The New Frontline of Journalism: Defending Against AI-Enhanced Information Operations and the Creation of Deepfakes

Undergraduate Researchers:
Samuel Lalor (Management Information Systems), Dan Ngo (Management Information Systems)

BACKGROUND

- Journalists are increasingly under attack from both state and non-state actors (Henrichsen & Shelton, 2022).
- However, little research has investigated the way that artificial intelligence (AI) technologies are being used to harm journalists.
- Specifically, we investigate the way that state and non-state actors are using AI-enhanced technologies such as deepfakes as information operations do discredit and harm journalists. The focus is not journalism as an overall industry, but journalists as a target for harm.
- This is a newly developed project, so our first step was to understand the existing literature with the below research question.

- **Primary Research Question**
  - *What is the current state of the literature on AI-enhanced information operations that target journalists?*

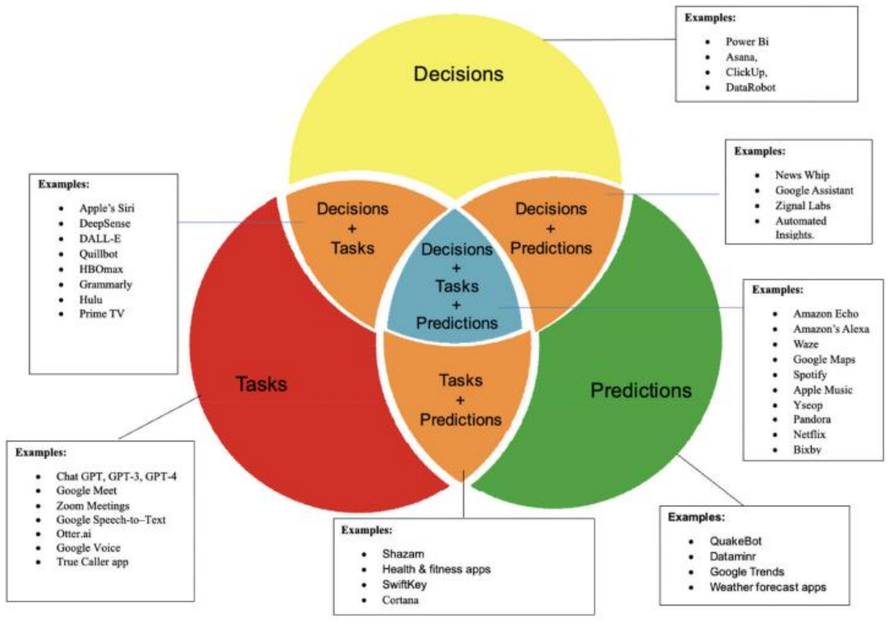

Figure 1. Venn Diagram Showing Examples of Artificial Intelligence Tools based on their levels of performance.

The Edward R. Murrow College of Communication

Carson College of Business
Washington State University

METHODS

- Since this is research in a new area we began with an annotated bibliography of the existing literature
- We first searched EBSCOhost with the following search terms:
  - *(information operations or deepfake or deep fake or shallowfake or shallow fake or AI or artificial intelligence) AND (journalism or journalist)*
- This initial search returned 490 results. We then reviewed each of the 490 abstracts to determine which papers to retain in the literature base.
- Much of the literature only mentioned journalism in passing, only 6 papers were directly relevant to the context of information operations against journalists.
- For each remaining paper the topic, method, research questions, theory, results, and an overall evaluation of the paper were recorded in an annotated bibliography.

ANNOTATED BIBLIOGRAPHY RESULTS

- How does AI benefit and harm journalists?
  - One of the benefits on the usage of AI is its ability to perform specific tasks, make decisions, communicate, interact, and forecast results, which leads to unique output (Gil de Zúñiga et al., 2024). However, there are ethical concerns surrounding AI, such as transparency, privacy, and the potential risks associated with increased machine autonomy and consciousness.
  - These are the specific ways that AI can be used in journalism:
    - AI technologies enable automated generation of news articles including it algorithms can process large datasets and produce news content swiftly (Túñez-López et al., 2021).
    - Using AI, Journalists can extract insights from vast amounts of information and Personalized content delivery becomes feasible, tailoring news to individual preferences (Túñez-López et al., 2021).
    - Shown in figure 1, each decision, tasks, and prediction application are the driving tools of AI and its quick process into making the unique output such as a news article.
  - AI can create several potential areas for future research, including the effects of computational journalism, audiences' reactions to AI-generated news versus journalist-written news, and the psychological effects of AI literacy on individuals' ability to distinguish between fake and real news (Gil de Zúñiga, et al., 2024).
- What is a deepfake?
  - Deepfakes are a multimodal disinformation that usually come in the form of video or photo where users would face-swap and change the audio of the interviewee using a software (Lee and Shin, 2022)
- How does it compare to other modalities?
  - This consisted of showing participants news that came in the form of deepfakes, text-photo (using a screen shot from the deepfake video), and text-only. The participants viewed deepfakes as the highest in source vividness than the other two (Lee and Shin, 2022).
  - This is due to the picture superiority effect, where messages presented in richer modalities are more likely to be understood than words (Lee and Shin. 2022).

FUTURE WORK

- This annotated bibliography is the first step of a larger project on AI-enhanced information operations and deepfakes in the context of journalists.
- Future research will focus on a full systematic literature review, socio-technical methods for identifying information operations against journalists, and a development of a typology of types of information operations.

REFERENCES

Lee, J., & Shin, S., (2022). Something that They Never Said: Multimodal Disinformation and Source Vividness in Understanding the Power of AI-Enabled Deepfake News. *Media Psychology*. 25:4, 531-546. DOI:10.1080/15213269.2021.2007489

Gil de Zúñiga, H., Goyanes, M., & Durotoye, T. (2024). A Scholarly Definition of Artificial Intelligence (AI): Advancing AI as a Conceptual Framework in Communication Research. Political Communication, 41(2), 317–334. https://doi.org/10.1080/10584609.2023.2290497

Túñez-López, José Miguel, César Fieiras Ceide, and Martín Vaz-Álvarez. "Impact of Artificial Intelligence on Journalism: Transformations in the Company, Products, Contents and Professional Profile." Communication & Society, vol. 34, no. 1, 2021, pp. 177-193

ACKNOWLEGEMENTS

This work is supported by funding for the VICEROY Northwest Institute for Cybersecurity Education and Research (CySER) provided by The Office of the Undersecretary of Defense for Research and Engineering, in collaboration with the Air Force Research Laboratory and Griffiss Institute.

WASHINGTON STATE UNIVERSITY