4/2/2025

## CYSE 201S

# Article Review #2: The rising threat of deepfake technology

### Introduction

As more news is being delivered through online websites and social media, misinformation is becoming a growing issue. Misinformation is especially prevalent in political environments, as bias dictates how people choose to present information. Normally, when seeing quotes on places such as social media, some people will search for a video to confirm the quote, since it is believed those can't be faked as easily. However, deepfakes refer to falsified content in the form of images, videos and audio with frightening realism (Bray et al., 2023, p. 1). In order to do this, Artificial Intelligence (AI) is leveraged to create the content since it can take inputs like real audio and video and create the most realistic looking deep fake (Bray et al., p. 2).

### **Social Science issues**

Dealing with deepfakes could burden the psychological load of many people, as it requires close examination and reasoning to determine if a video or image is a deep fake, since with the continued growth of AI, it will be harder to tell. It will also make people susceptible to tactics such as manipulation, which will overall worsen the trust they have associated with digital content. Therefore, it will have sociological effects as widespread media trust shrinks in addition to the psychological effects it has. For example, this could worsen the already declining trust in the government along with education. This effect could also allow people to claim real videos are actually deepfake, which will further worsen the trust of digital content. There will obviously be many ethical and legal concerns as well, since deep fakes are often used for revenge pornography and can be used in ways that could be seen as defamation.

#### **Research Questions and Procedures**

Page 3

The first research question that the authors have is if participants will be able to differentiate between deepfake images and real images. This will test the accuracy of the deepfakes in a question-answer method. The second and third research questions check if simple interventions improve deep fake detection accuracy and how the participants confidence and actual accuracy compare (Bray et al., 2023, p. 7). This will determine the relationship between confidence in the ability to detect fake content and actual ability, which could be an eye opener for both participants and researchers. The procedure involved gathering participants from the Prolific online platform, and they were then given a payment incentive if they got in the top 50 percent of all participants (Bray et al., 2023, p. 7). In addition, each person was shown 20 deepfakes and 20 real images in their test (Bray et al., 2023, p. 7).

### **Findings and conclusion**

The study found that participants had an accuracy of around 60 percent when detecting deepfake images, and they mostly used clues such as "Impossible" backgrounds (Bray et al., 2023, p. 8-9). Overall this study answers the much needed question of how prepared we are for a society where deepfakes are regularly used by malicious actors in order to spread fake news. It also evaluates the decision making of people who come across this content online, helping us estimate what type of damage deepfakes could cause. This study relates to the information about social forces discussed in the powerpoint presentations, because it has many social influences such as politics, media, and technology. This technology will be widely used in Politics because that ecosystem utilizes propaganda the most. Lastly, it serves as a wake up call to both participants and researchers about how vigilant we need to be in the future as the use of these technologies rise.

## Works cited:

Bray, S. D., Johnson, S. D., & Kleinberg, B. (2023). Testing human ability to detect 'deepfake' images of human faces. Journal of Cybersecurity, 9(1), tyad011. https://doi.org/10.1093/cybsec/tyad011