



AI4Debunk



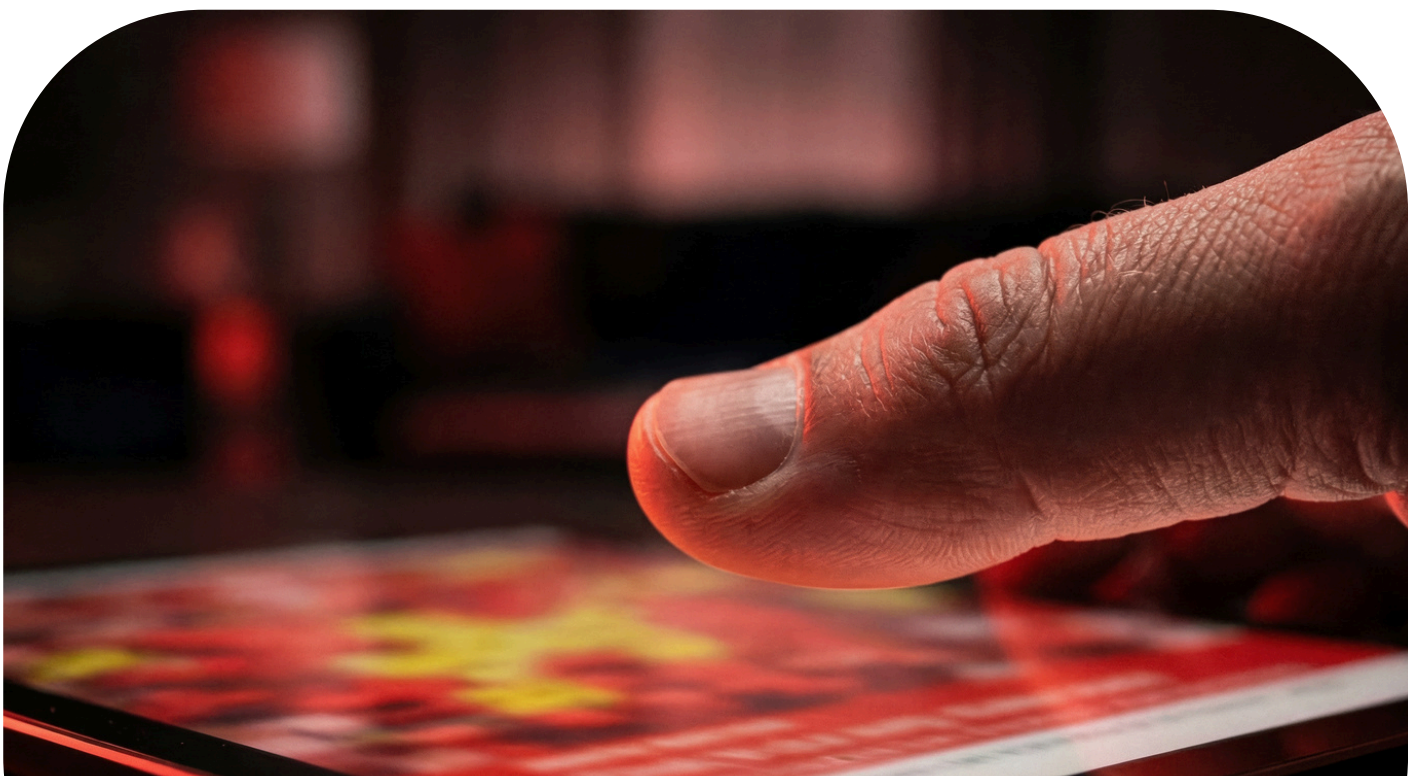
Alessia D'Andrea
Consiglio Nazionale
delle Ricerche



Arianna D'Ulizia
Consiglio Nazionale
delle Ricerche

Putting the Pieces Together: Why Linguistic, Media & AI Dynamics, and EU Policy Need to Come Together to Debunk Disinformation

In recent years, disinformation has gone from the occasional shots of "fake news" to a systemic threat that uses linguistic manipulation, media mechanisms, artificial intelligence, and transnational political strategies to form public opinion and undermine democratic processes. Two strands of research, one dealing with the "micro" mechanisms of disinformation and the other with high-level policy and governance responses, have often remained separate. It is time to bring them together.



