

From fact-checking and deepfake-spotting to critical media literacy: Disinformation interventions

Prof. UO, dr hab. Katarzyna Molek-Kozakowska,
University of Opole
Lucian Blaga University of Sibiu

Plan

- 1. Context of disinformation (need for resilience)**
- 2. Projects in EU supporting citizens and regulators**
- 3. Projects detecting and explaining manipulation, fake news and deepfakes**
- 4. Critical media literacy resources**

Context of disinformation

Current challenges

- information overload/deficiency, epistemic crisis,
 - technostress, technological dependency,
 - corporate/algorithmic control of media, echo chambers,
 - dis/misinformation, AI-powered fake news, etc.
-
- **Freedom of speech:** regulators are unwilling to introduce direct legal solutions (bans) to shield media users from these challenges and advocate for individual/social resilience against, e.g., disinformation;
 - **Resilience** to be shaped through **interventions**, to increase the capacity of a system (person, community, organization) to cope with adversarial circumstances and to recover after a challenging event.



This site is managed by:
Directorate-General for Communication

The Digital Markets Act is the EU's law
to make the markets in the digital sector
fairer and more contestable

Digital Services Act (DSA) overview

The DSA regulates online intermediaries and platforms such as marketplaces, social networks, content-sharing platforms, app stores, and online travel and accommodation platforms. Its main goal is to prevent illegal and harmful activities online and the spread of disinformation. It ensures user safety, protects fundamental rights, and creates a fair and open online platform environment.

POLICY AND LEGISLATION | Publication 13 February 2025

The Code of Conduct on Disinformation

The Code of Conduct aims to combat disinformation risks while fully upholding the freedom of speech and enhancing transparency under the Digital Services Act (DSA).

Code of Conduct on Disinformation



- **Demonetization:** cutting financial incentives and advertising revenues to the purveyors of disinformation;
- **Transparency of political advertising:** efficient labels, tags and libraries for political advertising;
- **Ensuring the integrity of services:** identify and eliminate tactics of malicious agents (fake accounts, bot-driven amplification, impersonation, deepfakes);
- **Empowering users and researchers:** safe design and free access;
- **Empowering the fact-checking community:** platforms to fund and enable the work of fact-checkers, extensions in various languages;
- **Taskforce for transparency,** reporting and monitoring.

Resilience interventions (1/2)



- **fact-checking instructions** to foster higher sensitivity to fake news and teach the verification protocol,
- **fact-checking tools/databases** for individual self-help and verification,
- **fact/spin-checking services** with established news organizations,
- image-doctoring and deepfake **spotting tools**,
- **monitoring services** with alerts/warnings related to viral fake news,
- **debunking sites** with current analyses and reporting of misleading content and the mechanisms of its dissemination,
- **prebunking** (inoculation) i.e. games, infographics, videos, graphic novels that explain the mechanisms behind mis/disinformation.

Resilience interventions (2/2)



- **AI-powered tools** to be developed for standardized verification against multiple sources and for manipulation detection,
- **awareness-raising and instructional materials** devoted to exposing, i.a. misleading argumentation, manipulation tricks, logical fallacies, language (mis)use,
- **learning/teaching materials**: lesson scenarios, exemplary materials, videos to be used in formal and informal education,
- **critical media literacy campaigns** targeting various groups of citizens,
- **recommendations** to policymakers and regulators.

EU supporting citizens and regulators

EU DisInfo Lab



EU DISINFOLAB CALLS FOR STRONG DSA ENFORCEMENT AND WELCOMES THE EU COMMISSION'S INVESTIGATION INTO META

With only five weeks left before the EU elections, the EU Commission's investigation on Meta will have to demonstrate if the DSA can deliver results to protect the elections from foreign interference and #disinformation spread through online platforms.

[Read more](#)

PAST WEBINARS - WATCH THE RECORDINGS!



