

# **TWin of Online Social Networks**

Deliverable D6.4

## **Citizen Lab Report & Materials**

Main Authors: Kira Wisniewski & Cosima Pfannschmidt



Funded by  
the European Union

## About TWON

TWON (project number 101095095) is a research project, fully funded by the European Union, under the Horizon Europe framework (HORIZON-CL2-2022-DEMOCRACY-01, topic 07). TWON started on 1 April 2023 and will run until 31 March 2026. The project is coordinated by the Universiteit van Amsterdam (the Netherlands) and implemented together with partners from Universität Trier (Germany), Institut Jozef Stefan (Slovenia), FZI Forschungszentrum Informatik (Germany), Karlsruher Institut für Technologie (Germany), Robert Koch Institute (Germany), Univerzitet u Begogradu - Institut za Filozofiju I Drustvenu (Serbia), Slovenska Tiskovna Agencija (Slovenia), Dialogue Perspectives e.V. (Germany).

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.



Funded by  
the European Union



**DIALOGUE  
PERSPECTIVES  
E.V.**

<b>Project name</b>	<b>TWin of Online Social Networks</b>
<b>Project acronym</b>	<b>TWON</b>
<b>Project number</b>	<b>101095095</b>
<b>Deliverable number</b>	<b>D6.4</b>
<b>Deliverable name</b>	<b>Citizen Lab Report &amp; Materials</b>
<b>Due date</b>	<b>31.03.2026</b>
<b>Submission date</b>	<b>26.03.2026</b>
<b>Type</b>	<b>R – Document, report</b>
<b>Dissemination level</b>	<b>PU – Public</b>
<b>Work package</b>	<b>WP6</b>
<b>Lead beneficiary</b>	<b>Dialogue Perspectives e.V. (DIA)</b>
<b>Contributing beneficiaries and associated partners</b>	<b>FZI Forschungszentrum Informatik (Germany), Universiteit van Amsterdam (UvA), Universität Trier (UT), Institut Jozef Stefan (JSI), Karlsruher Institut für Technologie (KIT), Robert Koch Institut (RKI), Univerzitet u Begogradu - Institut za Filozofiju I Drustvenu (UoB), Slovenska Tiskovna Agencija (STA)</b>

# Contents

- Figures ..... 4**
- Abbreviations ..... 4**
- 1. Introduction ..... 5**
- 2. Reports about the TWON Citizen Labs ..... 7**
  - 2.1. First Citizen Lab: “Navigating the Confluence of Real-Life Crisis and Digital Democracy” ..... 8
  - 2.2. Second Citizen Lab: “Free Speech or Free-for-All? Tackling Hate, Misinformation and Polarisation in the Age of AI and Tech-Oligarchs” ..... 10
  - 2.3. Third Citizen Lab: “Digital Remembrance. Plural Perspectives” ..... 13
  - 2.4. Fourth Citizen Lab: “Rebuilding Bridges: Dialogue, Trust and Solidarity Post October 7th” ..... 16
- 3. Reflection of Citizen Labs and Digital Citizenship .....20**
- 4. Final TWON Policy Recommendations .....21**
- 5. Conclusion .....26**
- 6. Appendix.....28**
  - 6.1. Agenda 1st Citizen Lab in Karlsruhe, Germany, 16 - 19 September 2024 ..... 28
  - 6.2. Agenda 2nd Citizen Lab in Vienna, Austria, 11 - 14 May 2025 ..... 29
  - 6.3. Agenda 3rd Citizen Lab in Chemnitz, Germany, 7 - 9 October 2025 ..... 30
  - 6.4. Agenda 4th Citizen Lab in Brussels, Belgium, 14 - 18 March 2026 ..... 31
  - 6.5. Workshop Results: Requirements for a Democratic Platform Alternative (2nd Citizen Lab, Vienna) ..... 32

# Figures

Figure 1. Participants of the Brussels Citizen Lab in March 2026

Figure 2. Participants of the Karlsruhe Citizen Lab Presenting Workshop Results on the Ethical Governance of a TWON

Figure 3. Participants of the Karlsruhe Citizen Lab

Figure 4. TWON Researchers at the Vienna Citizen Lab (from left to right: Damian Trilling, Fabio Sartori, Jonas Fegert, Kira Wisniewski, Cosima Pfannschmidt, Simon Munker)

Figure 5. Participants of the Vienna Citizen Lab at the Public Event “Free Speech or Free-for-All? Tackling Hate, Misinformation and Polarisation in the Age of AI and Tech-Oligarchs” (Foto Credits: Andreas Daniel Jakob)

Figure 6. The TWON Science Tent at Stadthallenpark, Chemnitz

Figure 7. Public Panel on "Practices of Rememberance in Digital Spaces" at the Documentation Centre on the NSU Complex

Figure 8. Simon Munker Presenting MacroTWONy at the Brussels Citizen Lab

Figure 9. Damian Trilling's Talk at the Brussels Citizen Lab

Figure 10. Cosima Pfannschmidt's Introduction to TWON, the Digital Public Sphere and Platform Mechanisms at the Brussels Citizen Lab

Figure 11. TWON Dissemination Event in Brussels with Katarina Barley, Vice-President of the European Parliament

# Abbreviations

<b>AI</b>	<b>Artificial Intelligence</b>
<b>DMA</b>	<b>Digital Markets Act</b>
<b>DSA</b>	<b>Digital Services Act</b>
<b>EU</b>	<b>European Union</b>
<b>GDPR</b>	<b>EU General Data Protection Regulation</b>
<b>NGO</b>	<b>Non-governmental organisation</b>
<b>OSN</b>	<b>Online Social Networks</b>
<b>TWON</b>	<b>Twin of Online Social Networks</b>
<b>TWONy</b>	<b>Demonstrator of the Twin of Online Social Networks</b>

# 1. Introduction

In recent years, the question of how European research projects generate societal and policy impact has gained increasing attention in both research funding and public debate. While impact is expected and formally required within EU-funded research frameworks, many projects face practical challenges in translating scientific results into broader societal relevance. Research consortia often struggle to identify appropriate points of engagement, communicate complex findings to non-academic audiences, and establish meaningful interactions with policymakers, civil society actors, and citizens beyond conventional dissemination activities.

The TWON project deliberately adopted a different approach from the outset. Given the societal relevance and political sensitivity of research on online social networks, polarization, and misinformation, the consortium considered it crucial not to conduct research behind closed doors and only communicate results at the end of the project. Instead, TWON aimed to create participatory formats that would enable early dialogue, iterative feedback, and mutual learning between researchers and broader societal stakeholders throughout the project lifecycle. The Citizen Labs, therefore, represent not only a dissemination activity but a core element of TWON's impact strategy. Within the framework of Task 6.3 of the Grant Agreement, the TWON consortium committed to conducting and disseminating four Citizen Labs, designed to involve a broader European public in the research process through participatory workshop formats. The Citizen Labs were intended to translate scientific findings into accessible formats for non-academic audiences while simultaneously providing a space where researchers could critically discuss emerging findings with citizens, practitioners, and civil society actors. This approach allowed the consortium to integrate societal perspectives early and continuously into research outputs and the development of policy recommendations.

Within the TWON project, Citizen Labs are implemented as a participatory engagement format that enables researchers, practitioners, and members of the public to jointly reflect on and discuss the societal implications of ongoing research. Being an innovative concept for citizen engagement in a research project, the methodology was developed, tested and adapted throughout the process. Rather than following a fixed methodological template, the Citizen Labs were deliberately designed as flexible and context-sensitive formats, adapted to different thematic and regional settings. Across the four iterations, the Citizen Labs combined elements of science communication, participatory policy dialogue, and hands-on experimentation with research demonstrators, thereby creating a structured interface between scientific research, societal experience, and policy reflection.

The successful implementation of this strategy was made possible through the collaboration with Future 500 (DialoguePerspectives e.V.), whose expertise in participatory dialogue formats and European civil society engagement enabled the project to reach diverse audiences across Europe. In parallel, the FZI Research Center for Information Technology, acting as dissemination leader within TWON, ensured that insights emerging from the Citizen Labs were systematically connected to scientific dissemination and policy communication activities. Together, these structures enabled TWON to establish an iterative interface between research, society, and policy actors.

