

Misinformation and social media

Objectives

- Define fake news
- Examine how fake news proliferated in the 2016 U.S. presidential election
- Determine which types of individuals are most susceptible to fake news
- Assess computational methods for detecting fake news
- Define deep fakes and explore the potential implications of this technology
- Discuss the importance of detecting fake news and how to prevent the spread of misinformation online

Assigned readings

- Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., & Lazer, D. (2019). Fake news on Twitter during the 2016 US presidential election. *Science*, 363(6425), 374-378.
- Pennycook, G., & Rand, D. G. (2018). Who falls for fake news? The roles of bullshit receptivity, overclaiming, familiarity, and analytic thinking. *Journal of personality*.
- Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explorations Newsletter*, 19(1), 22-36.
- The Agency. *The New York Times Magazine*.
- Deepfakes: A Looming Crisis for National Security, Democracy and Privacy?
- Is seeing still believing? The deepfake challenge to truth in politics
- The Word - Truthiness

Optional readings

- Full House of Mustaches
- Robert Downey Jr and Tom Holland in Back to the future - This is heavy!
- Pelosi videos manipulated to make her appear drunk are being shared on social media
- You Won't Believe What Obama Says In This Video!

Response paper prompt

Deep fakes are a form of synthetic media constructed using deep learning techniques in which an existing image or video of a person are replaced with another individual. Modern deep fakes are relatively easy to implement using open-source software, and have attracted widespread attention for their use in hoaxes, fake news, political misinformation, entertainment, and revenge porn. They are yet another example of fake or manipulated content created in an effort to misinform its audience.

Social media is a predominant platform for spreading misinformation and deep fakes. While many individuals will recognize the falsehoods presented in the content, many others readily accept this content as fact. Social media platforms have been reluctant to police content generated by its users, and only recently have begun to engage in limited forms of fact-checking. As of yet, most social media platforms do not prohibit false content but instead only label the content as fake or minimize its appearance on their network.

What responsibility do social media companies such as Facebook and Twitter have to regulate false content that is shared on their platforms? How should any such regulation be implemented (e.g. removing fake content, labeling content as fake, reducing its appearance on public feeds)? Should certain types of fake content be treated differently from others?