15 MAY: 2025: CLIMATE DISINFO RELOAD

From AI-generated greenwashing to transatlantic denial campaigns, climate disinformation is mutating fast. Join us as top experts from Global Witness and Ripple Research expose the biggest threats of 2025, and what we can still do about them.

[Read more](#)



8 MAY: INFLUENCE OF FOREIGN NARRATIVES ON CONTEMPORARY CONFLICTS IN FRANCE (IN FRENCH)

This session examines how foreign information manipulation influences public perception and democratic resilience in France. It explores key narratives related to recent conflicts and analyse their impact based on a new report by the Fondation Descartes.



30 APRIL: EXPLORING POLITICAL ADVERTISING DATA – WHO TARGETS ME?

In this session, Sam Jeffers will use tools and dashboards developed by Who Targets Me to explore how political advertising campaign spending, targeting, and messaging are tracked across platforms.

[Read more](#)

Responsible journalism in the era of AI
27-31 October 2025 | University of Opole

19073

disinformation cases collected and debunked

Our approach



Monitor



Identify



Document



Respond

DATABASE

Categories



European elections



Ukraine

288 articles

Show: 10 ▾

Disinformation Review



June 05, 2025

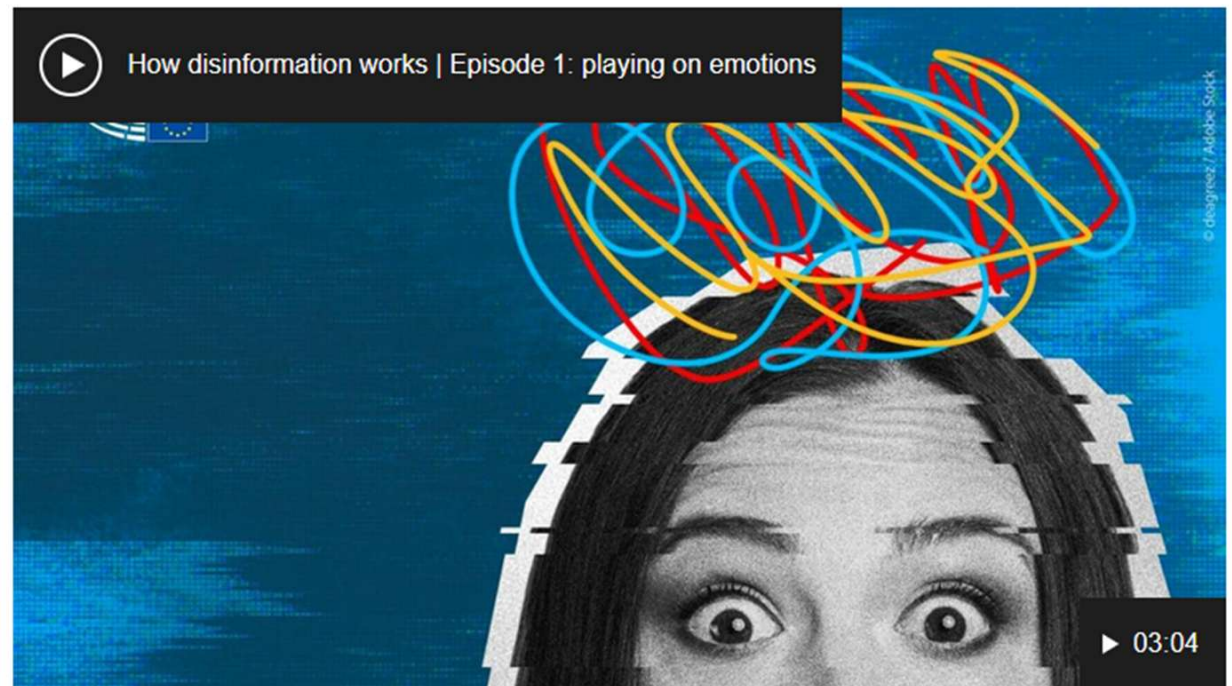
An unpleasant incident

The Kremlin downplays Ukraine's devastating attack on Russian military airfields, looking for distractions and feebly recycling old lies about terrorism and Anglo-Saxons instead.

