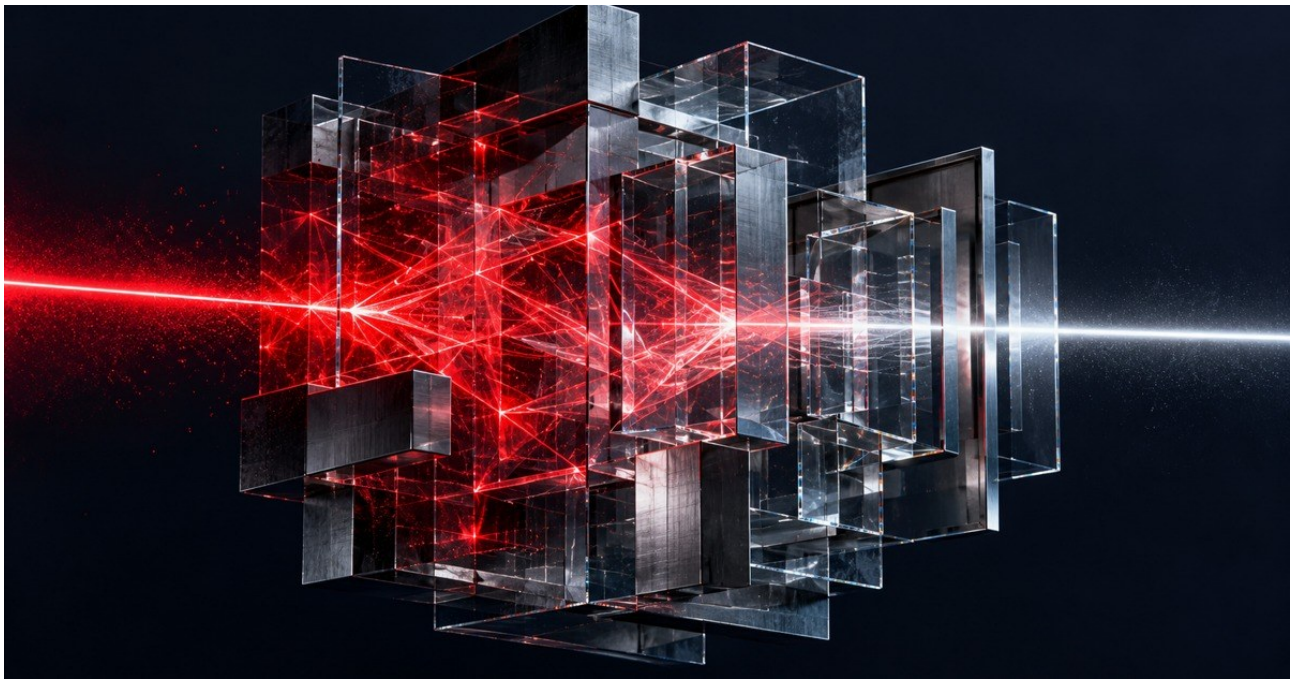
DISINFORMATION & DIGITAL RESILIENCE / EN

The architecture of deception: a systemic analysis of information disorder in the digital age

How platforms, the attention economy and cognitive vulnerabilities shape disinformation. Ivan-Oleksandr Yakymiv examines manipulation, generative AI and ways to build digital resilience.

Ivan-Oleksandr Yakymiv | 28.10.2026



Author analysis · edited from the author's text dated 28 December 2025, updated on 28 October 2026 with clarifications of research scope and regulation. Editorial illustrations are visual metaphors for the concepts discussed, rather than records of real events.

Foundations of contemporary disinformation

TRUTH editorial illustration · Foundations of contemporary disinformation

The digital transformation has created a public sphere in which social motivations, platform design and cognitive vulnerabilities interact. This article uses “the architecture of deception” as an analytical perspective on that interaction, rather than a claim that every digital service necessarily produces falsehood. Information disorder includes problems of truth, intent, context and distribution. Treating it only as a collection of isolated lies makes it harder to understand how misleading narratives acquire visibility and credibility. [1]

Three layers are especially useful. Society supplies political interests, economic incentives and the desire for group belonging. Technology supplies platforms, data analytics, bots and AI, making communication scalable and selectively visible. Psychology supplies emotions, attention limits and confirmation bias, which influence interpretation and sharing. Their interaction may favour manipulation, but outcomes depend on the service, its safeguards and the behaviour of its users.