The objectives of the Citizen Labs were therefore threefold: firstly, to integrate lived experiences into research processes; secondly, to educate participants about how online social networks and strengthen digital citizenship, and thirdly, to strengthen the translation of project results into policy and decision-making contexts. Public panels and outreach activities were integrated into each Citizen Lab to ensure that discussions extended beyond the workshop setting and reached broader audiences, including policymakers. Across all Citizen Labs, researchers from different TWON work packages were actively involved in the design

and implementation of the programme. TWON experts contributed to workshops, lectures, demonstrations, panel discussions, and moderation formats, ensuring a close integration of scientific insights and practical engagement. Sessions covered methodological foundations, ethical considerations, AI-supported simulations, network modelling, and alternative platform designs, allowing for continuous citizen feedback across all components of TWON and strengthening the real-world relevance of the project's outcomes.

The four Citizen Labs were implemented in different European contexts between 2024 and 2026, each building on the previous while adapting the format to its local setting. The first Citizen Lab, held in Karlsruhe (September 2024), marked the start of developing the TWON policy recommendations within a workshop-based setting. The Vienna Citizen Lab (May 2025) built on these discussions by combining expert input with practical experimentation, including hands-on experimentation with TWONy demonstrators and structured policy prototyping sessions. The Chemnitz Citizen Lab (October 2025) expanded the format into a publicly accessible urban setting, placing particular emphasis on low-threshold engagement with global and local communities, while further contextualising the policy discussions through themes of remembrance culture and societal inclusion. The final Citizen Lab in Brussels (March 2026) brought the process into a European policy environment, with a strengthened focus on dialogue, democratic resilience, and the relationship between platform dynamics and broader societal debates. It consolidated previous insights and concluded the work on the TWON policy recommendations. Public outreach components complemented the workshop formats across all events and facilitated engagement with broader audiences and policy stakeholders. Insights emerging from the Citizen Labs informed both the development of TWON's policy recommendations and the communication of the project's research results. The iterative structure of the four Labs allowed feedback from participants to be reflected in subsequent discussions, demonstrator development, and policy prototyping activities. The TWON Citizen Labs are all publicly documented on the TWON website ([twon-project.eu/citizen-labs/](https://twon-project.eu/citizen-labs/)).

This deliverable is structured as follows: following the introduction, Chapter 2 provides individual reports on all four Citizen Labs, outlining their formats, participants, and attached outreach activities. Chapter 3 reflects on the overarching insight and impact gained from implementing Citizen Labs as an integral part of the TWON project. Chapter 4 presents the final TWON policy recommendations and summarises their development process. The document concludes with a final summary, followed by an Appendix containing additional supporting materials and documentation.



Figure 1. Participants of the Brussels Citizen Lab in March 2026

## 2. Reports about the TWON Citizen Labs

The Citizen Labs were conducted by DialoguePerspectives e.V., a Berlin-based independent European platform for strengthening plurality and democracy in Europe through educational and civil society programmes. DialoguePerspectives develops impact-oriented civil society programmes to promote resilience in European society. Its work focuses on interreligious and worldview dialogue, combating antisemitism in Europe, developing and establishing a pluralistic European culture of remembrance, and communicating Jewish life in Germany through nationwide educational programmes. The European network of DialoguePerspectives e.V. includes over 500 young European leaders and more than 700 partners from European civil society. Its work focuses on the intersection of science, culture, politics, and socio-political practice. DialoguePerspectives e.V. is committed to strengthening the foundations of a resilient, pluralistic democracy in Europe. It specialises in establishing interdisciplinary, results-oriented European platforms for civil society and in implementing innovative programmes.

One of these programmes is Future 500 (formerly DialoguePerspectives), which brings together Europe's most promising talents from politics, business, academia, the arts, and civil society. Within DialoguePerspectives e.V., Future 500 was the programme responsible for the conception, organisation, and implementation of the TWON Citizen Labs. Future 500 develops concrete socio-political initiatives and implements them within European civil society. Through its work, Future 500 sustainably strengthens social cohesion across Europe by promoting intercultural and interreligious worldview dialogue, developing innovative approaches to counter social polarisation, and building long-term partnerships among politics, business, academia, and civil society. Based on best practices, collaborative research initiatives, and targeted public advocacy programmes, sustainable strategies for social integration, political stability, and democratic

resilience are developed. The programme especially brings together participants from diverse communities and backgrounds, encompassing individuals with 19 different religions and beliefs across 25 European countries. Through their unique perspectives and expertise, they contribute to fostering understanding, cooperation, and a pluralistic, democratic, and cohesive Europe. Within its commitment to TWON, DialoguePerspectives integrates its established expertise in fostering pluralistic dialogue and combating societal polarisation with a focused emphasis on digital democracy, hate speech, and disinformation.

The programme has a proven track record in formulating actionable calls to action and comprehensive policy briefs, as demonstrated through its European Leadership Workshops on topics such as “Plurality & Anti-Discrimination in the Workplace” and events like “Entering the Engine Room: Policy Briefs as a Means of Forging a Pluralistic Europe.” These initiatives include the development of policy recommendations aimed at advancing a cohesive and pluralistic Europe, methods and skills that are also employed within TWON.

In this context, DialoguePerspectives has prioritised educating participants on the dynamics of online platforms, the role of AI, and strategies such as pre-bunking and de-bunking to combat disinformation – an essential step towards strengthening and promoting a pluralistic and democratic European society. These efforts underscore the programme’s expertise and capacity to develop actionable policy recommendations for TWON contributing to shaping a cohesive Europe. A podcast episode on democracy in the digital age with Dr. Jonas Fegert (FZI Forschungszentrum Informatik) was recorded, focusing on the role of platforms in European democracies.

Various formats of science communication and discussion were employed throughout the four TWON Citizen Labs. These include talks and lectures, interactive workshops with hackathon-style working sessions, world-café-style small group discussions, silent reflections, guided museum tours, panel talks and a BarCamp. As the research states, methods and results are rather complex, and the Citizen Labs offered the time and space to convey some of this complexity, the applied formats proved to be more suitable than the initially foreseen one-pagers with condensed research results. The presentations held by TWON researchers at the Citizen Labs can be found in the appendix. Furthermore, for most of the Citizen Labs, the scientific process within the TWON project was still ongoing, so it would have been premature to produce one-pagers with final research results. However, to communicate research results to a non-scientific audience in line with tasks 6.3 and 6.4, policy briefs on regulating online social networks and the ethics behind digital twins have been created (see deliverable 6.5) and discussed with participants. Citizen Lab formats were iteratively adapted, learning from experiences at each iteration.

## **2.1. First Citizen Lab: “Navigating the Confluence of Real-Life Crisis and Digital Democracy”**

*Karlsruhe, Germany, September 16-19, 2024*

The Karlsruhe Citizen Lab “Navigating the Confluence of Real-Life Crisis and Digital Democracy” included expert inputs, moderated discussions, World Café sessions, BarCamp-style exchanges, and policy brainstorming workshops.

The programme featured research-based contributions by consortium members Prof. Damian Trilling (University of Amsterdam), Prof. Achim Rettinger (University of Trier), and Dr Eugen Pissarskoi (Karlsruhe Institute of Technology), addressing research approaches to social media dynamics, the ethical foundations of developing TWONs, and the role of AI and online platforms in digital democracy. Additional inputs covered regulatory frameworks such as the EU Digital Services Act, media literacy strategies, and approaches to pre-

bunking and de-bunking in online environments, thereby situating the discussions within current European policy debates.

Methodologically, the programme combined theoretical perspectives on digital democracy, AI, and platform regulation with participatory dialogue formats to integrate participants' lived experiences and professional expertise. The Karlsruhe Citizen Lab marked the starting point for the development of the TWON policy recommendations. At the conclusion of the four-day workshop, participants collaboratively drafted a policy brief outlining potential regulatory approaches, which subsequently informed further discussions within the TWON consortium. The Citizen Lab approach emphasised reflection, peer learning, and the joint development of concrete policy-oriented outputs.

