



**Feedback to the European Commission
on the “European Democracy Shield”**

Call for evidence - Ares(2025)2555098

May 23, 2025

Comments of:

The Future of Free Speech
2817 West End Ave #126 – 629
Nashville, TN 37203
www.futurefreespeech.org



The Future of Free Speech is an independent, nonpartisan think tank located at Vanderbilt University. We work to reaffirm freedom of expression as the bedrock of free and thriving societies through actionable research, empowering tools, and principled advocacy. The Future of Free Speech seeks to create a world where everyone’s right to freedom of expression is protected by law and reinforced by a culture that tolerates diverse viewpoints.

About this Feedback

This submission reflects the views of The Future of Free Speech and was drafted by Alexander Hohlfeld, external consultant to The Future of Free Speech, with input and review by Jacob Mchangama (Executive Director), Jordi Calvet-Bademunt (Senior Research Fellow), and Joan Barata (Senior Legal Fellow).

I. Introduction

We appreciate that the call for evidence highlights from the outset that “[t]he essence of democracy is that citizens can freely express their views and participate in democratic life, choose their political representatives, and have a say in their future”. As a non-partisan think tank advocating for freedom of expression, we would like to point out the pitfalls the proposed “European Democracy Shield” could pose to the right to freedom of expression, as enshrined in Article 11 of the Charter of Fundamental Rights of the European Union and Article 10 of the European Convention on Human Rights.

As we continuously emphasize, we are convinced that “in the digital age, it’s impossible to effectively shield citizens from hostile propaganda and disinformation without compromising the egalitarian and liberal values of democracy” (Mchangama, 2022b).

While we share your concerns about democratic backsliding, we caution that many measures taken against “disinformation” and “foreign information manipulation” could further contribute to this development instead of combatting it. As we recently demonstrated in mapping laws and regulations affecting freedom of expression in 22 democracies, we are witnessing a “global free speech recession”, in which “free speech restrictions are not just on the rise in authoritarian and semi-authoritarian states, but also in the liberal democracies”. Preventing “disinformation” is one of the most common grounds for speech-restrictive developments (The Future of Free Speech, 2023). We are very pleased to have the opportunity to contribute our views and hope that our concerns will be considered in the further process.

This submission structures as follows: the first part contextualizes general implications around tackling “disinformation” and “foreign information manipulation” (Chapter II-III). The second part evaluates parts of the current political framework in the EU (Chapter IV) before concluding with some general recommendations for the “European Democracy Shield” (Chapter V).



II. Contextualizing “Disinformation” and “Foreign Information Manipulation”

The call for evidence cites recent survey findings, noting that "EU citizens perceive the two most significant threats to be growing distrust and scepticism towards democratic institutions (36%) and false and/or misleading information in general circulating online and offline (34%)". However, perceived risks do not necessarily reflect actual risks. Therefore, it is crucial to consider research on the prevalence and impact of foreign information manipulation and interference, as well as disinformation. There is a significant discrepancy between policy and media discourses on the one hand and academic research on the other. While the risk perceptions about misinformation seem to be highest in liberal democracies, they seem unrelated to the prevalence of such content in the country (Knuutila et al., 2022). Empirical studies consistently show that the prevalence and impact of “disinformation” are often overstated and limited in scope (Acerbi et al., 2022; Allen et al., 2020; Altay et al., 2023; Bail et al., 2020; Budak et al., 2024; Eady et al., 2023; Fletcher et al., 2018; Golovchenko, 2020; González-Bailón et al., 2024; Grinberg et al., 2019; Guess et al., 2019, 2020; Jungherr & Schroeder, 2021). Researchers even emphasized that “not a single scientific study has demonstrated that beliefs in conspiracy theories and misinformation have, generally, increased over time” (Enders & Uscinski, 2022) and many scholars argue that the whole discourse around “disinformation” could be characterized as a “moral panic” (e.g. Anderson, 2021; Carlson, 2020; Hoffmann, 2023; Jungherr & Schroeder, 2021; Mitchelstein et al., 2020). Even the “EEAS East Stratcom Task Force” concluded that for the 2019 European Election, “available evidence has not allowed to identify a distinct cross-border disinformation campaign from external sources specifically targeting the European elections” (European Commission, 2019).

Research into AI driven “influence operations” also found no impact of AI on election results (Stockwell, 2024), and researchers emphasize that fears about the impact of generative AI on elections are generally overblown (Kapoor & Narayanan, 2024; Simon et al., 2023, 2024). This is not surprising, since decades of research on media research already emphasized that information does not function like “magical bullets” and opinions cannot be easily manipulated (Farkas & Schou, 2024; Mercier, 2020). The reasoning behind the conclusion in the Call for Evidence, that “recent experience has shown an intensification and wider spread of these threats”, remains obscure in this evidence.

