerspectives e.V.

Policy Recommendations

1. Support Research on Online Social Networks to Increase Transparency and Accountability

a. Create a European Research Infrastructure for Safe Online Social Network Research

Despite DSA Art. 40, obtaining data from platforms remains challenging. Requests must be very specific, yet the structure and scope of available data is unknown ex ante. Furthermore, data access comes with legal risks for research institutions. To enable and foster research on OSNs and provide legal security to researchers, a European OSN research infrastructure is needed that provides data, supports researchers, and offers a secure infrastructure for data processing. This independent research center should provide unified access for researchers to platform data and data donations. It redirects and stores data professionally in line with EU privacy laws, backed by a strong legal team. The center could build upon existing infrastructures such as GESIS and NFDI and act as a centralized network between the Commission, researchers, and NGOs.

b. Enable Effective Data Access under Art. 40 DSA

DSA Art. 40, its delegated act, and the Commission’s Data Access Portal are important steps towards improving platform research and enabling transparency, accountability, and oversight on platform mechanisms. However, instead of involving lengthy court proceedings, these rights need to be brought to life and effectively enforced to enable simple data access. Certain legal terminology should be clarified, such as “systemic risks”, “public data”, and “disinformation”. To allow basic research, these terms should not be laid out too narrowly. Looking ahead, researchers should be granted access to users’ complete viewing history rather than merely their active interactions with posts (likes, comments, shares) in order to fully study user behavior. Furthermore, researchers require access to platforms’ internal A/B testing results on the effects of platform mechanisms and, in the future, should be allowed to request specific A/B tests from platforms if they are ethically feasible. Completeness and reliability of datasets provided by platforms must be ensured through external audits.

c. Support Data Donation

Data donations by platform users are another important data source for platform research. The envisioned amendment to Art. 12(5) of the GDPR-Omnibus must be redrafted; in its current form, it would render data donation impossible as it allows the refusal of access requests for purposes other than the protection of data. Options for users to access and donate their data under their own terms must, on the contrary, be simplified. The proposed European Research Infrastructure could support this. Policymakers and civil society actors should support data donation with campaigns and their own donations. In addition, researchers should be incentivized to share datasets in standardized formats to enable greater comparability, replication, and cumulative knowledge building across studies.

d. Fund Research on OSNs

To understand the complex mechanisms and effects of OSNs, sufficient research funds are necessary. Funding should be long-term and sustainable to enable knowledge transfer, in-depth research, and career paths for junior researchers. Funding structures should reward interdisciplinary work, collaboration with non-academic partners from civil society, industry, and politics, and active science communication to the public to ensure research attends to societal needs and practical concerns.

Research is needed to quantify the undesired effects of platform design choices on citizens and societies, such as opinion and affective polarization, the dissemination of falsehoods, and the impact of foreign powers. Research is essential in providing factual evidence for design solutions for democratic social networks.

2. Effectively Enforce DSA & DMA to Shape Democratic Online Social Networks

In the past years, the EU has developed strong regulatory frameworks on the digital ecosystem, including the Digital Services Act, the Digital Markets Act and the AI Act. Now is the time to implement and enforce them effectively, to safeguard the democratic public sphere – despite pressure from across the Atlantic. We urge EU policymakers to use the tools at hand to shape safe online environments for all citizens, rather than excluding singular groups by prohibition. Recent investigations by the Commission into TikTok and Meta, in line with Art. 66 et seq. DSA, are a strong signal. Trusted Flaggers under the DSA, such as HateAid, who support the removal of harmful content online, must be protected by member states against intimidation and repression from foreign governments.

a. Take Concrete Legislative and Institutional Measures

The upcoming Digital Fairness Act offers a chance to strengthen consumer protection and develop instruments targeting Dark Patterns on platforms. Protection of user data remains crucial to both the consumer and to European sovereignty and must not be weakened through the Digital Omnibus. To ensure effective implementation and democratic oversight of platforms, the Commission should allocate additional dedicated staff. The success of the DSA's implementation should be monitored using clear, transparent metrics.

b. Ensure Transparency and Accountability

Especially the transparency requirements in the DSA (Art. 15, 24, 42 DSA) should be followed closely, including those relating to recommender systems (Art. 27 DSA) and content moderation practices (Art. 16 and Art. 23 DSA). Looking ahead, the platform algorithms should be made fully transparent and subject to external audits by technical experts to ensure accountability. Furthermore, classifiers on disinformation and hatespeech used by platforms for ranking and moderation purposes should also be made public and opened for public scrutiny. Synthetic content and deepfakes must always be clearly labeled in compliance with Art. 50 AI Act.

c. Enforce Interoperability Between Platforms

Very large online platforms (VLOPs) should be required to ensure interoperability in accordance with Article 6 of the DMA. This would allow users to engage with their preferred platform, rather than being forced towards the one with the largest user base. This is an important base for building and scaling successful alternatives to existing platforms, as network effects often prevent users from switching.

3. Force Platform Providers to Make their Platforms Safe by Design

In line with Art. 35 DSA (Mitigation of Risks), providers of very large online platforms should adapt the design of their products to make online environments safe by design for everyone. Our in-depth recommendations on improving platform mechanics can be found in TWON Deliverable 6.3. Similar to the existing guidelines on the protection of minors and guidelines on the mitigation of systemic risks for electoral processes, the Commission should issue guidelines on recommender systems and the design of very large online platforms (VLOPs), including the following design aspects:

a. Promote Content Diversity and Civility through Algorithm Design

Algorithms should be designed to deliver a curated mix of content that balances emotional tones and introduces users to various topic areas and political viewpoints, fostering a more inclusive and creative digital environment. False content, as well as strongly partisan or triggering content, should be downranked, while trustworthy news sources are labeled and prioritized. Our TWON simulations and TWONy demonstrators clearly show the strong effects on the spread of opinions by ranking algorithms functioning chronologically, randomly, or optimizing for emotionality.