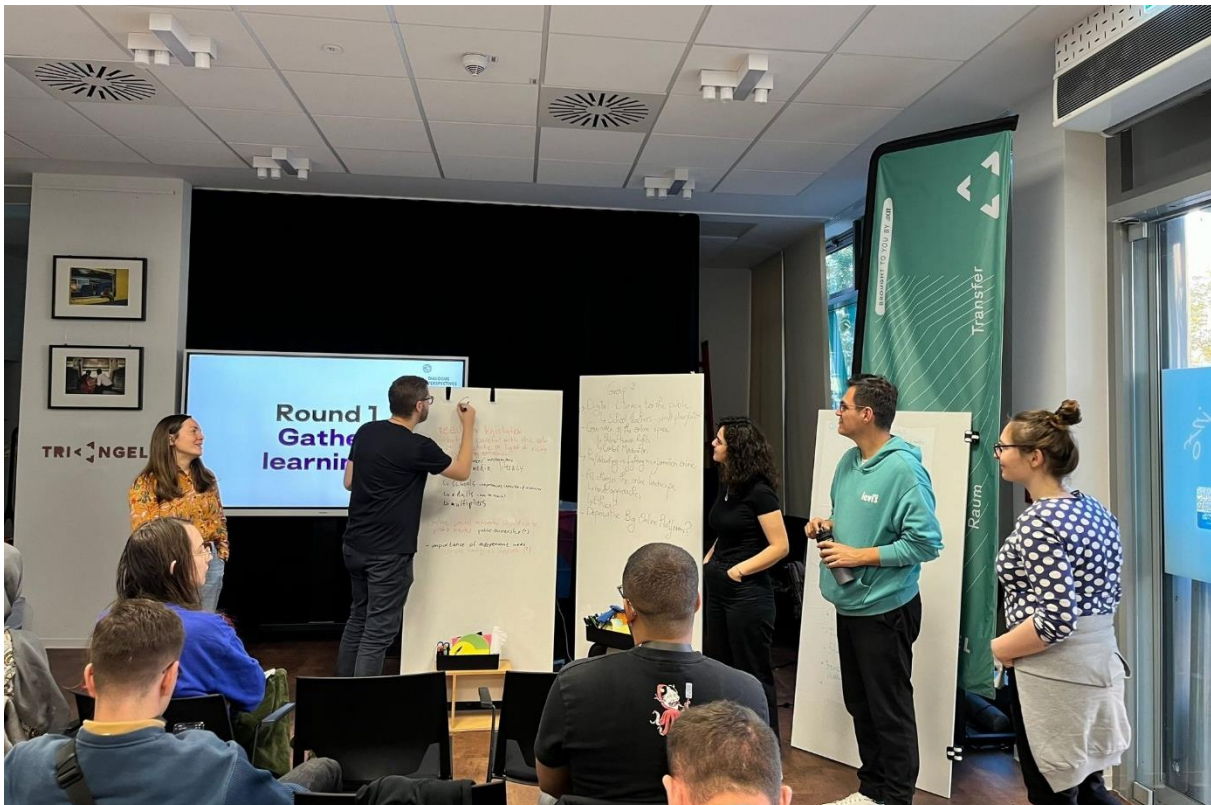


Figure 2. Participants of the Karlsruhe Citizen Lab Presenting Workshop Results on the Ethical Governance of a TWON

A public World Café-style session and a BarCamp were organised during the workshop in the Triangel Event Space Karlsruhe, explicitly designed to engage a broader audience and connect Citizen Lab discussions with public perspectives on polarisation, democracy, and digital transformation. Karlsruhe is home to the “European Joint Research Centre”, which serves as the European Commission’s primary hub for nuclear safety and security research. It is located at KIT North Campus and is thus closely connected to Karlsruhe Institute of Technology (KIT), one of the TWON consortium members.

Participants came from several European countries, including Denmark, Turkey, Sweden, Albania, Austria, and Italy. They represented diverse academic and professional backgrounds spanning media and communication studies, contemporary history, theology, education policy, youth work, and public sector strategy.



Figure 3. Participants of the Karlsruhe Citizen Lab

Several participants were engaged in academic research projects, including work within the Erasmus Mundus Master's programme Education Policies for Global Development and a DFG-funded research project on religion, digitality and confessionality. Research interests included the strategic communication of far-right political actors on social media, affect and memory in contemporary political discourse, and the role of digital platforms in shaping public debate.

Other participants brought professional experience from civil society initiatives addressing hate speech and democratic resilience, including involvement in the No Hate Speech Youth Campaign of the Council of Europe and the Italian National Network Against Hate Speech and Hate Crimes. Additional backgrounds included work in Europe Direct information centres, journalism and online editorial departments, democracy education and youth social work, as well as public sector consulting at Accenture Strategy & Consulting. Several participants were affiliated with institutions such as Humboldt University of Berlin and the University of Vienna.

## 2.2. Second Citizen Lab: “Free Speech or Free-for-All? Tackling Hate, Misinformation and Polarisation in the Age of AI and Tech-Oligarchs”

*Vienna, Austria, May 11-14, 2025*

The Vienna Citizen Lab titled “Free Speech or Free-for-All? Tackling Hate, Misinformation and Polarisation in the Age of AI and Tech-Oligarchs” combined expert-led workshops, participatory discussion formats, and hands-on experimentation. The programme included interactive lectures and research-based contributions

addressing social media dynamics, algorithmic amplification, AI systems, platform governance, and questions of foreign influence and digital manipulation.

Dr Jonas Fegert introduced the TWON project and its research objectives, while Prof. Dr Damian Trilling (University of Amsterdam) critically examined prevailing assumptions about echo chambers, filter bubbles, and the spread of disinformation, highlighting the methodological limits of current research on social media dynamics. Prof. Dr Achim Rettinger (University of Trier) addressed the intersection of AI agents and online discourse, discussing both risks and opportunities of algorithmic content curation in shaping public opinion and emotional responses.

A central methodological component was the practical demonstration of the TWONy platform within a Citizen Lab context. Led by Simon Munker (University of Trier) and Dr. Fabio Sartori (KIT), participants explored the micro and macro TWONy demonstrators and worked in small groups to simulate online interactions using AI-generated personas under different ranking mechanisms, analyse emotional dynamics, and compare outcomes with real-world social media behaviour. Participants were asked to provide feedback on the TWONy tools for their evaluation and improvement. Additional expert inputs covered algorithmic amplification of hate speech and misinformation (Dr. Even Kapros), platform accountability and recommender systems (Dr. Julia Neidhardt, TU Vienna), foreign influence and digital manipulation (Julia Smirnova, CeMAS), digital memory and activism (Dr. Stefania Manca, Italian National Research Council), and questions of social media, AI, and freedom of speech (Nuriyatul Lailiyah).

FZI researcher Cosima Pfannschmidt facilitated a workshop on envisioning democratic alternatives to existing online social networks, encouraging participants to reflect on governance structures, content moderation models, and ranking criteria for a future-oriented digital public sphere. A policy prototyping session focused on translating these discussions into concrete regulatory proposals. Participants worked on further developing the TWON policy recommendations, building on the initial work undertaken during the Karlsruhe Citizen Lab and testing emerging policy directions against empirical research insights and technical demonstrations.



*Figure 4. TWON Researchers at the Vienna Citizen Lab (from left to right: Damian Trilling, Fabio Sartori, Jonas Fegert, Kira Wisniewski, Cosima Pfannschmidt, Simon Munker)*

An open public evening event entitled “Free Speech or Free-for-All? Tackling Hate, Misinformation and Polarisation in the Age of AI and Tech-Oligarchs” was organised, featuring two panel discussions with experts from academia, civil society, and policy. The first panel, “Digital Democracy and the Power of Platforms: Policy, AI, and Accountability”, brought together Prof. Dr Achim Rettinger, Benjamin Fischer (CeMAS), Judith Peterka (Federal Chancellery of Germany and TWON Advisory Board), Natascha Strobl (Political Scientist and expert on right-wing extremism), and Dr Sebastian Heidebrecht (Centre for European Integration Research, Vienna University). The second panel, “Countering Hate and Information Manipulation – Strategies for a Safer Digital Sphere”, included Alina Bricman (Director of EU Affairs at B’nai B’rith International), Rosa Jellinek (Activist and Social Media Expert), Selin Aydın (Programme Manager CLAIM - Alliance Against Islamophobia and Anti-Muslim Hate), and Stefania Manca (Institute of Educational Technology, Italian National Research Council). The discussions were moderated by Dr. Jonas Fegert and addressed digital democracy, platform accountability, hate speech, and misinformation. The event aimed to disseminate TWON-related research themes to a broader public audience while fostering exchange with external stakeholders.



Figure 5. Participants of the Vienna Citizen Lab at the Public Event “Free Speech or Free-for-All? Tackling Hate, Misinformation and Polarisation in the Age of AI and Tech-Oligarchs” (Foto Credits: Andreas Daniel Jakob)

Vienna was chosen as the location, as it is home to a number of EU institutions, such as a representation of the European Commission, a connection office of the European Parliament, and the European Union Agency for Fundamental Rights (FRA).