The Anatomy of the Threat: Language, Media Mechanisms, and generative AI

Disinformation campaigns are achieved through various tiers, which are interrelated, making it very effective (D'Andrea et al., 2025a). First, at the level of language, manipulative content is used, which is based on emotions and polarization, affecting perceptions rather than informing. It is then spread through the various mechanisms of the media and the platform, which are based on the cognitive biases of the users, who are drawn to sensational and polarizing content. Furthermore, the appearance of generative AI makes this problem particularly important, as it is possible to produce a great deal of very believable false content, which is very difficult to detect (Shukla & Tripathi, 2024; Pilati & Venturini, 2025; Shoaib et al., 2023). This, of course, has very important consequences for the public sphere, affecting political opinions and decision-making processes.

The Role of AI: As Risk and as Tool

AI is a double-edged sword. On the one hand, generative AI facilitates large-scale disinformation campaigns (Shoaib et al., 2023; Feuerriegel et al., 2023). On the other hand, AI tools can detect, track, and counter manipulative content if used in a responsible manner (Curtis et al., 2025; Truică & Apostol, 2022; Leite et al., 2025). Techniques include cross-lingual detection, network analysis, or automated fact-checking pipelines (Peña-Alonso et al., 2025; Shukla & Tripathi, 2024). Moreover, explainable AI systems can contribute to the detection of deepfakes and the analysis of network diffusion patterns (Curtis et al., 2025; Truică & Apostol, 2022).

Why Policy Alone is not Enough — But Policy is Still Essential

In terms of policy, the European Union has put in place programs like the European Digital Media Observatory (EDMO) and the Digital Services Act to increase media literacy, accountability, and transparency among its member states (European Commission, 2024; Sánchez Gonzales et al., 2024; D'Andrea et al., 2025b). Collaborative networks that include fact checkers, journalists, and academic institutions are integral in counterbalancing this at scale (Frau Meigs et al., 2025). They are usually initiatives that respond to the problem and, thus, control the consequences of disinformation without addressing its cognitive, linguistic, or algorithmic roots (Leite et al., 2025). The policies that are formed without being backed by research in domains such as emotional framing, persuasive strategies, and AI-driven amplification have the risk of being too broad or scattered.

Merging Micro and Macro-Toward a Holistic Strategy

Effective mitigation of disinformation must be addressed both at the micro and at the macro levels:

- **Micro-level vigilance**
 - Citizens must become capable of spotting manipulative narratives, emotionally framed and/or algorithmically amplified content (Arribas et al., 2025).
- **Macro-level governance**
 - This would necessitate regulatory frameworks that provide for transparency, accountability of platforms, access to data to allow for independent analysis, and the overall coordination between EU member states and fact-checking networks (Pilati & Venturini, 2025; European Commission, 2024).

These levels need to be in conversation with one another: policy design should be informed by research on what makes disinformation persuasive, and research should be scaled via structural policy tools.

Toward Democratic Resilience: What Should We Do?

Three interlocking priorities emerge:

1. Encourage interdisciplinary research to help identify novel tactics, techniques, and procedures of disinformation that are continuously evolving to improve detection across languages, cultures, and platforms (D'Andrea et al., 2025a; Arribas et al., 2025).
2. Commit to large-scale media literacy and critical thinking, embedding it within schools, community programs, and public campaigns (Frau-Meigs et al., 2025; SánchezGonzales et al., 2024).
3. Introduce and enforce structural safeguards, such as due-transparency obligations for online platforms, crisis-response mechanisms, labelling of AI-generated content, and cooperative EU-wide oversight (European Commission, 2024; Pilati & Venturini, 2025).



Conclusion

Disinformation cannot be defeated by research or regulation alone; a more holistic approach that draws insight from linguistics, AI, media studies, and policy is essential. This is only through this integration at the level of micro understandings of persuasion with macro governance and regulatory measures that societies may build resilience to safeguard democratic processes (Shukla & Tripathi, 2024; Farooq & de Vreese, 2025).