One option would be to allow users to customize their algorithmic settings from a predefined set of options. This promotes personalized digital autonomy while ensuring baseline exposure to diverse perspectives. This, however, requires algorithmic transparency.

b. Implement Nudges for Responsible User Behavior in Platform Design

Platform design is decisive in incentivizing user behavior. To foster responsible behavior without infringing on user rights, platforms should implement specific nudges. These can include informational cues on content origin (e.g., whether the authoring account is anonymous or state-controlled) and inoculation and prebunking interventions to hamper the effect of mis- and disinformation. Explainable AI tools to classify and debunk disinformation in real time should be integrated to platforms, ensuring that users are informed about the credibility of the content they engage with. To avoid the impression of false consensus on an opinion, metrics should be displayed on active user reactions (likes, comments, shares) as well as on passive user interactions with a post (viewing the post without interaction). To limit toxic behavior and the spread of disinformation, friction could be added, for example by adding an extra confirmation click or a pop-up with humanizing cues.

c. Limit Psychologically Harmful Effects and Addictiveness

Limiting the psychologically harmful effects of online social networks on adults and youths is crucial. Design measures limiting addictiveness and compulsive behavior must be implemented, such as reducing the rate at which new material appears in news feeds and integrating deliberate interruptions after specific time frames, prompting users to take a break, or making the user experience unattractive through black-and-white mode. Platforms must be required to raise awareness about the harmful effects of content and include links to mental health help hotlines. To reduce platforms' financial incentives to maximize user screen time, the right to display advertisements could be restricted to the first hour a person spends on the platform each day.

4. Support the Development of European Public Platforms

To enable long-term European sovereignty over communication infrastructures, alternatives to existing platforms seem necessary. Den nächsten Satz ersetzen durch: These alternatives must be independent of foreign governments, opinionated private entities, and monopolistic structures.

The development of independent alternative platforms, aligned with EU standards, should be supported to counterbalance existing platform monopolies. Decentralized platforms, such as Mastodon, could provide more inclusive discourse, relying on pro-democratic and transparent algorithms. Publicly funded non-state entities, such as the Wikimedia Foundation, could manage these platforms to avoid state misuse.

To start the shift toward a European Public Platform, a democratic, participatory process should be set up at EU-level to define goals and requirements for such an alternative platform, ensuring alignment with both the user needs and societal goals.

Moreover, e-participation tools should be utilized to enable the meaningful participation of citizens in local discussions and decision-making processes at various levels within the state.

5. Evidence-Based Policymaking

Before legislation is enacted, legal sandboxing tools such as TWON-as-a-service should be used to simulate counterfactual scenarios and the effects of specific platform design options on democratic. Such simulation environments enable independent evaluations in contexts where real-world A/B testing is either not feasible or would raise ethical concerns. Policymakers should also consult researchers from the broad field of platform research to draw on established empirical evidence regarding platform design and regulation, thereby supporting effective and goal-oriented policymaking.

6. Promote Media Literacy and Knowledge on Platform Mechanisms

With AI-generated disinformation rising, targeted campaigns to improve media literacy are essential. It is recommended to launch media literacy campaigns and educational programs targeting various age groups, with a particular focus on empowering individuals to recognize and counteract disinformation and manipulation by AI tools. Tools such as the TWONy and TWONderland demonstrators can effectively support media literacy education. Digital literacy is not only a precondition for inclusive online participation and dialogue. Knowledge of platform mechanisms and their societal effects is also a prerequisite for digital citizenship and advocacy for the individual political interests on platform regulation.

4. Support Independent Journalism

A European fund should support independent media outlets that adhere to high-quality standards. This is important to foster fact-based information on OSNs, counter misinformation, and strengthen public discourse. It is essential to ease the economic pressure on media outlets resulting from the shift of advertising budgets away from the press and toward online social networks. To ensure transparency and accountability, this fund should be managed by an independent body.

Acknowledgements

We would like to extend our sincere thanks to all participants of our Citizen Labs and the Policy Hackathon for their thoughtful input, critical reflections, and dedicated engagement. Their contributions significantly shaped the recommendations presented in this document. The participants of the Policy Hackathon, held in Berlin on January 28, 2026, were:

Judith Peterka, Annette Zimmermann, Tobias Bacherle, Lena-Maria Böswald, Gregor Bauer, Clara Ruthardt, Yasmine Goldhorn, Malte Zaumseil, Victoria Godik, Jan Rau, Leonie Oehmig, Mauritius Dorn, Paula Charlotte Matlach, Eckehard Olbrich, Ina Ni, Isabel Bezzaoui, Maximiliane Linde, Kira Wisniewski, Alenka Gucek, Andreas Reitenbach, Fabio Sartori, Michael Mäs, Antoine Verrière, Alisa Kerschbaum, Corinna Oschatz, Damian Trilling, François t'Serstevens, Ljubisa Bojic, Vera Mevorah, Heiko Paulheim, Achim Rettinger, Christoph Hau, Nils Schwager, Raghvi Baloni, Simon Munker, Sjoerd Stolwijk, Jonas Fegert and Cosima Pfannschmidt.

We also thank Tahireh Panahi for her input on the Digital Services Act and on prebunking interventions:

[Dynamo Policy Paper, 2024](#)

TWON (project number 101095095) as a research project is fully funded by the European Union under the Horizon Europe framework (HORIZON-CL2-2022-DEMOCRACY-01-07).



**Funded by
the European Union**