Participants came from a range of European countries, including Ukraine, Sweden, Turkey, the Czech Republic, the United Kingdom, Greece, Spain, Italy, Kosovo, and Germany. They represented diverse professional and academic backgrounds spanning policymaking, civil society engagement, academia, digital communication, social media consultancy, and public administration. Several participants were actively engaged in research and policy work on digital rights, platform governance, migration and integration policy, democratic resilience, and the societal impact of online communication. Others contributed practical experience from civil society organisations working on hate speech prevention, digital literacy, minority rights, and advocacy at European and international levels. Participants were affiliated with institutions and organisations such as the Institute for Strategic Dialogue (ISD), the MENA Digital Rights Coalition, the Global Coalition for Tech Justice, Amnesty International, the German Federal Office for Migration and Refugees, SABRA, the Arolsen Archives, TU Berlin, the German Embassy in London, and the OECD. This diversity of institutional and professional perspectives contributed to an interdisciplinary exchange and supported the translation of TWON’s research findings into policy-relevant discussions.

### 2.3. Third Citizen Lab: “Digital Remembrance. Plural Perspectives”

*Chemnitz, Germany, October 7-9, 2025*

The Chemnitz Citizen Lab, titled “Digital Remembrance. Plural Perspectives” was embedded in a multi-day study trip combining workshops, site visits, artistic interventions, and a public Science Tent format. Core methods included interactive workshops on digital remembrance, memory politics, and digital democracy,

as well as the deployment of TWON demonstrators (“MicroTWONy”, “MacroTWONy” and “TWONderland”) in a publicly accessible urban space.

From the TWON consortium, Dr. Jonas Fegert (FZI), Dr. Fabio Sartori (KIT), Cosima Pfannschmidt (FZI), Nils Schager (University of Trier), and Simon Münker (University of Trier) participated in the Citizen Lab, presenting the TWONy demonstrators to the public and engaging in discussions on online social networks, democratic discourse, algorithmic ranking mechanisms, and ethical platform design. Cosima Pfannschmidt conducted a workshop on improving online social networks, based on the consortium’s recommendations on improving platform mechanics (deliverable D6.3).

The Citizen Lab / Science Tent was set up in Chemnitz’s central Stadthallenpark (city park), allowing visitors of the European Capital of Culture and local residents to engage with the demonstrators and discussions. Visitors and participants explored how different recommendation logics shape emotional dynamics, exposure to content, and public debate. Conversations focused on the societal impact of online social networks, regulatory approaches, and possibilities for more democratic and ethically grounded platform architectures.



Figure 6. The TWON Science Tent at Stadthallenpark, Chemnitz

The choice of Chemnitz as a location was linked to its status as European Capital of Culture 2025, which provided an opportunity to integrate the Citizen Lab into a broader European context and cultural programme and engage diverse audiences. At the same time, Chemnitz offered a specific local context shaped by debates on remembrance culture, right-wing extremism, and democratic participation.

In contrast to the more workshop-based settings of earlier Citizen Labs, the Chemnitz format placed particular emphasis on low-threshold public engagement in an open, non-institutional space. The programme further integrated guided tours and curator-led discussions linking digital methods with local

histories of remembrance, as well as policy-oriented workshops on political correspondence and platform improvement. Policy discussions were contextualised by connecting digital democracy and platform governance to questions of remembrance culture and societal inclusion, thereby broadening and validating the perspectives developed in previous Citizen Labs.

A public panel discussion entitled “Practices of Remembrance in Digital Spaces” was held at the Documentation Centre on the NSU Complex, joining with the CPPD festival series “Memory Matters”. The panel brought together Dr. Jonas Fegert (FZI), journalist Nhi Le, and social media activist Susanne Siegert (@keineerinnerungskultur), and was moderated by Benjamin Fischer (CeMAS). The discussion addressed how digital methods and online platforms influence contemporary remembrance culture, highlighting both opportunities – such as new educational formats and broader accessibility – and risks, including manipulation, authenticity concerns, and the amplification of hate speech and misinformation through platform mechanisms.



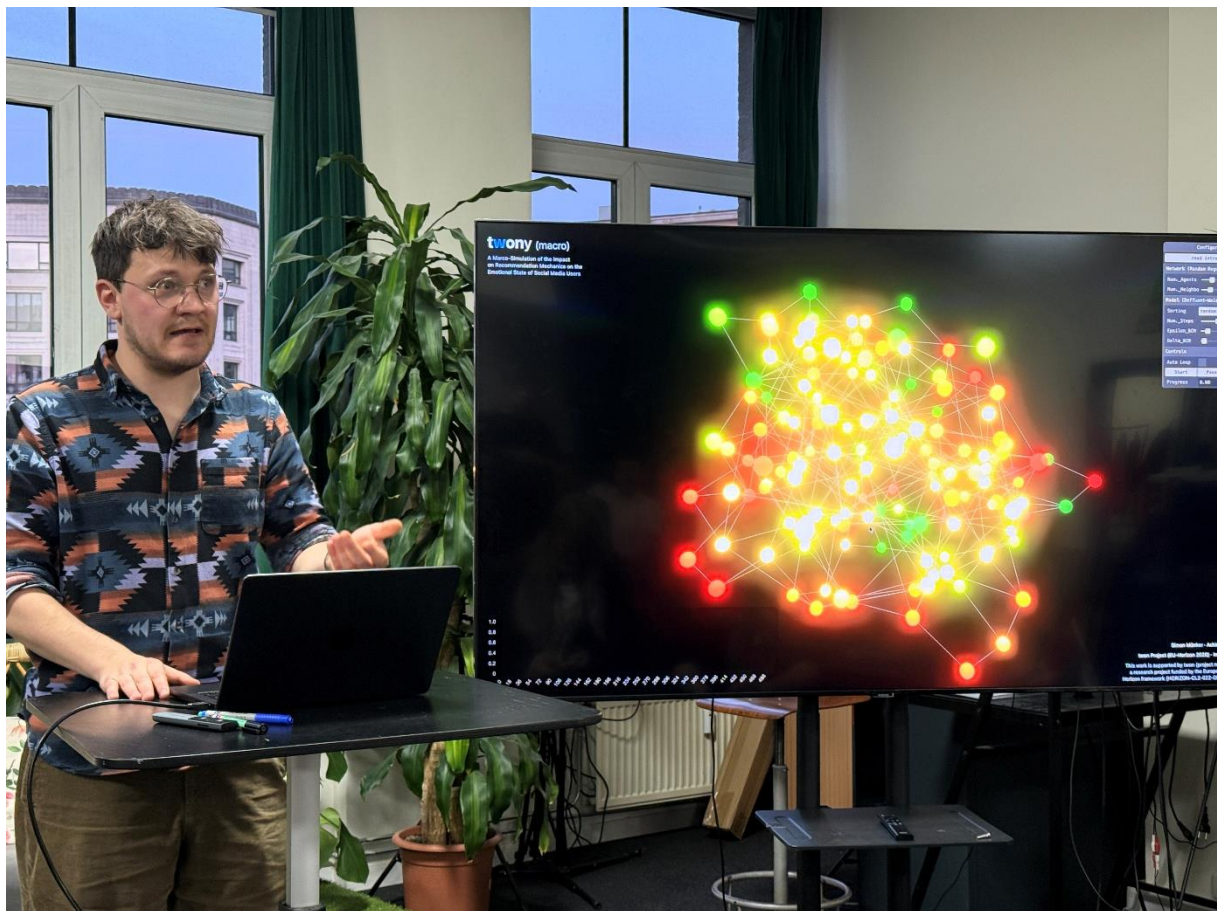
Figure 7. Public Panel on "Practices of Rememberance in Digital Spaces" at the Documentation Centre on the NSU Complex

Participants came from a range of European countries, including Hungary, the Czech Republic, Italy, Poland, Germany, Estonia, Greece, Slovakia, France, and Serbia. They represented diverse academic, journalistic, cultural, and civil society backgrounds, contributing perspectives from Central, Eastern, and Southern Europe. Several participants were engaged in research and academic work in fields such as history, religious and cultural studies, Jewish studies, and media studies. Their research focused on modern European history and cultural memory, religious pluralism and threats to democracy, contemporary antisemitism, and the relationship between media and democratic governance. Participants were affiliated with institutions including Charles University in Prague, the Hochschule für Jüdische Studien Heidelberg, the University of Wrocław, and the Terezín Initiative Institute, where their work encompassed educational programmes, archival digitisation, and contributions to Holocaust-related documentation projects. Other participants

brought professional experience from journalism, civil society advocacy, youth engagement, and cultural practice. These included work as a journalist covering infrastructure and public policy, leadership roles in organisations such as Young Transatlantics (Initiative Junger Transatlantiker e.V.), research and advocacy work on hate speech and radicalisation within digiQ (Bratislava), as well as artistic and curatorial practice focused on memory culture and minority narratives.