The high risk perception among the population is thus more likely to stem from exaggerated warnings about disinformation in media and policy discourses than the actual risk of “disinformation”. Since the perceived threat of "disinformation" seems to impact the satisfaction with democracy negatively (Jungherr & Rauchfleisch, 2022; Nisbet et al., 2021), the regulatory and media attention risks exacerbating the very issues it intends to mitigate.

Risk perceptions of “disinformation” are also strongly dependent on the so-called “third person” effect, believing others would be gullible for “disinformation” and “manipulation” but not themselves (Altay & Acerbi, 2023; Corbu et al., 2020). Ringing the alarm bell on "disinformation" or "foreign information manipulation and interference" could then lead to "casting divergent



opinions as a result of manipulation" and thus "weaken the support for crucial elements of liberal democracy and the legitimacy of electoral defeat" (Jungherr & Rauchfleisch, 2022).

Moreover, warnings about "disinformation", particularly in "elite discourse" (Van Duyn & Collier, 2019), also seem to lower trust in media in general and reduce the perceived authenticity of factually accurate information (Hameleers, 2023; van der Meer et al., 2023; Van Duyn & Collier, 2019). This is particularly concerning since people generally seem "able to discern true from false news and erred on the side of skepticism rather than credulity" (Pfänder & Altay, 2025).

We believe that the initiatives to counter "disinformation" and "foreign information manipulation" are carried out in the best faith, but would like to second the warning that "unbalanced warnings against perceived threats of disinformation on democracy might turn out to be the actual danger" (Jungherr & Rauchfleisch, 2022). Supposed widespread gullibility was used as an argument against democratic power across history, but decades of research show that people are way less gullible than usually assumed (Mercier, 2020). Thus, since one of the aims of the "European Democracy Shield" is to strengthen citizens' trust in democracy, democratic institutions, and trust in electoral outcomes, the focus on "disinformation" and "information manipulation" might damage this trust more than it strengthens it.

"Any evidence-based response to the problem of misperceptions must thus begin with an effort to counter misinformation about the problem itself. Only then can we design interventions that are proportional to the severity of the problem and consistent with the values of a democratic society" (Nyhan, 2020, p. 233).

III. Combatting "disinformation" as a threat to freedom of expression

The Call for Evidence defines "disinformation" as "false or misleading content that is spread with an intention to deceive or secure economic or political gain, and which may cause public harm". While such a definition seems logically appealing, in reality, neither the truthfulness, intention, nor potential harm can be neutrally assessed in most cases. Accordingly, the term "disinformation" is not only legally unsuitable (Appelman et al., 2022; Fathaigh et al., 2021) but also bears the intrinsic danger of being appropriated as a means of suppressing critical or divergent voices in political discourse. This tendency manifests at the individual level (Michael & Breaux, 2021; Ribeiro et al., 2017) and even more concerning at the political level. As noted by the former UN Special Rapporteur on Freedom of Expression the term "disinformation" "is an extraordinarily elusive concept to define in law, susceptible to providing executive authorities with excessive discretion to determine what is disinformation, what is a mistake, what is truth" (Kaye, 2020, p. 13).

However, even in clear-cut cases of "disinformation," censorship poses significant risks to democratic discourse and freedom of expression. This reflects a foundational principle of freedom of expression, as articulated by the successor Special Rapporteur Irene Khan:



“[[f]reedom of expression includes the freedom to seek, receive and impart information and ideas of all kinds, true or false, offensive or enlightened, regardless of frontiers and in any media of one's choice” (Khan, 2022, p. 11).

This principle appears increasingly under threat within the EU. In particular, the objective raised by Commission President von der Leyen in her speech on the “European Democracy Shield”, that “once we have detected malign information or propaganda, we need to ensure that it is swiftly removed and blocked” (von der Leyen, 2024) would impose significant constraints on the principle of freedom of expression and information and substantially conflicts with the raised objective in the speech to protect the values of democracy.

IV. Analysis of the Existing Political Framework

The ban of Russian State Media

In light of the empirical research outlined above, it is questionable whether the “Regulation concerning restrictive measures in view of Russia's actions destabilising the situation in Ukraine” (Council Regulation (EU) 2022/350 /Council Decision (CFSP) 2022/351) are compatible with international human rights law, particularly the right to freedom of expression and information as enshrined in Article 19 of the UN Declaration on human rights. We have particular doubts about whether the regulation meets strict tests of necessity and proportionality and whether the full ban is the “least intrusive instrument”. As previously noted “[t]here ought to be space to allow war critics, ordinary citizens, and scholars to debunk myths and counter disinformation” (Mendiratta & Alkiviadou, 2022). Since freedom of expression protects not only the right of the speaker but also the receiver of information, whether true or false, the ban affects every European citizen who seek to critically engage with or understand the narratives disseminated by Russian state media. Additionally, the decision provides a rationale and legitimacy for other countries to increase censorship or ban Western media, as it recently happened in Russia (Mchangama, 2022b; Mendiratta & Alkiviadou, 2022). It is important to recall that “for all its harms and costs, free speech and access to information is a competitive advantage when democracies engage in information wars with authoritarians” (Mchangama, 2022a).