Disinformation as a complex sociotechnical system

TRUTH editorial illustration · Disinformation as a complex sociotechnical system

A simple sender–receiver model cannot capture all the actors involved in an influence operation. The people who devise a narrative may differ from those who package it as media or distribute it. State agencies, commercial operators, organised communities and ordinary users can contribute at different stages. An apparently spontaneous wave of attention can therefore have a much more complicated origin.

Feedback loops help explain this architecture: engagement produces behavioural signals, which can inform subsequent recommendations and distribution. However, a loop is a conceptual model, not a validated universal equation predicting belief. Nor should filter bubbles be treated as an inevitable experience of all internet users. Recommendation systems, personal choices and social networks interact differently in different settings. Claims about effects on elections or public health require context-specific evidence. [1]

Three faces of information disorder

TRUTH editorial illustration · Three faces of information disorder

Wardle and Derakhshan distinguish misinformation, disinformation and malinformation. Misinformation involves false information shared without an intention to cause harm. Disinformation combines falsehood with deliberate harmful intent. Malinformation uses genuine information to cause harm, for example by exposing private information. These distinctions make it possible to discuss more than whether a statement is factually correct. [1]

Creation, production and distribution should also be distinguished. The originator of a message may not be the person who turns it into a media product or shares it publicly. Intent can change along that chain: someone may innocently circulate material designed by another actor to mislead. True material can likewise be stripped of context and incorporated into a false narrative. Effective analysis therefore asks who acted, at what stage and with what evidence of intent.

From ancient rhetoric to print

TRUTH editorial illustration · From ancient rhetoric to print

Manipulation predates digital technology. Yet rhetoric should not be equated with deception. Aristotle examined the means of persuasion while also emphasising truth, justice and reasoned argument. His Rhetoric offers tools for understanding communication; it does not establish that truth is irrelevant to persuasion. [3]

The growth of print made messages easier to reproduce and circulate beyond immediate institutional control. Pamphlets, polemics and eventually a competitive press expanded both access to knowledge and opportunities for misleading persuasion. Sensational communication could attract audiences, but historical media should not be treated as identical to contemporary recommendation systems. The continuity concerns interests and persuasion; the infrastructure and scale are different.

The industrialisation of persuasion

TRUTH editorial illustration · The industrialisation of persuasion

Twentieth-century mass communication brought organised propaganda into specialised state and commercial institutions. Radio and cinema made it possible to address large audiences at the same time. During the Cold War, influence operations also sought to erode trust in institutions and undermine an opponent's credibility, rather than simply win agreement with a single proposition.

A related contemporary mechanism is saturation: repeated and contradictory messages can exhaust an audience's ability or willingness to verify them. Paul and Matthews describe a Russian "firehose of falsehood" model characterised by multiple channels, rapid repetition and a lack of commitment to consistency. This is a specific analytical model, not a complete account of every propaganda campaign. It suggests why an effective response may need to address distribution patterns as well as individual false claims. [4]