## 2.4. Fourth Citizen Lab: “Rebuilding Bridges: Dialogue, Trust and Solidarity Post October 7th”

*Brussels, Belgium, March 14-18, 2026*



*Figure 8. Simon Munker Presenting MacroTWONy at the Brussels Citizen Lab*

The Brussels Citizen Lab, entitled “Rebuilding Bridges: Dialogue, Trust and Solidarity Post October 7th”, was implemented as a multi-day workshop, bringing together European professionals, civil society actors, and researchers in a policy-oriented setting. As the political centre of the European Union and a key site of ongoing debates on digital regulation and platform governance, Brussels provided an appropriate environment for situating the final Citizen Lab within a European policy context.



Figure 9. Damian Trilling's Talk at the Brussels Citizen Lab

The programme combined expert inputs, interactive workshops, moderated dialogue formats, and simulation-based experimentation. A central focus was the interplay between online social networks, public discourse, and democratic cohesion, particularly in the context of societal polarisation and crisis. Research-based contributions addressed the influence of algorithmic curation, AI-driven communication, and platform dynamics on political narratives and interreligious relations, while dialogue-oriented sessions explored how trust and solidarity can be rebuilt in digitally mediated environments.

A key methodological component was the live demonstration and application of the TWONy platform. Participants explored AI-generated personas and different ranking mechanisms (e.g. chronological versus emotion-based ranking) in small-group simulations, analysing emotional dynamics and comparing simulated outcomes with real-world social media behaviour. From the TWON consortium, Prof. Damian Trilling (University of Amsterdam), Prof. Achim Rettinger (University of Trier), Dr. Jonas Fegert (FZI), Cosima Pfanschmidt (FZI), as well as Simon Munker and Christoph Hau (Trier University) contributed to the workshop through expert inputs, moderation, and technical demonstrations.



Figure 10. Cosima Pfanschmidt's Introduction to TWON, the Digital Public Sphere and Platform Mechanisms at the Brussels Citizen Lab

The Brussels Citizen Lab placed a strengthened emphasis on dialogue despite crises, connecting TWON's research on social media dynamics with broader civil society debates on democratic resilience and constructive communication. It also constituted the final working setting within the Citizen Lab framework in which the TWON policy recommendations were reviewed, refined, and consolidated.

As part of the workshop, a public panel discussion entitled "Voices Rising: Rebuilding Bridges. Dialogue, Trust and Solidarity Post-October 7th" was held on 16 March 2026. The event brought together Camila Piastro (European Union of Jewish Students), Barbara von Freytag (Journalist and Political Analyst), Furkan Yüksel (Political Educator), and Prof. Dr Achim Rettinger (Trier University), and was moderated by Igor Mitchnik.

The discussion addressed how AI, online platforms, and media narratives shape public discourse and interreligious relations, and examined possibilities for rebuilding dialogue and trust in polarised societies. The event was open to the public and embedded in the broader Future 500 network workshop.



*Figure 11. TWON Dissemination Event in Brussels with Katarina Barley, Vice-President of the European Parliament*

In line with Task 6.3, the Citizen Lab results and the scientific TWON results were presented to and discussed with European decision makers and policy makers at the European Commission site at a dissemination event entitled “From Research to Sovereignty in the Digital Public Sphere – Insights from the TWON Project”. The programme presented the project’s research results and policy implications in a Brussels policy context, focusing on platform governance, digital sovereignty, and evidence-based regulation.

The participation of the Citizen Lab cohort enabled a direct connection between the workshop discussions and the wider project dissemination activities, prompting reflections on the Citizen Lab process and its contribution to the development of TWON policy recommendations. In this way, the event illustrated the link between the participatory Lab formats and other work packages and dissemination efforts within the project.

Participants came from a range of European countries, including Austria, Georgia, Germany, Hungary, Ireland, Poland, the United Kingdom, the Netherlands, and Spain. They represented diverse professional and academic backgrounds spanning political education, civil society advocacy, public policy, academia, journalism, and interfaith engagement. Several participants were actively engaged in work addressing antisemitism, democratic resilience, digital hate speech, and interreligious dialogue. Professional experience included leadership roles in Jewish community organisations, engagement with European and international institutions, local political office and youth parliamentary work, as well as counter-hate speech training within the framework of the Council of Europe. Academic backgrounds covered fields such as sociology, history, Holocaust and genocide studies, and media and cultural studies, with a focus on narratives, identity, and social cohesion in digitally mediated environments. This diversity of institutional, academic, and practical perspectives contributed to an interdisciplinary exchange and strengthened the connection between TWON’s research on digital discourse and broader European debates on dialogue, trust, and democratic resilience.

### 3. Reflection of Citizen Labs and Digital Citizenship

The Citizen Labs conducted within the framework of the TWON project can be considered a strong success overall in relation to the objectives defined in the proposal and to the broader ambition of fostering digital citizenship in Europe. As envisaged, the Labs were designed to involve a European public in the research process, educate citizens about the functioning and mechanisms of online social networks (OSNs), and jointly reflect on concrete policy implications. Across four workshops in different European contexts, we were able to establish participatory spaces in which researchers, civil society actors, minority representatives, activists, and policy stakeholders engaged in in-depth and discussion-oriented dialogue.

DialoguePerspectives e.V. as owner of task 6.3 proved very effective in successfully implementing the Citizen Labs.

A central strength of the Citizen Labs was their ability to reach a highly diverse audience from across Europe. Participants included individuals from minority communities, activists, practitioners with lived experience of online hate, polarisation, and community-building in digital spaces. This diversity significantly enriched the discussions. The feedback we received, particularly from members of minority groups, underlined how valuable it was to reflect on platform mechanisms not only as abstract governance tools but as structures that shape everyday experiences online. Participants shared how they encounter hate speech, exclusion, and misinformation, but also how they use digital platforms to build supportive communities and foster dialogue. This grounded perspective strengthened the empirical and normative dimensions of our research.

The thematic focus of the individual Citizen Labs further contributed to this richness. Discussions in Karlsruhe were centred on platform regulation, the Digital Services Act and questions of enforcement and pre- and debunking of misinformation. In Vienna, the focus on the “TWONy” demonstrators highlighted innovative approaches to science communication. Chemnitz connected platform mechanisms to questions of remembrance culture and how historical narratives are negotiated online. In Brussels, platform mechanisms were discussed in relation to enabling dialogue in times of polarization and improving democratic debate culture. Across these settings, participants engaged critically with the ways in which algorithmic curation, moderation systems, and regulatory frameworks influence information exposure, participation, and polarisation.

In doing so, the Citizen Labs made a tangible contribution to fostering digital citizenship. Within TWON, digital citizenship is understood as the ability to participate in and actively shape digital environments while respecting and strengthening human dignity, human rights, and democracy. The Citizen Labs strengthened participants’ understanding of how platform mechanisms shape their information diet and online interactions. They enhanced awareness of rights online, including privacy, data protection, freedom of expression, and the responsibility to respect the rights of others. They also encouraged critical reflection on well-being online, and the social and emotional dimensions of digital engagement. Importantly, digital citizenship was addressed not only as a set of technical skills but as a question of how we choose to live together and uphold shared democratic values in digital spaces. In this respect, the Citizen Labs clearly contributed to the intended outcome of enhancing capacities for digital citizenship through citizen education, platform literacy, and grassroots participation.

Another important achievement was the interdisciplinary and cross-sectoral character of the discussions. The Citizen Labs successfully prevented the TWON project from remaining within a solely academic bubble, without contact to practice. Researchers presented their findings, but practitioners, civil society representatives, and external experts from different disciplines were equally central to the dialogue. This exchange ensured that theoretical insights were confronted with practical realities and policy constraints. It

also contributed to community-building around online social networks, bringing together people with different perspectives who would not usually meet in a single forum. The format was deliberately discussion-oriented, allowing for genuine exchange rather than one-directional dissemination. At the same time, the programme of the Citizen Labs was continuously adapted to account for new participants, and the evolving state of the TWON research. This flexibility increased relevance and responsiveness.

The process also generated important learning effects for the project itself. Through the Citizen Labs, it became clear that explaining platform mechanisms and alternative designs in purely abstract terms can limit accessibility. Participants expressed a need for more tangible illustrations, which led to the insight that a demonstrator would be a valuable addition to future research and engagement activities. This lesson has methodological implications beyond TWON and highlights the importance of making complex socio-technical systems experientially understandable.