The Digital Services Act (DSA)

Even though the DSA only mentions “disinformation” in its Recitals, “disinformation” appears to constitute a central enforcement priority with respect to the systemic risk obligations (Art. 34 & 35). While the DSA has been carefully drafted and, according to many experts, cannot require platforms to take down legally protected speech under the systemic risk obligations (Husovec, 2024), it is highly concerning that some enforcement actions appear to have gone beyond the DSA's legal boundaries and pressured platform providers to take down non-illegal content labelled as “disinformation”. These concerns were addressed in an Open Letter to you in August last year (The Future of Free Speech, 2024).



While Recitals 86 and 90 of the DSA require particular consideration to the impact on freedom of expression in their risk assessments, we observed a general lack of such considerations in most Risk Assessments published by the platforms. On the contrary, platform risk assessments focused heavily on content-related risks (Chapman, 2025; DSA Civil Society Coordination Group, 2025), with particularly strong reference to “disinformation” related risks. The problem here is that the DSA provides “no clarity regarding the consideration of specific safeguards to avoid excessive and inadequate restrictions on users’ rights to freedom of expression” (Barata, 2024, p. 16)

Even more worrying, many of the mitigations suggested in the “Code of Practice on Disinformation”, which recently got transferred to a Code of Conduct under the DSA have serious implications for the right to freedom of expression, which are not taken into sufficient consideration.

The Code of Conduct on Disinformation

The Code of Conduct on Disinformation is built on the contradiction between the note that “fundamental rights must be fully respected in all actions taken to fight Disinformation” and that a “balance [...] must be struck between protecting fundamental rights and taking effective action to limit the spread and impact of otherwise lawful content”. Thus, it assumes that fundamental rights must be restricted to limit the spread and impact of lawful content. However, lawful content should remain beyond the reach of state interference and limiting the reach can constitute an act of censorship. The High-Level-Expert Group on Disinformation also recommended that “[a]ny form of censorship either public or private should clearly be avoided” (HLEG, 2018). The Code of Conduct on Disinformation violates this principle. Censorship does not only consist of removal of content; deliberately reducing the reach can also constitute an act of censorship. A comparison with real-world counterparts reveals this clearly:

“In well-functioning democracies, limiting the reach of a newspaper by fixing its selling price through law or limiting the number of copies that may be printed of a given book or magazine would be as unconstitutional as censoring it.” (Del Campo et al., 2024)

Platforms fulfilled the objective of limiting the reach of specific legal content in pre-emptive compliance in a way that significantly endangers freedom of expression. As explained above, truthfulness and intent are almost impossible to assess in most cases. Thus, many platforms entirely refrained from considering the intention and further outsourced the responsibility of assessing “truthfulness” to fact-checking organizations. Based on their judgements, platforms like Meta have reduced the reach of certain posts. TikTok even removed videos that fact-checkers could not verify completely from their recommendation system. This grants enormous power to fact-checkers and has a similar effect to outright removal since it is estimated that about 90-95% of the views on TikTok come from recommendations (WSJ, 2021). Limiting the reach is a direct intervention in seeking and receiving information and thus negatively affects the right to freedom of expression. Once again, it is helpful to compare it with an example from other media:



“When a person asks for another person to send them messages, and a third party intervenes to block those messages, that is censorship. Even if you call it “freedom of reach,” it’s still censorship.” (Doctorow, 2022)

Thus, to respect freedom of expression, “as a general rule, disseminating disinformation, propaganda, and similar types of undesirable speech must be tackled via non-repressive mechanisms such as media ethics, communication policies, reinforcement of public service media, or promoting media pluralism” (Barata & Calvet-Bademunt, 2023). We recommend refraining from approaches requiring providers of online platforms to deliberately reduce the reach of certain content since it sincerely affects legally protected speech, and the assessment is often arbitrary.

Fact-Checking Initiatives

Despite its prominence in the Call for Evidence, fact-checking should not be viewed as a silver bullet against disinformation. A former Meta employee, who interviewed 20 trust and safety professionals, pointed out that the “fact-checking regime was not about truth. Instead, it evolved into a desperate stopgap and PR tactic that was doomed from the start” and concluded that “[t]he collapse of social media fact-checking is an admission that top-down truth enforcement doesn’t work at internet scale” (Wagner, 2025). The problem with fact-checking is that “fact-checkers” are not free from “confusion, bias, and arbitrariness” (Uscinski, 2023) and “often attempt to check statements that are not facts and cannot be verified as true or false” (Uscinski & Butler, 2013). Unsurprisingly, fact-checkers have issued numerous contested or erroneous decisions in recent years (Mchangama, 2025). Freedom of expression advocates argued that fact-checking “censored truthful information and, worse, reduced public confidence in the information contained on social media” (Lukianoff & Goldstein, 2025).