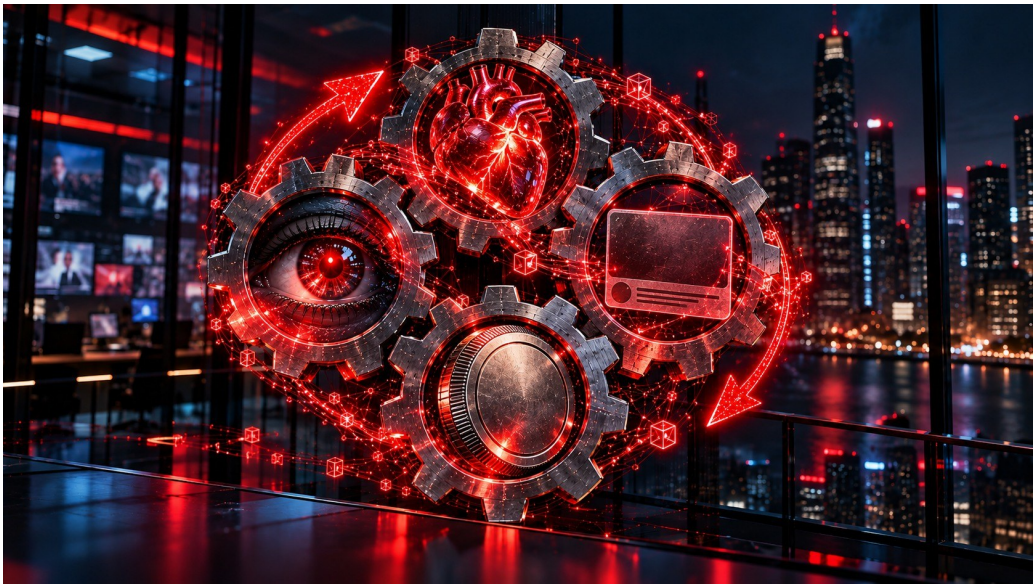
Web 2.0 and computational propaganda

TRUTH editorial illustration · Web 2.0 and computational propaganda

Networked communication expanded participation in publishing. Users with mobile devices became nodes in a global distribution system, while platforms acquired considerable power over visibility. A message can now circulate through a mixture of personal relationships, recommendation systems and coordinated activity, making the boundary between organised influence and everyday participation difficult to trace.

Computational propaganda uses automated systems, algorithms and human coordination to shape political communication. Bots and organised accounts may inflate apparent popularity or push a narrative towards wider visibility. Such activity can create an impression of spontaneous public support, but automation alone does not prove that a campaign has changed people's beliefs. Research must examine both the organisation of distribution and its actual effects. [5]

The attention economy

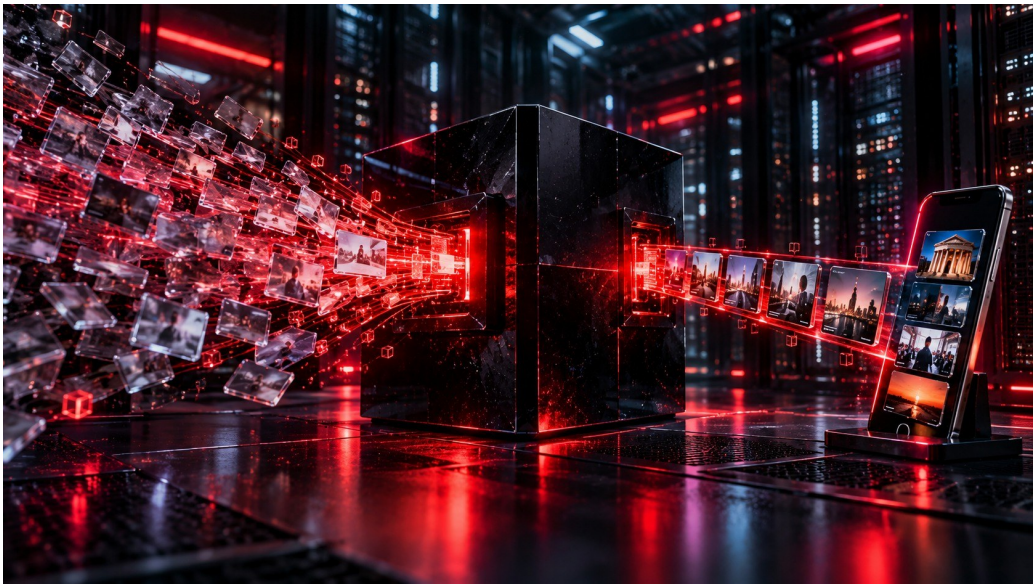


TRUTH editorial illustration · The attention economy

Many advertising-funded services compete for attention. Commercial value can depend on activity and time spent, which can create tensions between an engaging message and an accurate one. This does not mean platforms never account for quality. It means their incentives and evaluation criteria deserve scrutiny when assessing how misleading material becomes visible.

Recommender systems use many signals, including engagement. In some circumstances, emotionally charged material may benefit from this arrangement. That is a reason to investigate objectives, safeguards and outcomes, rather than assume that every algorithm always favours falsehood. Systemic analysis asks whether the distribution environment rewards harmful manipulation and what changes would reduce that incentive without undermining legitimate expression.

Surveillance capitalism and algorithmic opacity



TRUTH editorial illustration · Surveillance capitalism and algorithmic opacity

Shoshana Zuboff's theory of surveillance capitalism provides a critical framework for understanding the collection and commercial use of behavioural data. Such data can be used to predict behaviour and attempt to influence it. The framework should not be confused with proof that profiling gives an actor unlimited power over an individual. [11]

Opaque recommendation systems make independent scrutiny harder. A user may not understand why a message appeared or which interests shaped its distribution. Better explanations, appropriate research access and evaluation of actual outcomes can improve accountability. At the same time, personalised persuasion should not be assumed effective merely because it is technically possible: its influence must be measured rather than inferred from the existence of data.