Overall, the Citizen Labs fulfilled the multi-level impact envisaged in the proposal. On a first level, citizens were directly reached and engaged intensively with questions of platform governance and democratic debate. On a second level, participants are expected to act as multipliers within their own communities. On a third level, the inclusion of stakeholders and policymakers, particularly in public panels, ensured that discussions resonated beyond the workshops themselves and contributed to broader societal and policy debates. The Citizen Labs demonstrated that citizens are willing and able to engage with complex questions about online social networks when provided with inclusive and well-structured formats. They also showed that participatory approaches can meaningfully enrich empirical research and strengthen its societal relevance, motivating their repeated inclusion in future research projects. At the same time, The Citizen Labs showed that translating complex methodological questions and empirical findings on polarisation, misinformation, and OSN governance into concise, accessible educational formats is challenging and requires significant effort. Balancing scientific nuance with clarity for non-specialist audiences is demanding, but possible and fruitful.

## 4. Final TWON Policy Recommendations

The following policy recommendations have been developed throughout the TWON project duration and iteratively refined after discussions with researchers, policymakers, participants of Citizen Labs, and social media professionals. They were presented to EU policymakers at the final TWON dissemination event in Brussels on March 17, 2026, entitled “From Research to Sovereignty in the Digital Public Sphere – Insights from the TWON Project”.

### **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 policymakers 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, reinforcing affective polarization, while 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 risk 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 to strengthen European digital sovereignty and democratic resilience. It calls for effective enforcement of existing EU digital laws, robust and standardized research access to platform data, 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 has developed an innovative “digital twin” approach: instead of 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, Karlsruhe Institute of Technology (KIT), University of Belgrade, University of Trier, FZI Forschungszentrum Informatik, 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

Obtaining access to data from platforms remains a challenging process despite DSA Art. 40: 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 OSN and provide legal security to researchers, a European OSN research infrastructure is needed that provides data, supports researchers, and offers secure infrastructure for data processing. This independent research centre should enable unified access for researchers to data from platforms and data donations. It redirects and stores data professionally in line with EU privacy laws, backed by a strong legal team. The centre 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 for improving platform research and enabling transparency, accountability and oversight on platform mechanisms. However, these rights need to be brought to life and effectively enforced, to enable simple access, rather than lengthy court proceedings. Certain legal terminology should be clarified, such as "systemic risks", "public data", and "disinformation". These terms should not be laid out too narrowly to allow basic research. Looking ahead, researchers should be granted access to users' complete view history, rather than merely active interactions with posts (likes, comments, shares), to be able to fully study user behaviour. Furthermore, researchers require access to platforms' internal A/B testing results on the effects of platform mechanics and should in the future 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) in the GDPR-Omnibus must be redrafted: in its current form it would make data donation impossible, as it allows the refusal of access requests for purposes other than the protection of data. Options for users to access their data and donate it 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 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 OSN

To understand the complex mechanisms and effects of OSN, 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**

The EU has in the past years developed strong regulatory frameworks on the digital ecosystem, including the Digital Services Act, Digital Markets Act and 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 on TikTok and Meta by the Commission in line with Art. 66 et seq. DSA, are a strong signal. Trusted Flaggers under the DSA, such as HateAid, who support in removing 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 consumer and European sovereignty and should not be loosened via the Digital Omnibus. To support implementation and democratic oversight of platforms, personnel deployed on this at the Commission should be increased. The success of DSA implementation should be tracked with clear metrics.

### **b) Ensure Transparency and Accountability**

Especially the transparency requirements in the DSA (Art. 15, 24, 42 DSA) should be followed closely, including on 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 hate-speech 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 labelled, complying with Art. 50 AI Act.

### **c) Enforce Interoperability Between Platforms**

Very large online platforms should be forced to ensure interoperability, in line with Art. 6, DMA. This way, people are free to engage with their platform of choice, not forced towards the one with the biggest user base. This is an important base for building and scaling successful alternatives to existing platforms, as network effects are often stopping 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 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 labelled and prioritized. Our TWON simulations and TWONy demonstrators clearly show the strong effects on the spread of opinions of 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 differing perspectives. However, this requires algorithmic transparency.

#### **b) Implement Nudges for Responsible User Behaviour in Platform Design**

Platform design is decisive in incentivizing user behaviour. To foster responsible behaviour, without infringing 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 in 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 not only on active user reactions (likes, comments, shares) should be displayed, but also on passive user interactions with a post (viewed the post without interaction). To limit toxic behaviour and the spread of disinformation, friction could be added, e.g. 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 for adults, as well as for youths is crucial. Design measures limiting addictiveness and compulsive behaviour 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 harmful effects of content and link to mental health help hotlines. To remove the monetary incentive for platforms to keep users online, the right to displaying advertisement could be limited to the first hour per day a person spends on the platform.

### **4. Support the Development of European Public Platforms**

To enable long-term European sovereignty over communication infrastructures, alternatives to existing platforms seem necessary. These should be independent from 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 user needs and societal goals.

Moreover, e-participation tools should be used to enable meaningful participation of citizens in local debates and actual decision-making at different levels of the state.

## 5. Evidence-Based Policymaking

Make use of legal sandboxing tools, such as TWON-as-a-service, to simulate counterfactual scenarios and the effects of specific platform design options on democratic discourse, before implementing legislation. 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. 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 condition for inclusive online participation and dialogue. Knowledge on platform mechanisms and their societal effects are also a prerequisite for digital citizenship and advocacy for the individual political interests on platform regulation.

## 7. Support Independent Journalism

Create a European fund to support independent media outlets that adhere to high-quality standards. This is important to foster fact-based information on OSN, counter misinformation and strengthen public discourse. It is necessary to alleviate the economic pressure on media outlets caused by the shift of advertisement budgets from press to OSN. This fund should be managed by an independent body to ensure transparency and accountability.

## 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 Münker, 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](#)).

## 5. Conclusion

The Citizen Labs substantially advanced the objectives of the TWON project under Task 6.3 by creating participatory spaces in which research perspectives from the different work packages (WPs), lived

experience, and professional practice could be brought into structured exchange. Across four Citizen Labs in different European contexts, the project successfully translated complex socio-technical questions on online social networks, polarisation, misinformation, and platform governance into accessible and discussion-oriented formats. The iterative design of the Labs, moving from workshop-based settings to public and policy-oriented contexts, enabled both continuity and contextual adaptation, while maintaining a shared methodological and thematic core.

Additionally, the iterative structure of the Citizen Labs enabled insights from the discussions to feed back into the TWON project itself. Reflections from participants informed the refinement of the TWONy demonstrators, the user modelling for our simulations, and contributed to the refinement of policy implications throughout the project. In particular, the dialogue with practitioners, civil society actors, and participants with lived experience of online hate and polarisation helped contextualise the research findings and ensured that the emerging policy recommendations were grounded in societal realities.

The Citizen Labs also provided valuable insights into how complex research content can be communicated effectively to non-academic audiences. Feedback across the Labs underscored that tangible and experience-based methods - such as demonstrators and simulation-based exploration - can significantly increase accessibility and the quality of discussion when communicating complex research. At the same time, the Labs confirmed the effort required to balance scientific nuance with clarity for non-specialist audiences, particularly while research and tool development are still ongoing.

A key added value of the Citizen Labs was their contribution to strengthening digital citizenship in practice. Participants engaged with questions of how algorithmic curation, ranking mechanisms, and moderation practices shape public discourse and democratic cohesion, and reflected on rights and responsibilities in digital environments. Through dialogue-based formats and practical experimentation with demonstrators, the Labs supported a deeper understanding of platform dynamics and facilitated informed discussion of regulatory approaches and alternative design principles. Public outreach activities complemented the workshop formats and helped situate the Citizen Labs within broader public and policy debates.

In conclusion, the Citizen Labs significantly advanced the objectives of the TWON project. They meaningfully enriched research relevance, empowered citizens to reflect on and influence platform governance, and strengthened democratic participation in the digital sphere. Well-facilitated participation can strengthen dialogue across sectors and contribute to evidence-informed policy development on digital democracy. The overall experience confirms that embedding research on online social networks in sustained dialogue with diverse European publics is both feasible and highly valuable for future research projects and for the promotion of digital citizenship in Europe.