EP: Six tactics used to fool us

Table of contents

- [Playing on emotions](#)
- [Polarisation](#)
- [Flooding the information space](#)
- [Taking advantage of the confirmation bias](#)
- [Manipulating context](#)
- [Attacking and silencing critical voices](#)



People who post disinformation use strong emotions to try to manipulate us into spreading their message



How disinformation works | 3: Flooding the information

European Parliament • 141K views • 1 year ago



How disinformation works | 2: Sowing division

European Parliament • 113K views • 1 year ago



How disinformation works | 1: Playing on emotions

European Parliament • 183K views • 1 year ago



How disinformation works | Episode 6: Silencing critical voices

European Parliament • 21K views • 2 months ago



How disinformation works | Episode 5: Presenting things out of context

European Parliament • 178K views • 3 months ago



How disinformation works | 4: Confirming existing beliefs

European Parliament • 46K views • 4 months ago



European
Parliament

- **EMOTIONAL
CONTENT**
TRIGGERS
ENGAGEMENT

- **CHECK**
INFORMATION
THAT MAKES YOU FEEL
ANGRY, AFRAID OR EXCITED

- **PAUSE AND RESIST**
THE REFLEX TO REACT
WITHOUT THINKING

and resist the reflex to react
without thinking.



02:50 / 02:55

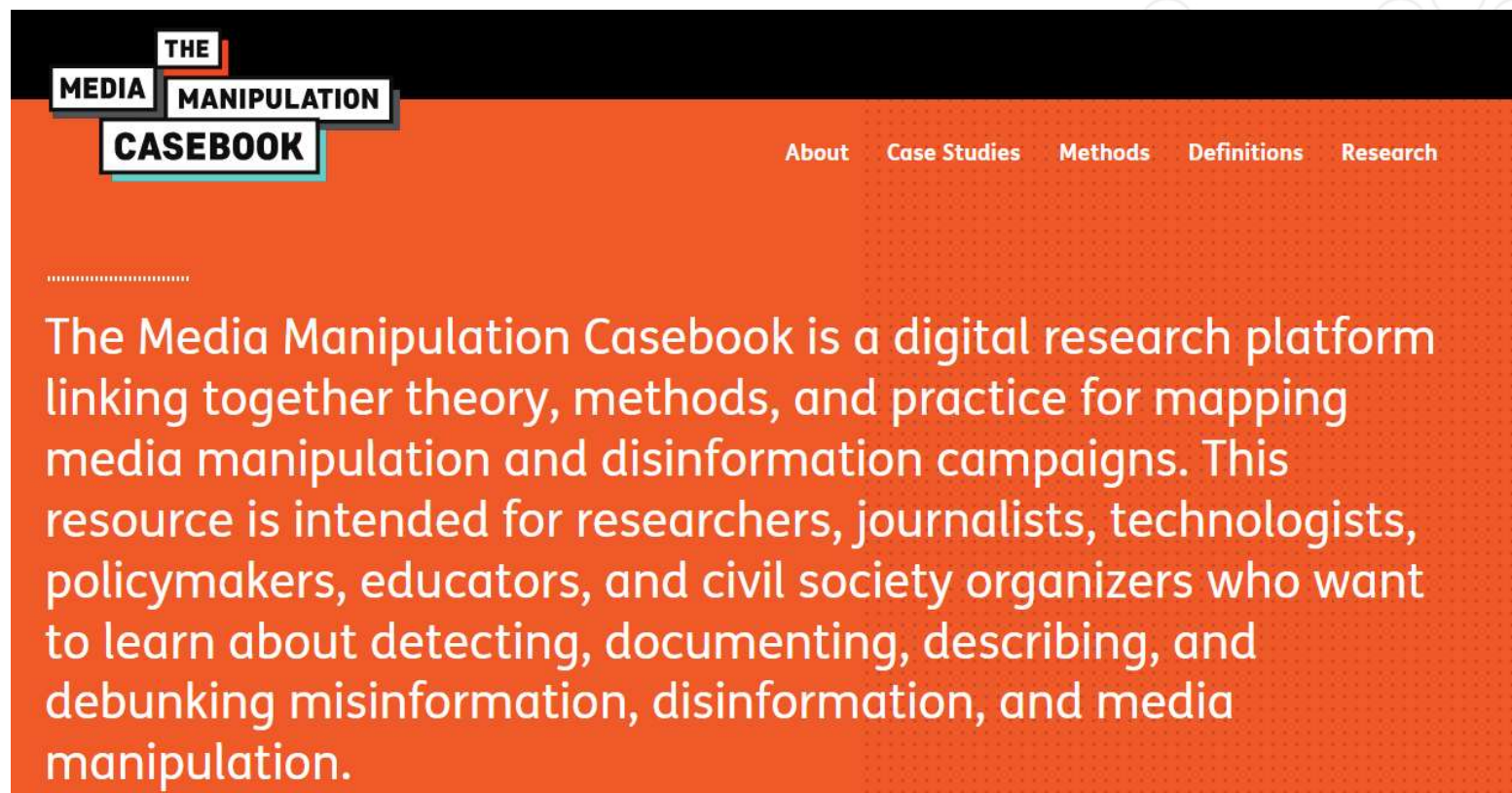


People who post disinformation use strong emotions to try to manipulate us into spreading their message

Detecting and explaining manipulation

The Media Manipulation Casebook

<https://mediamanipulation.org/>



As the practice of media manipulation evolves in response to changes in policies, regulations, and politics, so do the tactics. Below are some of the ways campaign operators take advantage of our networked media ecosystem.



Keyword Squatting



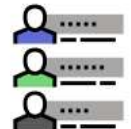
Recontextualized media



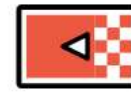
Misinfographics



Evidence collages



Viral sloganeering



Cheap fake



advertising



