

Disinformation in EU politics through AI

MA Igor Emmanuel da Silva Carnaúba

What is AI?

Where do you see AI in your real life?

Microtargeting: Data Analysis, Psychometrics, and Public Sentiment

- **Microtargeting:** This is the practice of using data (often harvested from social media or other online platforms) to analyze personal details, behaviors, and preferences of individuals or small groups. AI enables highly precise targeting of political messages or propaganda to these groups, exploiting their specific fears, biases, or interests.
- **Psychometrics:** Psychometrics is the measurement of psychological traits and values—like personality, attitudes, and opinions—using data analysis. AI can use psychometric data to predict how people will react to certain messages, then tailor content to maximize its persuasive power.
- **Public Sentiment Analysis:** AI tools can scan vast amounts of online communication to determine what people are feeling or how they might react to news or events. This helps disinformation actors adjust their strategies in real-time for greater impact.

Automated Disinformation: The Proliferation of Bot Technology

- Bots and Botnets: Bots are automated programs that can generate or share content online without human intervention. A botnet is a network of such bots, operating in concert to push large volumes of disinformation, amplify certain narratives, or fake the appearance of widespread support or outrage.
- Automated Disinformation: By using bots, disinformation campaigns can quickly spread messages, overwhelm genuine discourse, and disrupt the flow of factual information. Automation makes these campaigns faster, more persistent, and harder to trace.

AI-Facilitated Forgery: The Threat of 'Deepfake' Technology

- Deepfakes: Deepfakes are highly convincing fake images, videos, or audio recordings created using AI techniques like machine learning and neural networks. They can show people saying or doing things they never did, making it extremely difficult for ordinary individuals to discern what is real.
- Machine Learning and Neural Networks: These are types of AI that learn from data. In the context of disinformation, they can be trained to create realistic forgeries or to generate new, emotionally resonant fake content that mimics genuine speech or writing styles.
- Authenticity and Convincing AI-produced Content: AI can now produce content so realistic that it becomes hard to distinguish misinformation from real news or authentic footage, undermining public trust in media and institutions.

Is it real or fake?

<https://www.dw.com/en/real-or-fake-fact-checking-quiz-fake-news-ai-fake-ai-image-detection-debunking-verification/a-72104848>

More definitions

- **AI Hallucinations and Misinformation:** AI systems, particularly large language models and chatbots, can generate content that is not factually accurate (“hallucinations”) or can be used to deliberately mislead, manipulating human decisions and exploiting cognitive vulnerabilities.
- **Need for Ethical Frameworks and Regulation:** The paper stresses the urgent need for robust, ethical standards and updated legal/regulatory frameworks to guide AI development and deployment. Responsible AI usage involves collaboration among all stakeholders—researchers, developers, policymakers, and users.
- **Balancing Innovation and Societal Values:** While AI brings unprecedented opportunities for progress and innovation, it is needed to ensure these advancements align with ethical principles and societal values. Society needs regular training, engagement, and sharing of experiences among AI users; open dialogue; and the development of best practices to mitigate risks.

Let's trick AI

- Try to craft prompts/questions that cause AI to generate misleading information (“hallucinations”).
- Discuss why these errors occur and what this reveals about the strengths and limits of AI tools.

LET'S PLAY!