## 6. Appendix

### 6.1. Agenda 1st Citizen Lab in Karlsruhe, Germany, 16 - 19 September 2024

#### Navigating the Confluence of Real- Life Crisis and Digital Democracy

##### 16 September

**Welcome** (Neta-Paulina Wagner, Programme Manager DialoguePerspectives)

**Welcome Session: The Role of Social Media and AI in the Confluence of Real-Life Crisis and Digital Democracy - A Technical Perspective** (Achim Rettinger (Director FZI Research Center for Information Technology))

##### 17 September

**The Limits of Research on Social Media Dynamics** (Dr. Damian Trilling, University of Amsterdam)

**The Possibilities and Limitations of the EU's Digital Services Act** (Melanie Döring, Project Coordinator "Digital Policy Lab" at Institute for Strategic Dialogue, Berlin (ISD))

**Pre- & De-bunking in the Online Realm** (Marisa Wengeler, Senior Educator Business Council for Democracy at Institute for Strategic Dialogue, Berlin (ISD))

**What's the Impact on Us? Fault Lines, Polarisation and Issues that Impact Dialogue and Pluralism** (Public World Café)

##### 18 September

**Can Two Wrongs make a Right? – An Ethical Reflection on the Idea of Creating Twins of Online Social Networks?** (Eugen Pissarskoi, PhD & researcher at the Department of Sociology at the KIT)

**Europe- Asia and Internet & Democracy – a comparison** (Nuriyatul Lailiyah, PhD & researcher at University Diponegoro, Indonesia)

**How do we Communicate the Needed Changes?** (Neta-Paulina Wagner, Programme Manager DialoguePerspectives)

**Public Bar Camp** @ Tage für Demokratie (Neta-Paulina Wagner, Programme Manager DialoguePerspectives)

##### 19 September

**Policy Brainstorm & Concrete Output** (Neta-Paulina Wagner, Programme Manager DialoguePerspectives)

**Wrap-up and Goodbyes** (Neta-Paulina Wagner, Programme Manager DialoguePerspectives)

## 6.2. Agenda 2nd Citizen Lab in Vienna, Austria, 11 - 14 May 2025

### Free Speech or Free-for-All? Tackling Hate, Misinformation and Polarization in the Age of AI and Tech-Oligarchs

#### 11 May

**Introduction** (Kira Wisniewski, Programme Manager Future 500)

**TWON – The Project, the Goals, the Need** (Dr. Jonas Fegert, FZI Research Center for Information Technology, TWON)

#### 12 May

**The Limits of Research on Social Media Dynamics** (Prof. Dr. Damian Trilling, Vrije Universiteit Amsterdam (VU) and University of Amsterdam (UvA), TWON)

**Algorithmic Amplification of Hate Speech and Misinformation** (Dr. Ing. Even Kapros, Strategic Designer and Researcher on HCI, UX, and Ethics, CEO and founder of Endurae, Strategic Advisor with Project Arc)

**Foreign Influence: Digital Manipulation by Authoritarian States and Paths to Platform Accountability** (Julia Smirnova, Senior Researcher at Center für Monitoring und Analyse, Berlin (CeMAS))

**What's Working, What's Not: Recommender Systems and Platform Accountability** (Dr. Julia Neidhardt, Head of the CD Lab for Recommender Systems, UNESCO Co-Chair for Digital Humanism, TU Vienna)

**What Would the Perfect Online Social Network of the Future Look Like? Requirements for a Democratic Platform Alternative** (Cosima Pfannschmidt, FZI Research Center for Information Technology, TWON)

#### 13 May

**Can AI-Agents replace us? And why this can be Beneficial for Society** (Prof. Dr. Achim Rettinger, Computational Linguistics, Trier University, FZI Research Center for Information Technology, TWON)

**Demonstration of TWONy** (Simon Munker, Trier University, TWON & Dr Fabio Sartori, KIT Karlsruhe Institute of Technology, TWON)

**Voices That Echo: Holocaust Memory, Digital Activism and Peacebuilding in the Social Media Age** (Dr. Stefania Manca, Institute of Educational Technology, Italian National Research Council)

**Public Event “Free Speech or Free-for-All? Tackling Hate, Misinformation and Polarisation in the Age of AI and Tech-Oligarchs”** (Moderation: Dr. Jonas Fegert, FZI Research Center for Information Technology, TWON)

**Panel I: Digital Democracy and the Power of Platforms: Policy, AI, and Accountability** (Prof. Dr. Achim Rettinger, FZI Research Center for Information Technology, TWON, Benjamin Fischer, CeMAS, Judith Peterka, TWON Advisory Board, Natascha Strobl, Political Scientist, Expert on Right-wing Extremism and the New Right, Dr. Sebastian Heidebrecht, Centre for European Integration Research, Vienna University)

**Panel II: Countering Hate and Information Manipulation – Strategies for a Safer Digital Sphere** (Alina Bricman, Director of EU Affairs at B'nai B'rith International, Rosa Jellinek, Activist, Social Media Expert, Selin Aydın, Programme Manager CLAIM-Allianz, Stefania Manca, Institute of Educational Technology, Italian National Research Council)

14 May

**Social Media, AI, Disinformation, and Freedom of Speech** (Nuriyatul Lailiyah, Assistant Professor Communication Department Faculty of Social and Political Sciences Diponegoro University, Semarang Central Java)

**World Café: Local Realities in the Digital Sphere** (IGGÖ - Islamische Glaubensgemeinschaft in Österreich, JöH - Jüdische österreichische Hochschüler:innen, SEEDS - Security Education by Empowering Democratic Strength)

**Policy Prototyping: From Vision to Draft** (Cosima Pfannschmidt, FZI Research Center for Information Technology, TWON & Benjamin Fischer, CeMAS)

**Closing Reflection** (Kira Wisniewski, Programme Manager Future 500)

## 6.3. Agenda 3rd Citizen Lab in Chemnitz, Germany, 7 - 9 October 2025

### Digital Remembrance. Plural Perspectives – Chemnitz Study Trip

7 October

**Dagesh-Studio on the Road #Sukkot Edition, Concert & Artist Talk** (Alex Stolze & Daniel Laufer)

Welcome (Kira Wisniewski, Programme Manager Future 500)

8 October

**Vietnamese Diaspora in Chemnitz: Memories, Histories and Imaginaries** (Theo Döppers, Chemnitz University of Technology)

**TWON Science Tent @ Stadthallenpark** (Simon Münker, TWON, Trier University, Cosima Pfannschmidt, TWON, FZI Research Center for Information Technology, Fabio Satori, TWON, KIT Karlsruhe Institute of Technology, Nils Schwager, TWON, Trier University)

**Joint Lunch with DialoguePerspective's Coalition for Pluralistic Public Discourse** (CPPD)

**Why Memory is Political: Digital & Multimedial Resistance** (Dan Thy Nguyen & Nina Reiprich, Directors of Studio Marshmallow)

Public Event "Practices of Remembrance in Digital Spaces" (Dr Jonas Fegert, FZI, Nhi Le, journalist, Susanne Siegert, social media activist @keineerinnerungskultur, Moderation: Benjamin Fischer, CeMAS), at the Documentation Center on the NSU Complex

9 October

#### Policy Lab

**Making your Voice Heard - Effective Political Correspondence** (Julia Löffler, Policy Advisor Alliance 90/The Greens)

**How Can We Improve Existing Online Social Networks? Recommendations for Improving Platform Mechanics** (Cosima Pfannschmidt, TWON, FZI Research Center for Information Technology)

**Guided Tour through the Documentation Center on the NSU Complex Introduction to the Digital Memorial Space of the Documentation Center on the NSU Complex Curator-led Tour** by Daniel Laufer, Dagesh-Studio  
on the Road – #Sukkot Edition, Zwischen Schutz und Fragilität

**Wrap-Up** (Kira Wisniewski, Programme Manager Future 500)

## **6.4. Agenda 4th Citizen Lab in Brussels, Belgium, 14 – 18 March 2026**

**Rebuilding Bridges: Dialogue, Trust and Solidarity Post October 7th 14 - 18 March 2026 | Brussels, Belgium**

### **14 March**

**Welcome** (Kira Wisniewski, Programme Manager Future 500)

**Asset Mapping – Getting to Know the Group** (Maximiliane Linde, Researcher and Former Board Member of DialoguePerspectives e.V.)

**TWON – The Project, the Goals, the Need** (Cosima Pfannschmidt, FZI Research Center for Information Technology, TWON)

### **15 March**

**From Silence to Words in Social Discourse: How Do We talk Again?** (Great Golberg, Psychologist and Psychotherapist & Lili Zahavi, | Filmmaker and Political Trainer)

**The Limits of Research on Social Media Dynamics** (Prof. Dr. Damian Trilling, Vrije Universiteit Amsterdam (VU) and University of Amsterdam (UvA), TWON)

**Community Work in Conflict and Post-Conflict Settings: Experiences from Eastern Europe (Focus on Ukraine)** (Igor Mitchnik, Executive Director Austausch e.V.)