Building a “European network of fact-checkers” as it is proposed in the Political Guidelines for the European Commission 2024-2029, is thus unlikely to advance the stated aims.

Since “[t]he epistemic strength of open societies is the ability to engage arguments and viewpoints in public and settle on the most plausible” (Jungherr, 2024, p. 10), a shortcut of imposing the most plausible argument top-down does not reflect this epistemic strength. Democracies should instead focus on more participatory, deliberative models, that leverage the potential of debate, different viewpoints and free speech instead of attempting to suppress it.

V. Key Recommendations for Safeguarding Freedom of Expression in the European Democracy Shield

From this, we derive the following conclusions for the European Democracy Shield:

-
- I. *Refrain from exaggerated warning about “disinformation” and “foreign information manipulation”*
-



Exaggerated warnings can have the opposite effect and contribute to the crisis it aims to mitigate. As pointed out, it is not necessarily “disinformation” itself that undermines trust in democracy and the acceptance of electoral results, but rather exaggerated warnings in media and policy discourse. While according to academic research the prevalence and impact of “disinformation” is rather low, unfounded warnings lead people to view other opinions as a result of manipulation which can lower the acceptance of the democratic process.

II. Rebuild mutual trust through communication, transparency and promoting a free-speech culture.

Most researchers agree that the spread of “disinformation” is rooted in a deeper crisis: a lack of trust in institutions. But trust is a two-way street. Communication that treats people as gullible or easily manipulated only deepens mistrust. Rebuilding trust requires institutions to demonstrate trust in citizens through transparency, openness, and respectful dialogue

Recent measures such as the ban of Russian State Media or efforts to deliberately reduce the spread of certain information have undermined this goal. These actions risk reinforcing the issue of distrust they aim to combat and should be re-evaluated in light of their long-term implications for democratic legitimacy.

While Taiwan has been the “largest target of foreign information disinformation for the last ten years” (Mchangama, 2024), their approach of dealing with “disinformation” provides crucial lessons and can be seen as an inspiration for liberal democracies. Rather than censoring speech or blocking access, Taiwan embraced openness, transparency and empowerment of citizens. As behavioral scientist Tzu-wei Hung emphasized, “Free speech is not the cost but the key to counteract disinformation” (Aspinwall, 2025).

The “Open Government National Action Plan” (National Development Council, 2025) offers many insights for the proposed “European Democracy Shield”. First and foremost is the emphasis on promoting “Open Data and Freedom of Information”. Trust in institutions cannot simply be taken for granted but needs to be earned daily through transparency, accountability and inclusion. By “radically trusting citizens” (Bertelsmann Stiftung, 2020), as former Taiwanese Minister of Digital Affairs Audrey Tang advocated, Taiwan has strengthened its resilience to disinformation: “By embracing transparency and encouraging public participation, Taiwan strengthens its defenses against propaganda and disinformation, which thrive when information is opaque and people lack confidence in one another and their public institutions” (Caster, 2024). The lesson Taiwan holds for Europe is, that “[w]hen you radically trust citizens, citizens will trust you back” (Curtis, 2024). Crucially this includes trusting citizens also in the exposure to “disinformation” or “foreign manipulation information” and the absence of restrictions on their “freedom to seek, receive and impart information and ideas of all kinds”.



III. Refrain from top-down approaches in “fact-checking” and encourage crowdsourced alternatives

Traditional approaches of fact-checking and their current embedding in platform architecture do not reflect the epistemic strength of open societies and freedom of expression as we pointed out. That’s why we recommend focusing on crowdsourced alternatives using the “wisdom of crowds” (Allen et al., 2021). Such systems “combine the participatory, open and interactive nature of social media with features that encourage thoughtful listening, an understanding of conversational dynamics and the careful emergence of an understanding of shared views and points of rough consensus” (Weyl et al., 2024, p. 283)

Crowd-sourced approaches leverage the potential and power of democratic debate and freedom of expression instead of suppressing it. This reflects in higher trust: Researchers found that “[a]cross both sides of the political spectrum, community notes were perceived as significantly more trustworthy than simple misinformation flags” (Drolsbach et al., 2024). In contrast to traditional fact-checking, crowdsourced approaches “encourage the exploration of diverse political information” (Weyl et al., 2024, p. 278) and due to their design they can even “harness partisan motives” to increase the volume and quality of notes (Martel & Allen, 2024). Many crowdsourced approaches have safeguards in place, to correct biases by taking support of diverse user groups into account (Weyl et al., 2024, p. 278).