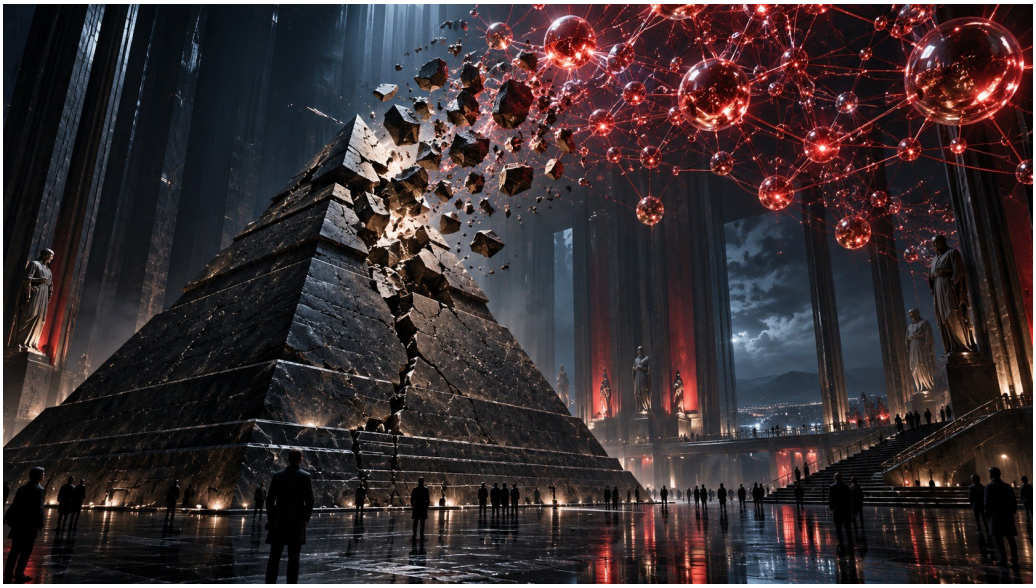
Human beings as amplifiers

TRUTH editorial illustration · Human beings as amplifiers

Vosoughi, Roy and Aral studied roughly 126,000 verified stories circulated on Twitter between 2006 and 2017. Within that dataset, false stories spread farther and faster than true ones. The result concerns a particular historical platform and sample; it is not a universal measurement of the entire internet or of present-day X. [2]

The researchers examined novelty and emotional responses as possible explanations. False stories were more novel in their analysis, and replies more often expressed fear, disgust or surprise. These patterns do not prove that every false story is more viral. Bots also accelerated both true and false information similarly in the study, so they did not explain the observed difference. This does not imply that bots are harmless in other settings. Confirmation bias and group identity remain useful concepts, but should not replace investigation of specific cases. [2]