recontextualized media



testimonial



trolling



viral sloganeering



copypasta



evidence collages



keyword squatting



memes



swarming

4 8 A B C D E F G H I K L M N O P R S T U V W Y

DISRUPT

EN ▾

BAD NEWS

From fake news to chaos! How bad are
you? Get as many followers as you
can.

FALSIFY

START

ABOUT

SHARE

TILT

Resilience Series **RAAKE**



Commissioned by the Cybersecurity and Infrastructure Security Agency
Written by Clint Watts and Farid Haque

Resilience Series **BUG BYTES**



CLINT WATTS
FARID HAQUE

Detecting fake news and deepfakes

CODE OF STANDARDS



The **European Code of Standards for Independent Fact-Checking Organisations** is a set of criteria designed to ensure that organisations fact-checking misinformation and disinformation adhere to the highest standards of methodology, ethics and transparency in order to best serve the public interest. Once verified as compliant with the professional standards set out in this Code, organisations become members of the European Fact-Checking Standards Network (EFCSN).



Prebunking at Scale

Prebunking at Scale aims to enhance public resilience against dis- and misinformation by proactively addressing emerging false claims and narratives before they gain widespread traction by combining the strengths of two of the most widely used interventions against the spread of disinformation: prebunking and debunking. The former is often not specific enough, and the latter is only applicable retroactively. Our project will reconcile these approaches and leverage their strengths to most effectively counter disinformation.



Ask FactCheck

Have a question for FactCheck? [Ask us](#). Please also see our [Ask SciCheck archives](#) for questions answered by our SciCheck team.



SciCheck

Fact-checking science-based claims.



Facebook Initiative

Debunking viral claims.



Players Guide 2024

The special interest groups behind the TV ads.



Viral Spiral

Don't get spun by internet rumors.