IV. Support spaces for democratic discourse, viewpoint diversity and competing values

The heart of democracy is not truth but popular sovereignty, which rests on the foundation of freedom of expression and access to information. Thus, “[s]aving democracy is [...] not about arming against fake news and disinformation – at least not primarily. It is instead, and perhaps more importantly, about creating genuine spaces for politics: that is, spaces for contestation, for political difference, and for pluralism” (Farkas & Schou, 2024). The next European Initiative should put these principles in the spotlight in order to safeguard European democracies.

V. Include freedom of expression experts in decision-making

The Call for Evidence mentions that “the Commission is planning to present an EU civil society strategy” that “will support, protect and empower civil society organizations and human rights defenders and will put forward actions to step up the Commission’s engagement with them and to ensure that they are better protected in their work”. Given the impact the European Democracy Shield can have on freedom of expression and the concerns raised in relation to related measures



like the DSA, the European Commission should ensure that freedom of expression advocates are systematically consulted. This would reduce the risk that the European Democracy Shield has unintended effects on fundamental rights.

References

- Acerbi, A., Altay, S., & Mercier, H. (2022). Research note: Fighting misinformation or fighting for information? *Harvard Kennedy School Misinformation Review*. <https://doi.org/10.37016/mr-2020-87>
- Allen, J., Arechar, A. A., Pennycook, G., & Rand, D. G. (2021). Scaling up fact-checking using the wisdom of crowds. *Science Advances*, 7(36), eabf4393. <https://doi.org/10.1126/sciadv.abf4393>
- Allen, J., Howland, B., Mobius, M., Rothschild, D., & Watts, D. J. (2020). Evaluating the fake news problem at the scale of the information ecosystem. *Science Advances*, 6(14), eaay3539. <https://doi.org/10.1126/sciadv.aay3539>
- Altay, S., & Acerbi, A. (2023). People believe misinformation is a threat because they assume others are gullible. *New Media & Society*, 14614448231153379. <https://doi.org/10.1177/14614448231153379>
- Altay, S., Berriche, M., & Acerbi, A. (2023). Misinformation on Misinformation: Conceptual and Methodological Challenges. *Social Media + Society*, 9(1), 20563051221150412. <https://doi.org/10.1177/20563051221150412>
- Anderson, C. W. (2021). Fake News is Not a Virus: On Platforms and Their Effects. *Communication Theory*, 31(1), 42–61. <https://doi.org/10.1093/ct/qtaa008>
- Appelman, N., Dreyer, S., Bidare, P. M., & Potthast, K. C. (2022, May 16). Truth, intention and harm: Conceptual challenges for disinformation-targeted governance. *Internet Policy Review*. <https://policyreview.info/articles/news/truth-intention-and-harm-conceptual-challenges-disinformation-targeted-governance/1668>
- Aspinwall, N. (2025, May 20). Taiwan Learned You Can't Fight Fake News by Making It Illegal. *Foreign Policy*. <https://foreignpolicy.com/2024/01/16/taiwan-election-china-disinformation-lai/>
- Bail, C. A., Guay, B., Maloney, E., Combs, A., Hillygus, D. S., Merhout, F., Freelon, D., & Volfovsky, A. (2020). Assessing the Russian Internet Research Agency's impact on the political attitudes and behaviors of American Twitter users in late 2017. *Proceedings of the National Academy of Sciences*, 117(1), 243–250. <https://doi.org/10.1073/pnas.1906420116>
- Barata, J. (2024). *The Future of Free Speech: Old Threats and New Challenges* (SSRN Scholarly Paper No. 4982214). Social Science Research Network. <https://doi.org/10.2139/ssrn.4982214>
- Barata, J., & Calvet-Bademunt, J. (2023, October 30). *The European Commission's Approach to DSA Systemic Risk is Concerning for Freedom of Expression*. Tech Policy Press. <https://techpolicy.press/the-european-commissions-approach-to-dsa-systemic-risk-is-concerning-for-freedom-of-expression/>