Authority, knowledge and institutional trust

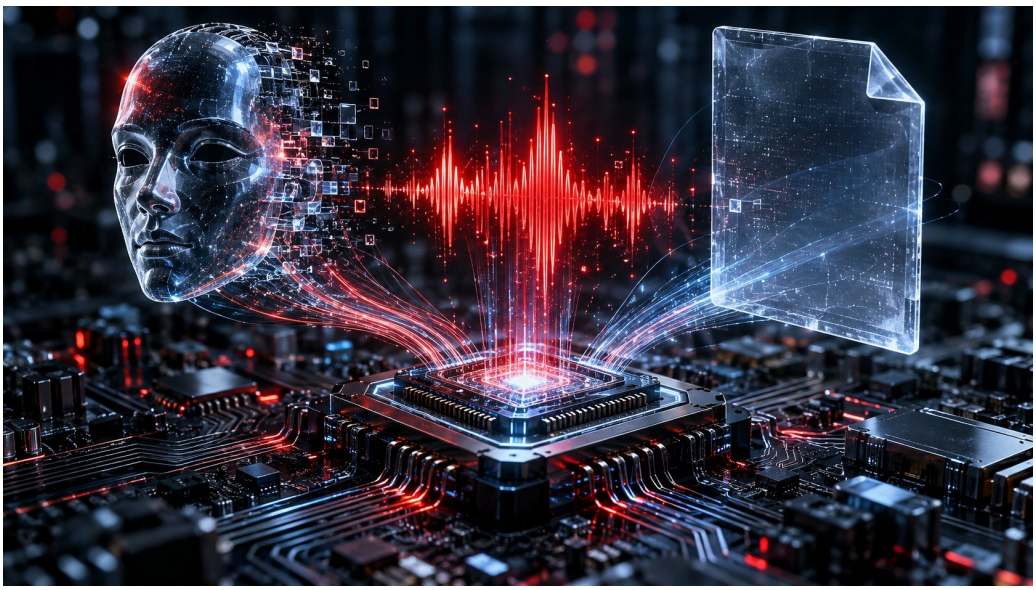


TRUTH editorial illustration · Authority, knowledge and institutional trust

Information disorder can weaken the shared evidential basis for public discussion. Popularity may be mistaken for expertise, while universities, newsrooms and scientific institutions may be portrayed as merely partisan actors. The danger is not just a mistaken belief about one event but the loss of agreed procedures for testing competing claims.

Trust should not require treating institutions as infallible. Transparent methods, corrections and independent scrutiny make authority accountable. In a crisis, people need to distinguish uncertainty within a legitimate investigation from the claim that no reliable knowledge is possible. Rebuilding that distinction is a practical task for journalism, education and public communication.

Generative AI and synthetic media

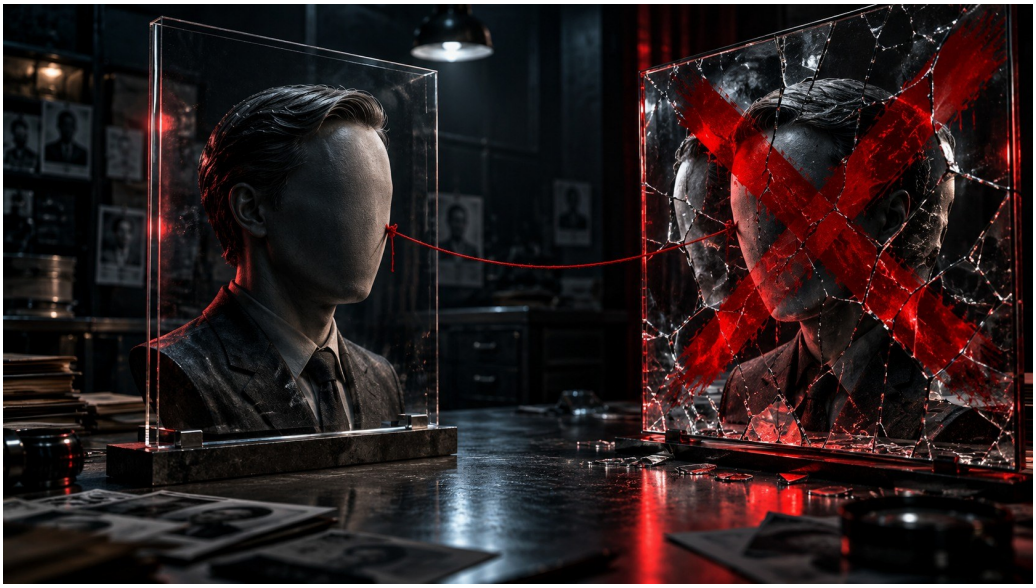


TRUTH editorial illustration · Generative AI and synthetic media

Generative tools can lower the cost of producing persuasive text, images, audio and video. They may help impersonate a public figure, manufacture statements or scale manipulative interactions. The fact that content is synthetic does not itself make it deceptive; the problem concerns false attribution, misleading context and harmful use.

Risks include coordinated conversations that simulate grassroots activity and unreliable content entering datasets used by AI systems. These are reasons to investigate provenance and data quality, rather than assume every model automatically reproduces every narrative it encounters. AI can also assist verification by finding patterns and tracing distribution. No detector guarantees authenticity: automated findings still require context, corroboration and examination of the original material.

The liar's dividend



TRUTH editorial illustration · The liar's dividend

Chesney and Citron describe the "liar's dividend": awareness of convincing digital fabrication can help someone deny authentic evidence. A person need not produce a fake to benefit from the public's uncertainty about whether a genuine recording is real. The possibility of deception becomes a resource for avoiding scrutiny. [6]

This risk does not establish that denial always works better than apology or silence. Such comparisons require specific experimental evidence. A more robust response is to preserve originals, document provenance and seek independent corroboration. Public education should teach how to evaluate evidence, rather than promote the blanket belief that nothing seen online can be trusted.