Project

Detect DeepFakes: How to counteract misinformation created by AI



0:00 / 0:10



CC
Matt



In this Project:

Overview

Updates

People

Publications

How to Spot Political Deepfakes

AI literacy—and a healthy dose of human intuition—can take us pretty far.

via KelloggInsight · July 1, 2024
in Affective Computing · Opera of the Future · Viral Communications +1 more
#artificialIntelligence

Becoming Human - AI's Next Evolution Part Two: AI & We

Media Lab researcher Nikhil Singh and alums Aruna Sankaranarayanan and Matt Groh discuss ways that deepfakes may affect the 2024 elections.

via Channel News Asia · June 2, 2024
in Affective Computing · Opera of the Future · Viral Communications +1 more
#artificialIntelligence #politics #ethics

A deep dive into deepfakes

On BBC's The Forum, host Iszi Lawrence tries a tool developed by Media Lab researchers that tests users' ability to detect deepfakes.

via BBC's The Forum · May 27, 2023
in Affective Computing · Responsive Environments · Media Lab Research Theme: Life with AI
#human-computer interaction #artificialIntelligence #social media +2 more

- **Focus on the face.** Are facial features altered or inconsistent?
- **Observe the cheeks and forehead.** Does the skin appear unusually smooth or wrinkled? Are the skin's age characteristics consistent with the hair and eyes?
- **Pay attention to the eyes and eyebrows.** Do the shadows fall where you'd expect them to? Does the video accurately capture the natural physics of light?
- **Check the glasses.** Is there a glare, or perhaps too much glare? Does the glare shift appropriately as the person moves?
- **Inspect the facial hair or lack of it.** Does the beard or mustache look authentic? Does the video add or remove facial hair or make it appear natural?
- **Notice any facial moles.** Does the mole look convincing?
- **Watch the blinking.** Is the person blinking too frequently or too little?
- **Focus on the lip movements.** Do the lip movements align naturally with speech?



<https://support.google.com/websearch/answer/9116649>

☰ Google Search Help

🔍 Describe your issue

[Help Center](#)

[Community](#)

[Improve your Google Account](#)

Remove explicit non-consensual fake imagery from Google

We recognize it can be distressing when explicit non-consensual fake imagery is discoverable in Google search results. This article is intended to support you through the process to request removal of such content from Google search results.

Requirements

For us to consider the content for removal, it must meet all of these requirements:

1. You're identifiably depicted in the imagery.
2. The imagery in question is fake and falsely depicts you nude or in a sexually explicit situation.
3. The imagery was distributed without your consent.

Critical media literacy resources

Need help dealing with violent or distressing online content? [Learn more](#)

[Key topics](#) | Fake news and misinformation

< Key topics

The eSafety Guide

Cyberbullying



Adult cyber abuse



Image-based abuse



Illegal and restricted online content



Domestic violence



Technology-facilitated abuse



Workplace bullying and harassment



Online tools and features



How to stay safe online



Digital wellbeing



Sexual violence

Gendered violence

Fake news and misinformation

Fake news and misinformation

Information, images and videos posted online can be untrue or misleading, so check the facts before sharing them.

Sometimes stories, scenes and sounds are completely made up or altered so they seem real.

How can I tell if information is true?

Look out for content that seems very one-sided.

Ask yourself these questions:

- Is it from a trustworthy source, known for being honest about information?
- Does the headline match the content?
- Are different views included, to give the content balance?
- Do the quotes make sense and match the rest of the story, or do they seem to be missing the wider context?
- Does the content seem to be unbelievable or 'too good to be true'?
- Is there enough evidence and reasoning in the story to justify the claims or conclusion?