- Bertelsmann Stiftung. (2020). *Digital Democracy—What Europe can learn from Taiwan*. Bertelsmann Stiftung. <https://www.bertelsmann-stiftung.de/doi/10.11586/2022043>
- Budak, C., Nyhan, B., Rothschild, D. M., Thorson, E., & Watts, D. J. (2024). Misunderstanding the harms of online misinformation. *Nature*, 630(8015), 45–53. <https://doi.org/10.1038/s41586-024-07417-w>
- Calvet-Bademunt, J. (2025, April 2). The AI Election Panic: How Fear-Driven Policies Could Limit Free Expression | TechPolicy.Press. *Tech Policy Press*. <https://techpolicy.press/the-ai-election-panic-how-feardriven-policies-could-limit-free-expression>
- Carlson, M. (2020). Fake news as an informational moral panic: The symbolic deviancy of social media during the 2016 US presidential election. *Information, Communication & Society*, 23(3), 374–388. <https://doi.org/10.1080/1369118X.2018.1505934>
- Caster, M. (2024, January 20). Confronting Digital Authoritarianism Through Digital Democracy: Lessons From Taiwan. *The Diplomat*. <https://thediplomat.com/2024/01/confronting-digital-authoritarianism-through-digital-democracy-lessons-from-taiwan/>
- Chapman, P. (2025, January 9). Advancing Platform Accountability: The Promise and Perils of DSA Risk Assessments. *Tech Policy Press*. <https://techpolicy.press/advancing-platform-accountability-the-promise-and-perils-of-dsa-risk-assessments>
- Corbu, N., Oprea, D.-A., Negrea-Busuioc, E., & Radu, L. (2020). ‘They can’t fool me, but they can fool the others!’ Third person effect and fake news detection. *European Journal of Communication*, 35(2), 165–180. <https://doi.org/10.1177/0267323120903686>
- Curtis, P. (2024, July 22). How Taiwan bucked a global trend – and restored voters’ trust in politics. *The Guardian*. <https://www.theguardian.com/commentisfree/article/2024/jul/22/taiwan-bucked-global-trend-trust-politics-hired-protesters>
- Del Campo, A., Zara, N., & Ugarte, R. Á. (2024). *Are Risks the New Rights? The Perils of Risk-based Approaches to Speech Regulation*. CELE. unpublished confidential draft
- Doctorow, C. (2022, December 10). *Pluralistic: Freedom of reach IS freedom of speech (10 Dec 2022) – Pluralistic: Daily links from Cory Doctorow*. <https://pluralistic.net/2022/12/10/e2e/>
- Drolsbach, C. P., Solovev, K., & Pröllochs, N. (2024). Community notes increase trust in fact-checking on social media. *PNAS Nexus*, 3(7), pgae217. <https://doi.org/10.1093/pnasnexus/pgae217>
- DSA Civil Society Coordination Group. (2025, March). *Initial Analysis on the First Round of Risk Assessments Reports under the EU Digital Services Act*.
- Eady, G., Paskhalis, T., Zilinsky, J., Bonneau, R., Nagler, J., & Tucker, J. A. (2023). Exposure to the Russian Internet Research Agency foreign influence campaign on Twitter in the 2016 US election and its relationship to attitudes and voting behavior. *Nature Communications*, 14(1), 62. <https://doi.org/10.1038/s41467-022-35576-9>



- Enders, A. M., & Uscinski, J. (2022, December 16). What's the Harm? An interrogation of the societal impact of conspiracy theories and potential remedies. *Knight First Amendment Institute at Columbia University*. <https://knightcolumbia.org/content/whats-the-harm>
- European Commission. (2019). *Action Plan against disinformation: Report on progress*. Publications Office of the European Union. https://commission.europa.eu/document/download/c1ae5566-f115-429d-b1af-99a7ef1f72c2_en?filename=factsheet_disinfo_elex_140619_final.pdf
- Farkas, J., & Schou, J. (2024). *Post-truth, fake news and democracy: Mapping the politics of falsehood* (Second edition). Routledge.
- Fathaigh, R. Ó., Helberger, N., & Appelman, N. (2021). The perils of legally defining disinformation. *Internet Policy Review*, 10(4). <https://policyreview.info/articles/analysis/perils-legally-defining-disinformation>
- Fletcher, R., Cornia, A., Graves, L., & Nielsen, R. K. (2018). *Measuring the reach of “fake news” and online disinformation in Europe*. Reuters Institute. <https://reutersinstitute.politics.ox.ac.uk/our-research/measuring-reach-fake-news-and-online-disinformation-europe>
- Golovchenko, Y. (2020). Measuring the scope of pro-Kremlin disinformation on Twitter. *Humanities and Social Sciences Communications*, 7(1), 1–11. <https://doi.org/10.1057/s41599-020-00659-9>
- González-Bailón, S., Lazer, D., Barberá, P., Godel, W., Allcott, H., Brown, T., Crespo-Tenorio, A., Freelon, D., Gentzkow, M., Guess, A. M., Iyengar, S., Kim, Y. M., Malhotra, N., Moehler, D., Nyhan, B., Pan, J., Rivera, C. V., Settle, J., Thorson, E., ... Tucker, J. A. (2024). The Diffusion and Reach of (Mis)Information on Facebook During the U.S. 2020 Election. *Sociological Science*, 11, 1124–1146. <https://doi.org/10.15195/v11.a41>
- Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., & Lazer, D. (2019). Fake news on Twitter during the 2016 U.S. presidential election. *Science*, 363(6425), 374–378. <https://doi.org/10.1126/science.aau2706>
- Guess, A., Nagler, J., & Tucker, J. (2019). Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Science Advances*, 5(1), eaau4586. <https://doi.org/10.1126/sciadv.aau4586>
- Guess, A., Nyhan, B., & Reifler, J. (2020). Exposure to untrustworthy websites in the 2016 US election. *Nature Human Behaviour*, 4(5), Article 5. <https://doi.org/10.1038/s41562-020-0833-x>
- Hameleers, M. (2023). The (Un)Intended Consequences of Emphasizing the Threats of Mis- and Disinformation. *Media and Communication*, 11(2), 5–14. <https://doi.org/10.17645/mac.v11i2.6301>
- HLEG. (2018). *A multi-dimensional approach to disinformation: Report of the independent High level Group on fake news and online disinformation*. Publications Office of the European Union.



