

**AI Misinformation and Disinformation Spread Through Social Media Platforms: An
Individual's Responsibility to Understand the Risks and Benefits**

Ryan MacLean

70329263

Faculty of Arts, University of British Columbia

Poli 100: Introduction to Politics

Dr. Yves Tiberghien

February 17, 2025

Introduction:

On February 4th, 2025, Stanford University hosted another addition to its weekly Cyber Policy Center Winter Seminar Series. Fortunately, I had the pleasure of attending this unique and highly educational event while listening to the expertise of Indiana University's very own director of social media, Filippo Menczer. Menczer took his vast expertise in social media algorithms and AI implementation to analytically quantize AI's positive, negative, and destructive effects on information spread and human health. The hour-long seminar highlights several distinctive instances where AI has been used to misinform individuals in serious manners, altering their once sound decisions based on science and reason into illogical actions performed out of intellectual manipulation.

I advocate based on Menczer's demonstration of AI manipulation of important information, that it's an individual's responsibility, especially if they don't handle AI regularly, to familiarize themselves with developing AI dangers online. Additionally, those individuals must understand how these dangers influence their logical political insights through the spread of misinformation and disinformation. Menczer's refreshing holistic viewpoint on the use of AI creates three central themes within his presentation, each focusing respectively on the quantified measures of "the good, bad, and the ugly" (Menczer, 2025) free implementation of AI on our daily stream of information. My report will provide a key summary of these extensive themes and will further explore one's responsibility to self-educate on AI before sufficient regulations are implemented.

Summary:

The Positive Use of AI on Information Verification:

AI tools like ChatGPT are capable of identifying false online claims with an accuracy of around 90% (Menczer, 2025). While this is impressive, Menczer explains this AI identification doesn't always agree with highly trained politically independent human fact-checking organizations. When it comes to AI confirming true online claims, it doesn't perform to the same level only correctly identifying true online claims 15% of the time. This is mainly because "it's unsure" in 65% of cases needing further info to confirm the information's validity. (Menczer, 2025). While the accuracy isn't perfect it can be a powerful tool to verify online information quickly and increase human belief along with the number of online shares; helping spread reputable news within social circles (Menczer, 2025).

The Negative side effects of AI fact-checking:

Menczer's next central theme focuses on how "We may try to use AI for a good application, but it has an unintended negative consequence" (Menczer, 2025). AI fact-checkers aren't particularly good at verifying true claims. Along with the 65% rate of uncertainty there's also a 20% likelihood AI tools like ChatGPT will label true online facts as false (Menczer, 2025). This can create serious confusion and misunderstanding in addition to diminished trust in reputable news sources. AI tools may also state information to be true despite being false (Menczer, 2025). Individuals exposed to inconclusive ChatGPT fact-check information are 10% more likely to believe and 9% more likely to share this information even though it might not be 100% true. This can create accidental widespread misinformation networks as people not only share this false news online but also in real life (Menczer, 2025).

The AI use of bad actors to create destructive circumstances:

While individuals must consider the negative externalities of AI use when fact-checking online information, Menczer also warns of the potential harm coming from the purposeful spread of disinformation through bad actors and their AI tools. As mentioned by Menczer, websites such as Children's Health Defense use disinformation to argue false standpoints on serious topics like the approval and administration of COVID-19 vaccines for children. They spew unreviewed information within their articles and then cite unsupportive academically dense materials that are extremely difficult for non-experts to digest and understand. Lazy individuals may not look to read that cited information and instead take the article's false claims at face value leading to harmful and devastating consequences like death from COVID-19 in this case (Menczer, 2025). AI is used by bad actors to increase the outreach of these harmful sites through techniques like flooding (creating thousands of posts and deleting them later) or using AI-generated faces to appear more authentic and human-like. These misuses trick social media algorithms into bypassing their anti-bot systems while reaching the for you pages of individuals who are already more likely to share articles containing disinformation (Menczer, 2025). This can result in the manipulation of millions of politically sound opinions, helping sway votes in democratic elections.

Discussion:

I really appreciate Menczer approaching AI in its totality. I feel there are many misconceptions about AI and the hypothetical and realistic dangers it presents. However, the fact that Menczer decided to highlight AI's usefulness when implemented properly just shows the potential good these tools can be used for in the future. Yes, misuse and the spread of misinformation and disinformation will always be present while AI exists, but the creation of "thoughtful regulations that embrace the powerful promise of artificial intelligence – while

mitigating the risk that it overpowers humanity” (Altman, 2023) can ensure the maximization of AI benefits while protecting people’s health and critical political views.

AI is already here and not going away anytime soon, so as we discussed in class “specialized education in one domain (Engineering or Science) [is] not enough” (Tiberghien, 2025). It needs to be every individual’s responsibility, especially those not regularly working with AI, to understand potentially how this technology can be used against them and for them. Social media’s AI-promoted “fake and outrage-inducing content can now attain a level of reach and influence that was not possible before 2009” (Haidt, 2022). This means if you refuse to educate yourself about current misuses of AI, it will influence your everyday choices and beliefs, especially now when those required AI regulations are still developing. Without taking a critically informed viewpoint, your political perspective can be shifted to the political extremes where the spread of disinformation through bad characters thrives (Menczer, 2025).

Conclusion:

Menczer’s seminar clearly shows that without education capable of tackling this subject, waiting for proper and effective AI regulations to save people’s health and critical thinking abilities is not enough. Even in the early days of AI misuse, people have succumbed to devastating disinformation, putting their health, their money, and the political safety of their country (currently America) in harm's way. To prevent more loss and more hurt we need to collectively educate ourselves no matter what area of life an individual is in because AI is the present and future. The promoting of our understanding when it comes to AI will bring fresh solutions to the complex issue of how to regulate AI, and together we can use this enlightened thinking to create a safer internet where we humans don’t have to constantly worry about whether what we’re seeing is real or just another piece of AI-generated disinformation.

References:

- Haidt, J. (2022, April 11). “Why the past 10 years of American Life Have Been Uniquely Stupid: it’s not just a phase.” *The Atlantic*. <https://fisherp.mit.edu/wp-content/uploads/2022/05/Why-the-Past-10-Years-of-American-Life-Have-Been-Uniquely-Stupid-The-Atlantic.pdf>
- Kelly, S. (2023, October 31). Sam Altman warns AI could kill us all. But he still wants the world to use it. *CNN*. <https://www.cnn.com/2023/10/31/tech/sam-altman-ai-risk-taker/index.html>
- Menczer, F. (2025, February 12). AI and social media manipulation: the good, the bad and the ugly. *YouTube*. https://www.youtube.com/watch?v=_ptfOZNY8Gk
- Tiberghien, Y. (2025, January 23). Can Humanity Survive the Age of Artificial Intelligence. *PowerPoint*. https://canvas.ubc.ca/courses/152873/pages/week-3?module_item_id=7717519