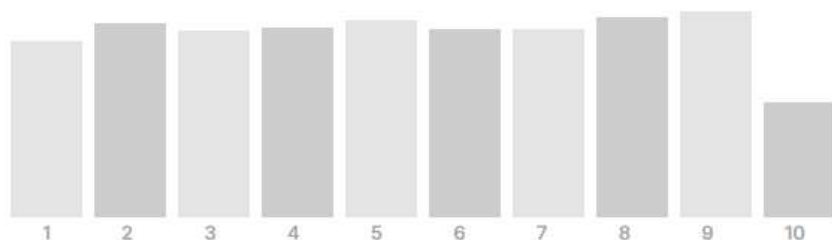
VIDEO: How to identify a catfish



VoxCheck is a Ukrainian fact-checking project. Since 2018, VoxCheck has been a signatory to the Code of Ethics of the Poynter Institute's International Fact-Checking Network (IFCN), the largest professional fact-checking community in the world. And also a member of Meta's Third-Party Fact-Checking Program, under which we rate and refute false content shared by users of the platform. You can find more information about our policies and principles on the official VoxCheck [website](#).

Structure: disinformation narratives, fakes and cases

The database consists of thousands of blatantly fake or manipulative messages, which we call disinformation cases. In order to classify them, we combined all the cases based on a common theme into fakes. From fakes, we formed narratives based on a key common message. That is, different cases of disinformation are examples of the same fake.



FAKES' DYNAMICS, 2023

2022 2023



GENERAL STATISTICS

Narrative concerning weapons provided to Ukraine by Western countries

Allies want to divide Ukraine

Black transplantology in Ukraine

Narrative concerning reforms in Ukraine

Narrative concerning foreign mercenaries



ITALY

GERMANY

POLAND

SLOVAKIA

HUNGARY

CZECHIA

White list

Demagog

Pravda

Konkret24

TVN24

Onet.pl

Wyborcza

Rzeczpospolita

Wirtualna Polska

Defence24

Polski Instytut Spraw Międzynarodowych

PAP

OKO.press

Black list

Sputnik Polska

Niezależny Dziennik Polityczny

Wolne Media

News Front

Myśl Polska

Sieć Voltaire

Bibula

Geopolitica.ru

Bumerang Polski

NEon24

Media Narodowe

salon24

Responsible journalism in the era of AI

27-31 October 2025 | University of Opole

DisInfoResist

Projekt naukowo-interwencyjny na temat budowania
odporności na dezinformację wśród młodzieży
opracowany we współpracy z



CORECON



Sources



- America's Cyber Defence Agency (<https://www.cisa.gov/>) Real Fake graphic novel
- Australian Government's e-safety commissioner website (<https://www.esafety.gov.au/>)
- Bad News Game: An interactive game that puts players in the shoes of a fake news creator to teach them about the tactics used in disinformation. (<https://www.getbadnews.com/>)
- EU Disinfo Lab (<https://www.disinfo.eu/>)
- EU vs. DisInfo: European External Action Service's East StratCom Task Force (<https://euvsdisinfo.eu/>)
- European Parliament: Spotting disinformation: Six tactics used to fool us
<https://www.europarl.europa.eu/topics/en/article/20250227STO27081/spotting-disinformation-six-tactics-used-to-fool-us>
- FactCheck.org: US-based non-partisan, non-profit organization that monitors the factual accuracy of statements by public figures and organizations. (<https://www.factcheck.org/>)
- MIT Media Lab's affective computing: DetectDeepFakes Project (<https://www.media.mit.edu/>)
- Media Manipulation Casebook (<https://mediamanipulation.org/>)
- Propaganda Diary: A reference base of fake narratives circulated in European media regarding Ukraine
<https://russiandisinfo.voxukraine.org/en>



Funded by
the European Union



forthem.

Fostering Outreach
within European Regions,
Transnational Higher Education
and Mobility



CORECON

The **C**overage and **R**eception of
the Russian-Ukrainian **C**onflict

Responsible journalism in the era of AI
27-31 October 2025 | University of Opole

Katarzyna Molek-Kozakowska



molekk@uni.opole.pl

CORECON

website



grants.ulbsibiu.ro/corecon