- Hoffmann, C. P. (2023). The Role of Moral Panics in Media Transformation: An Examination of the “Techlash.” In A. Godulla & S. Böhm (Eds.), *Digital Disruption and Media Transformation: How Technological Innovation Shapes the Future of Communication* (pp. 41–54). Springer International Publishing. https://doi.org/10.1007/978-3-031-39940-4_4
- Husovec, M. (2024). The Digital Services Act’s red line: What the Commission can and cannot do about disinformation. *Journal of Media Law*, 16(1), 47–56. <https://doi.org/10.1080/17577632.2024.2362483>
- Jungherr, A. (2024). Foundational questions for the regulation of digital disinformation. *Journal of Media Law*, 16(1), 8–17. <https://doi.org/10.1080/17577632.2024.2362484>
- Jungherr, A., & Rauchfleisch, A. (2022). *Negative downstream effects of disinformation discourse: Evidence from the US* [Preprint]. SocArXiv. <https://osf.io/a3rzm>
- Jungherr, A., & Schroeder, R. (2021). Disinformation and the Structural Transformations of the Public Arena: Addressing the Actual Challenges to Democracy. *Social Media + Society*, 7(1), 2056305121988928. <https://doi.org/10.1177/2056305121988928>
- Kapoor, S., & Narayanan, A. (2024, December 13). We Looked at 78 Election Deepfakes. Political Misinformation Is Not an AI Problem. *Knight First Amendment Institute*. <http://knightcolumbia.org/blog/we-looked-at-78-election-deepfakes-political-misinformation-is-not-an-ai-problem>
- Kaye, D. (2020). *Disease pandemics and the freedom of opinion and expression. Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression* (No. A/HRC/44/49). https://www.ohchr.org/Documents/Issues/Opinion/A_HRC_44_49_AdvanceEditedVersion.docx
- Khan, I. (2022). *A/77/288: Disinformation and freedom of opinion and expression during armed conflicts—Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression* (No. A/77/288). <https://www.ohchr.org/en/documents/thematic-reports/a77288-disinformation-and-freedom-opinion-and-expression-during-armed>
- Knuutila, A., Neudert, L.-M., & Howard, P. N. (2022). Who is afraid of fake news? Modeling risk perceptions of misinformation in 142 countries. *Harvard Kennedy School Misinformation Review*. <https://doi.org/10.37016/mr-2020-97>
- Lukianoff, G., & Eduardo, A. (2024, March 20). The misinformation crisis isn’t about truth, it’s about trust. *The Eternally Radical Idea*. <https://eternallyradicalidea.com/p/the-misinformation-crisis-isnt-about>
- Lukianoff, G., & Goldstein, A. (2025, January 10). METACOMMENTARY: Why Meta’s move to community notes is good. *The Eternally Radical Idea*. https://eternallyradicalidea.com/p/metacommentary-why-metas-move-to?utm_campaign=post