References

1. Arribas, C. M., Gertrudix, M., & Arcos, R. (2025). Preventive Strategies Against Disinformation: A Study on Digital and Information Literacy Activities Led by Fact-Checking Organisations. *Open Research Europe*, 5, 122.
2. Curtis, T. L., Touzel, M. P., Garneau, W., Gruaz, M., Pinder, M., Wang, L. W., ... & Pelrine, K. (2025). Veracity: An Open-Source AI Fact-Checking System. *arXiv preprint arXiv:2506.15794*.
3. D'Andrea, A., Fusacchia, G., & D'Ulizia, A. (2025a). Linguistic insights, media mechanisms and role of AI in dissemination and impact of disinformation. *Journal of Information, Communication and Ethics in Society*.
4. D'Andrea, A., Fusacchia, G., & D'Ulizia, A. (2025b). Policy review: countering disinformation in the digital Age-Policies and initiatives to safeguard democracy in Europe. *Information Policy*, 30(1), 82-91.
5. European Commission. (2024). True or false? How to defend yourself against disinformation. Commission news. https://commission.europa.eu/news-and-media/news/true-or-false-how-defend-yourself-against-disinformation-2024-10-23_en?
6. Farooq, A., & de Vreese, C. (2025). Deciphering authenticity in the age of AI: how AI-generated disinformation images and AI detection tools influence judgements of authenticity. *AI & SOCIETY*, 1-12.
7. Feuerriegel, S., DiResta, R., Goldstein, J. A., Kumar, S., Lorenz-Spreen, P., Tomz, M., & Pröllochs, N. (2023). Research can help to tackle AI-generated disinformation. *Nature Human Behaviour*, 7(11), 1818-1821.
8. Frau-Meigs, D., Corbu, N., & Osuna-Acedo, S. (2025). Fighting disinformation through media literacy: From current practices to policy making in Europe. *InMedia. The French Journal of Media Studies*, (10).
9. Leite, J. A., Razuvayevskaya, O., Scarton, C., & Bontcheva, K. (2025, May). A cross-domain study of the use of persuasion techniques in online disinformation. In *Companion Proceedings of the ACM on Web Conference 2025* (pp. 1100-1103).
10. Peña-Alonso, U., Peña-Fernández, S., & Meso-Ayerdi, K. (2025). Journalists' perceptions of artificial intelligence and disinformation risks. *Journalism and Media*, 6(3), 133.
11. Pilati, F., & Venturini, T. (2025). The use of Artificial Intelligence in counter-disinformation: A world wide (web) mapping. *Frontiers in Political Science*, 7, 1517726.
12. Sánchez Gonzales, H. M., González, M. S., & Alonso-González, M. (2024). Artificial intelligence and disinformation literacy programmes used by European fact-checkers. *Catalan Journal of Communication & Cultural Studies*, 16(2), 237-255.
13. Shukla, A. K., & Tripathi, S. (2024). AI-generated misinformation in the election year 2024: measures of European Union. *Frontiers in Political Science*, 6, 1451601.
14. Shoaib, M. R., Wang, Z., Ahvanooy, M. T., & Zhao, J. (2023, November). Deepfakes, misinformation, and disinformation in the era of frontier AI, generative AI, and large AI models. In *2023 international conference on computer and applications (ICCA)* (pp. 1-7). IEEE.
15. Truică, C. O., & Apostol, E. S. (2022). Misrobaerta: transformers versus misinformation. *Mathematics*, 10(4), 569.



Alessia D'Andrea

Consiglio Nazionale delle Ricerche

IRPPS Researcher. Her research focuses on web technologies, human-computer interaction, social sciences, marketing/virtual business, and health.



Arianna D'Ulizia

Consiglio Nazionale delle Ricerche

IRPPS researcher. Her research interests include social computing and human-computer interaction, particularly natural and multimodal language processing and user-centered interaction design.

SUPPORTED BY



Funded by
the European Union

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 or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.