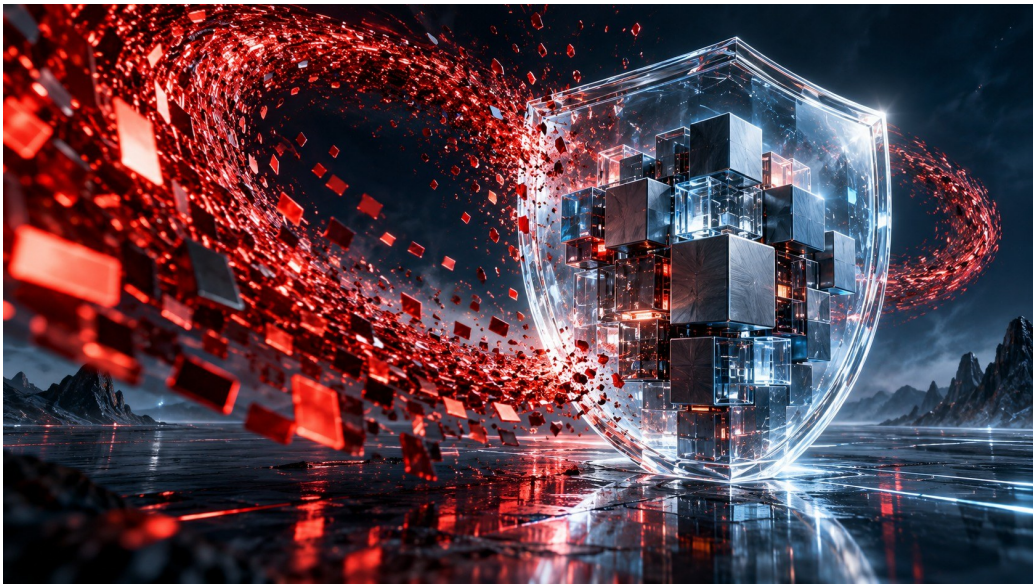
Digital Services Act and platform accountability

TRUTH editorial illustration · Digital Services Act and platform accountability

The EU Digital Services Act establishes transparency and accountability obligations whose scope varies by service category and size. Very large platforms and search engines face additional systemic-risk assessment and mitigation duties. It does not make every false statement unlawful or guarantee that misleading information will disappear. [7]

The regulation covers transparency of advertising and recommender parameters, independent audits for the largest services and conditional data access for vetted researchers. Very large services using recommendation systems must offer at least one option not based on profiling. This is not a universal ban on personalised advertising. DSA obligations also address deceptive interfaces, with scope limits and interaction with other laws. Accountability must remain compatible with fundamental rights, including freedom of expression. [7]