- Martel, C., & Allen, J. N. L. (2024). *Harnessing partisan motives to combat misinformation*. <https://osf.io/3ngbt/>
- Mchangama, J. (2022a, March 14). The Problem With Banning Russian Disinformation. *The Daily Beast*. <https://www.thedailybeast.com/the-problem-with-banning-russian-disinformation/>
- Mchangama, J. (2022b, August 12). In A War of Ideas, Banning Russian Propaganda Does More Harm Than Good. *TIME*. <https://time.com/6205645/russian-propaganda-censorship-history/>
- Mchangama, J. (2024, April 5). How to Fight Misinformation Without Censorship. *Persuasion*. <https://www.persuasion.community/p/how-to-fight-misinformation-without>
- Mchangama, J. (2025, January 9). Don't Mourn the Fact-Checkers. *Persuasion*. <https://www.persuasion.community/p/dont-mourn-the-fact-checkers>
- Mendiratta, R., & Alkiviadou, N. (2022, April 15). Banning Russia Today and Sputnik in Europe is a bad idea. *Human Rights Here*. <https://www.humanrightshere.com/post/banning-russia-today-and-sputnik-in-europe-is-a-bad-idea>
- Mercier, H. (2020). *Not born yesterday: The science of who we trust and what we believe*. Princeton University Press.
- Michael, R. B., & Breaux, B. O. (2021). The relationship between political affiliation and beliefs about sources of “fake news.” *Cognitive Research: Principles and Implications*, 6(1), 6. <https://doi.org/10.1186/s41235-021-00278-1>
- Mitchelstein, E., Matassi, M., & Boczkowski, P. J. (2020). Minimal Effects, Maximum Panic: Social Media and Democracy in Latin America. *Social Media + Society*, 6(4), 2056305120984452. <https://doi.org/10.1177/2056305120984452>
- National Development Council. (2025). *Open Government National Action Plan*. https://www.ndc.gov.tw/en/Content_List.aspx?n=0DA7FCB068C7ECF5
- Nisbet, E. C., Mortenson, C., & Li, Q. (2021). The presumed influence of election misinformation on others reduces our own satisfaction with democracy. *Harvard Kennedy School Misinformation Review*. <https://doi.org/10.37016/mr-2020-59>
- Nyhan, B. (2020). Facts and Myths about Misperceptions. *Journal of Economic Perspectives*, 34(3), 220–236. <https://doi.org/10.1257/jep.34.3.220>
- Pfänder, J., & Altay, S. (2025). Spotting false news and doubting true news: A systematic review and meta-analysis of news judgements. *Nature Human Behaviour*, 1–12. <https://doi.org/10.1038/s41562-024-02086-1>
- Ribeiro, M. H., Calais, P. H., Almeida, V. A. F., & Meira Jr, W. (2017). “Everything I Disagree With is #FakeNews”: Correlating Political Polarization and Spread of Misinformation (No. arXiv:1706.05924). arXiv. <https://doi.org/10.48550/arXiv.1706.05924>
- Simon, F. M., Altay, S., & Mercier, H. (2023). Misinformation reloaded? Fears about the impact of generative AI on misinformation are overblown. *Harvard Kennedy School Misinformation Review*. <https://doi.org/10.37016/mr-2020-127>



Simon, F. M., McBride, K., & Altay, S. (2024, September 3). AI's impact on elections is being overblown. *MIT Technology Review*.

<https://www.technologyreview.com/2024/09/03/1103464/ai-impact-elections-overblown/>

Stockwell, S. (2024). *AI-Enabled Influence Operations: Threat Analysis of the 2024 UK and European Elections*. <https://cetas.turing.ac.uk/publications/ai-enabled-influence-operations-threat-analysis-2024-uk-and-european-elections>

The Future of Free Speech. (2023). *The Free Speech Recession Hits Home. Mapping Laws and Regulations Affecting Free Speech in 22 Open Democracies*. <https://futurefreespeech.org/the-free-speech-recession-hits-home/>

The Future of Free Speech. (2024, August 21). *Open Letter to Thierry Breton on The DSA's Threats to Free Speech*. <https://futurefreespeech.org/open-letter-to-thierry-breton-on-the-dsas-threats-to-free-speech/>

Uscinski, J. E. (2023). What Are We Doing When We Research Misinformation? *Political Epistemology*, 2, 2–14. https://www.academia.edu/97403059/Political_Epistemology_No_2

Uscinski, J. E., & Butler, R. W. (2013). The Epistemology of Fact Checking. *Critical Review*, 25(2), 162–180. <https://doi.org/10.1080/08913811.2013.843872>

van der Meer, T. G. L. A., Hameleers, M., & Ohme, J. (2023). Can Fighting Misinformation Have a Negative Spillover Effect? How Warnings for the Threat of Misinformation Can Decrease General News Credibility. *Journalism Studies*, 0(0), 1–21.



<https://doi.org/10.1080/1461670X.2023.2187652>

Van Duyn, E., & Collier, J. (2019). Priming and Fake News: The Effects of Elite Discourse on Evaluations of News Media. *Mass Communication and Society*, 22(1), 29–48.

<https://doi.org/10.1080/15205436.2018.1511807>

von der Leyen, U. (2024, May 14). *A Strong Europe in the World*. Copenhagen Democracy Summit, Copenhagen. <https://admin.epp.eu/files/uploads/2024/05/Ursula-von-der-Leyen-Democracy-Summit-Copenhagen-Denmark-14-May-2024.pdf>

Wagner, L. (2025, March 3). How Tech Created the Online Fact-Checking Industry. *Pirate Wires*. <https://www.piratewires.com/p/how-tech-created-the-online-fact-checking-industry>

Weyl, E. G., Tang, A., & Community, . (2024).  數位 Plurality: *The Future of Collaborative Technology and Democracy*. Independently published. <https://www.plurality.net/>

WSJ. (2021, July 21). Inside TikTok's Algorithm: A WSJ Video Investigation. *Wall Street Journal*. <https://www.wsj.com/articles/tiktok-algorithm-video-investigation-11626877477>