**Demonstration of TWONy** (Christoph Hau, Trier University, TWON & Simon Munker, Trier University)

### **16 March**

**How Polarisation Works: Recognising and Interpreting Political Narratives** (Atahan Demirel, Policy Advisor on Anti-Discrimination in the Berlin House of Representatives)

**Invoking the Past, Legitimising the Present: The Use of Collective Memory in Digital Diplomacy** (Researcher and Former Board Member of DialoguePerspectives e.V.)

**Historicising the Present: Antisemitism and Racism in Israel-Palestine Social Media Discourses** (Furkan Yüksel, Lecturer in Historical-Political Education)

**From Social to AI Media - Personalised AI Agents for the Detriment or Benefit of Online Dialogue?** (Prof. Dr. Achim Rettinger, Computational Linguistics, Trier University, FZI Research Center for Information Technology, TWON)

**Public Event “Voices Rising: Rebuilding Bridges. Dialogue, Trust and Solidarity Post-October 7th** (Camila Piaastro, European Union of Jewish Students, Barbara von Freytag, Journalist, Political Analyst, Furkan Yüksel, Political Educator, Prof. Dr. Achim Rettinger (Trier University, TWON, Moderation: Igor Mitchnik, Executive Director, Austausch e.V.)

### 17 March

**Together Instead of Against Each Other: An Attempt to a Differentiated Dialogue on the Israeli-Palestinian conflict (A German Perspective)** (Sophie Orentlikher, Socio-Political Educator and Clinical Social Worker, Centre for Applied Research on Education and Diversity at the Catholic College of Aachen & Mohammed-Arfan Ashmawi, Social Educator and Clinical Social Worker, Research Assistant at the Centre for Applied Research on Education and Diversity at the Catholic College of Aachen)

**Media Literacy in a Digital World** (Amie Liebowitz, Journalist, Broadcast Presenter and Media Consultant  
World Café

**Public Evening Event: From Research to Sovereignty in the Digital Public Sphere – Insights from the TWON Project**

### 18 March

**TWON Hackathon: From Utopia to Policy - Reimagining Online Social Networks** (Cosima Pfannschmidt, FZI Research Center for Information Technology, TWON)

**Wrap-Up** (Kira Wisniewski, Programme Manager Future 500)

## **6.5. Workshop Results: Requirements for a Democratic Platform Alternative (2nd Citizen Lab, Vienna)**

During a workshop entitled “What Would the Perfect Online Social Network of the Future Look Like? Requirements for a Democratic Platform Alternative” at the Vienna Citizen Lab, participants engaged in structured small-group discussions to imagine an ideal online social network (OSN) that adheres to democratic values and actively safeguards democratic societies and discourse. The workshop - led by Cosima Pfannschmidt, FZI - invited participants to reflect on purpose, governance, content moderation, algorithms, ownership, and mental health aspects of online social networks. Other TWON researchers present at the Citizen Lab also joined the discussions. The results provide insight into citizens’ normative expectations and practical requirements for future platform design, and provide use cases for simulations to conduct with the TWON tool.

Across groups, participants expressed a strong desire for online social networks that serve both personal and societal functions. On a personal level, the ideal OSN should enable connections with peers, hobby communities, and professional networks. It should provide space for communication, dialogue, discussion, coordination, and organization. Participants emphasized the importance of being heard and of having access to education, career opportunities, science exchange, fundraising, and even dating. At the societal level, the platform should function as a trustworthy news aggregator, foster exposure to diverse perspectives, and support societal cohesion through community-building. Several groups stressed the importance of

“discovering new realities” and being constructively exposed to “the other side,” rather than remaining within homogeneous networks. While entertainment was considered legitimate, participants warned against dopamine-driven short-form formats and addictive design features. Functions that should clearly not be part of a democratic OSN included microtargeting, mobilization for destructive purposes, hate speech, bullying, stalking, foreign information manipulation and interference (FIMI), and any form of polarizing content. There was also strong resistance to the idea of a single organization or government acting as the ultimate arbiter of “truth.”

In discussing hate speech and disinformation, participants converged on the need for a multi-actor governance model. Decisions on content moderation should not be left exclusively to companies or governments. Instead, suggestions included the creation of an independent, non-governmental body, advisory teams composed of accredited institutions, NGOs and civil society organizations, and mechanisms for rotation to avoid concentration of power. Fact-checking should be conducted by established and credible institutions, ideally in a decentralized and transparent structure. Automation was seen as a necessary support tool to assist human reviewers, particularly to pre-screen large volumes of content, but not as a fully autonomous decision-maker. Sanctions for harmful content were considered necessary, yet enforcement should involve both companies and regulators, with civil society playing a role in oversight and accountability. Participants debated anonymity: while non-anonymous participation could reduce disinformation and abuse, anonymity was recognized as essential for protecting vulnerable groups. This tension was not resolved, but highlighted as a key design trade-off.

Regarding content, participants supported a diverse mix including entertainment, journalistic, academic, and civic content, alongside limited advertising. Political content should be diversified but not artificially amplified. AI-generated content and bots should be clearly labeled, and the use of filters, text or voice alterations should be transparently indicated. Influencer and creator content should be regulated more clearly, particularly where commercial interests are involved. Several groups suggested separating certain functions, such as shopping, from core social interaction spaces, while ensuring interoperability between platforms to facilitate user migration and avoid lock-in effects. Customizable feeds by category (e.g., business, education, entertainment) were widely supported. Participants called for stronger user control over content ranking and for prioritizing educational long-form content over purely attention-grabbing short formats. Proposals included screen-time warning indicators (“brain rot” alerts) and clearer boundaries for advertising, potentially through subscription-based or hybrid models with higher-quality and restricted ads.

The discussion on recommendation algorithms revealed a strong demand for transparency, user agency, and safety-by-design principles. Many participants argued that users should have meaningful control over their algorithmic feed and even the possibility to choose between different algorithms. However, concerns were raised that full personalization could trap individuals in “rabbit holes.” To counter this, participants proposed built-in “truthfulness filters” to downrank misinformation and safeguards to ensure exposure to diverse viewpoints. Core ranking criteria should include truthfulness, safety, diversity, localization (e.g., language awareness), and the avoidance of traumatic or excessively harmful content. Emotionality and engagement metrics alone should not determine visibility. Transparency about how algorithms function was repeatedly described as essential, and oversight mechanisms involving users and independent actors were seen as necessary.

On platform ownership and governance, participants favored decentralized, cooperative, or public-interest-oriented models over purely profit-driven corporations. Suggestions ranged from co-owned public-private structures to supranational European governance frameworks. Preventing monopolies was seen as critical; anti-trust measures, strong regulation of very large online platforms, and lighter-touch regulation for smaller platforms were proposed to balance innovation and competition. Interoperability, ease of migration, and

unified log-on systems were identified as structural requirements for a healthy digital ecosystem. Participants also called for a “codex” for digital platforms that would apply at both structural and individual levels, combined with stronger and timelier enforcement of existing rules, comparable to GDPR enforcement. Some groups proposed incentive-based systems, such as distributing funds based on innovation and compliance rates, and reducing fines in cases of immediate compliance.

Mental health and well-being emerged as a central design principle. Participants advocated for non-addictive algorithms that do not optimize for maximum time spent online. One concrete policy idea suggested limiting advertisement revenue to the first hour of daily usage to remove economic incentives for excessive engagement. Platforms should provide reminders after prolonged usage, including mental health check-ins (“Are you okay?”), prompts to take breaks, and suggestions for offline activities. Transparency about content moderation and algorithmic functioning was also seen as contributing to well-being. Further proposals included disabling public visibility of “likes,” offering different protective functions for vulnerable groups, providing links to helplines and support services, and combating malicious deepfakes. Participants emphasized that solutions must combine personal responsibility (e.g., mindfulness) with societal regulation and structural platform design changes, particularly in user interface and user experience (UI/UX) design that nudge toward healthier engagement patterns. Reward-based mechanisms were often preferred over purely punitive approaches.

Overall, the workshop results reveal a coherent vision of democratic online social networks grounded in plural governance, transparency, user agency, interoperability, and mental health protection. Participants consistently emphasized that platform design is not neutral but shapes democratic culture, social cohesion, and individual well-being. The discussions demonstrate that citizens are capable of articulating nuanced positions on complex trade-offs between freedom, safety, innovation, and regulation. Their input provides valuable guidance for future research, platform design considerations, and policy recommendations within the TWON project.



## Contact us

**Damian Trilling**

*Project Coordinator*

☎ +31 62 782 7904

✉ [d.c.trilling@uva.nl](mailto:d.c.trilling@uva.nl)

📍 University of Amsterdam  
Postbus 15791  
1001 NG Amsterdam



Funded by  
the European Union