Prebunking and psychological inoculation

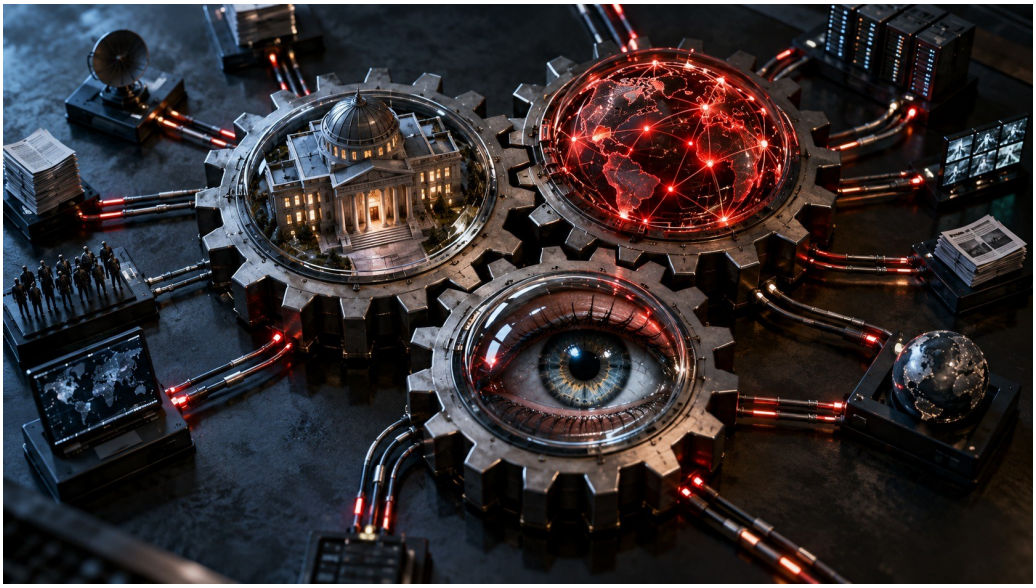


TRUTH editorial illustration · Prebunking and psychological inoculation

Psychological inoculation prepares people to recognise manipulation before encountering it in a persuasive setting. Short educational materials can demonstrate emotional manipulation, false dilemmas or scapegoating. This complements fact-checking: recognising a technique does not by itself establish whether a factual claim is true. [8]

Experimental research reports improvements on some resilience measures, but effects depend on intervention design, measurement and time. Work comparing inoculation and accuracy prompts also shows why recognising manipulation should not be confused with reliably judging real headlines. Combining approaches can help, while a one-off lesson should not be advertised as permanent immunity. Effective programmes need reinforcement, independent evaluation and attention to both true and false material. [8] [9] [10]

Towards an ecosystem of trust



TRUTH editorial illustration · Towards an ecosystem of trust

A structural problem needs several kinds of response. Platform accountability, independent journalism, research and media education address different parts of the same environment. Governments, civil society, researchers and services should be able to cooperate without removing the independent scrutiny that makes such cooperation credible.

Priorities include evaluating recommendation systems for information quality and diversity, teaching AI and verification literacy, and improving cross-border standards for transparency. Success should be measured through observable outcomes rather than the volume of removed content alone. The aim is an environment in which people can assess evidence and make collective decisions on a shared factual basis, while protecting open debate and the right to challenge authority.

About the author

Ivan-Oleksandr Yakymiv is President of the Board of Stowarzyszenie Centrum Wsparcia “Just Do It”, founder and publisher of TRUTH Magazine, leader of CODI-DW (Combating Disinformation in the Digital World), and founder of CODI Institute (Combating Disinformation). CODI Institute is a project and training initiative focused on digital resilience, critical thinking, AI in disinformation, and platform architecture against online manipulation.

Ivan-Oleksandr Yakymiv

Sources and further information

1. Claire Wardle & Hossein Derakhshan (2017), Information Disorder — Council of Europe
https://edoc.coe.int/en/media/7495-information-disorder-toward-an-interdisciplinary-framework-for-research-and-policy-making.html

2. Soroush Vosoughi, Deb Roy & Sinan Aral (2018), The spread of true and false news online — Science
https://doi.org/10.1126/science.aap9559

3. Aristotle, Rhetoric, Book I — Internet Classics Archive, MIT
https://classics.mit.edu/Aristotle/rhetoric.1.i.html

4. Christopher Paul & Miriam Matthews (2016), The Russian Firehose of Falsehood Propaganda Model — RAND
https://www.rand.org/pubs/perspectives/PE198.html

5. Computational Propaganda — Oxford Internet Institute
https://www.oii.ox.ac.uk/research/projects/computational-propaganda/

6. Bobby Chesney & Danielle Citron (2019), Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security — California Law Review
https://www.californialawreview.org/print/deep-fakes-a-looming-challenge-for-privacy-democracy-and-national-security

7. Regulation (EU) 2022/2065 — Digital Services Act, Articles 25–27, 34–40
https://eur-lex.europa.eu/eli/reg/2022/2065

8. Jon Roozenbeek et al. (2022), Psychological inoculation improves resilience against misinformation on social media — Science Advances
https://pmc.ncbi.nlm.nih.gov/articles/PMC9401631/

9. Gordon Pennycook et al. (2021), Shifting attention to accuracy can reduce misinformation online — Nature
https://doi.org/10.1038/s41586-021-03344-2

10. Gordon Pennycook et al. (2024), Inoculation and accuracy prompting increase accuracy discernment in combination but not alone — Nature Human Behaviour
https://doi.org/10.1038/s41562-024-02023-2

11. Shoshana Zuboff (2019), The Age of Surveillance Capitalism — author's introduction
https://shoshanazuboff.com/book/about/

Contact

https://truthmagazine.eu/articles/ivan-oleksandr-yakymiv/ | redaktor@truthmagazine.pl

Online version: https://truthmagazine.eu/publikacje/architecture-of-deception/