INFORMATION DARK SPACES: CONCEPTUALIZING AND CHARACTERIZING INVISIBLE VULNERABILITIES IN OUR INFORMATION ENVIRONMENT

Date: November 13, 2023

Disclaimer: This briefing note contains the encapsulation of views presented by the speaker and does not exclusively represent the views of the Canadian Association for Security and Intelligence Studies.

KEY EVENTS

On November 13, 2023, Dr. Caroline Orr Bueno presented *Information Dark Spaces* for this year's West Coast Security Conference. The presentation was followed by a question-and-answer period with questions from the audience and CASIS Vancouver executives. The key points discussed were the factors where Information Dark Spaces (IDS) emerge, the definition of IDS and its comparison to data voids and data deficits, and actors that exploit IDS.

NATURE OF DISCUSSION

Dr. Orr Bueno provided a critical look at the creation and impact of IDS, focusing on factors that contribute to misinformation and disinformation. Given the advances in the digital space, she explained the various ways IDS is being exploited through the usage of data and search engine tools. Dr. Orr Bueno affirmed the necessity of proactively identifying IDS before malicious actors do so in order to prevent vulnerabilities in the information environment. Such vulnerabilities can generate misinformation and disinformation.

BACKGROUND

Presentation

Dr. Orr Bueno provided the conceptual foundation of IDS, which stems from the idea that misinformation/disinformation and harmful content are a product of

fluid and dynamic forces in the information environment, rather than static. She explained that there are several forces that create vulnerabilities in information environments. These forces include the surge in demand for information, especially during the aftermath of a breaking news event (such as mass shooting, terrorist attack, or environment crises) where the supply of information cannot be met. The time between the absence of information and the dissemination of verified, credible, and reliable information is an example of IDS that is commonly exploited. Other factors include situational factors such as disasters and emergencies where people often seek information as a way to deal with high stress and uncertainty. Individual and algorithmic factors are unseen forces where both individual's specific keyword searches and the platform on which they entered it on, steer individuals to certain search results which could lead to harmful content, though it is seemingly unrelated to the search.

Dr. Orr Bueno explained that her work on IDS grows out of a collaboration with her colleague Rhyner Washburn at UMD and prior work in infodemiology, data deficits, and data voids, noting that the World Health Organization was involved in infodemiology, particularly during the COVID-19 pandemic where IDS prominently emerged. She defined data voids as search queries that do not return relevant results and data deficits as the supply of credible information that does not keep up with demand. Furthermore, Dr. Orr Bueno defined IDS as a combination of both data void (*Type 1*) and data deficit (*Type 2*), along with what she labels as *Type 3*, information that exists but is unknown, concealed, or hasn't been used or analyzed, and *Type 4*, information that appears to exist, but is AI generated or is synthetic data. Nonetheless, she explained that external events could cause uncertainty and generate questions which yield information searches, and thus create IDS where there are vulnerabilities in the information environment for mis/disinformation to be created.

Dr. Orr Bueno provided examples of how IDS is being exploited, identifying the key actors involved: data brokers, extremists, and China. Data brokers are able to reverse engineer search results in order to identify persons in anonymized data and this data can then be sold. She stated that this is a particular concern for members of the military, the intelligence community, and national security agencies in the case that their identities are revealed. The second actor identified were extremists who use tactics like copycat websites, document dumps, and information laundering to steer people towards specific websites and content. Copycat websites are commonly used when certain keywords are trending, and websites are created to look like credible news sources that people might trust. Document dumps refer to a release of lots of information to flood the information space to steer the narrative in one direction before news organizations are able to

analyze the information as credible. Information laundering, or the recycling of conspiracy theories, refers to using pre-existing content as breaking news events in order to resurface the content by a strategic use of keywords or search terms. The last actor identified was China, who use tactics such as SEO manipulation, URL, and keyword squatting to exploit search platforms, particularly during the pandemic. She explained that the search result for 'For Detrick' was dominated by Chinese state media results as a way to spread biolab conspiracy theories in order to blame the US for the pandemic.

Dr. Orr Bueno shared a case study of the East Palestine Ohio train derailment to illustrate IDS. She recognized all four types of IDS that contributed to the event, including the lack of reliable and credible on-the-ground reports, a rapid spread of information on social media to fill the information void (despite it being false information), industry officials concealing information from local authorities (creating distrust, uncertainty and more information searching), and AIgenerated articles and images frequently promoting fear mongering narratives.

Dr. Orr Bueno provided a list of recommendations for preventing the exploitation of vulnerabilities in the information environment, highlighting the need to conceptualize, operationalize, and categorize IDS and to consider measurements and methodologies for proactively identifying IDS. She described what data sources to use, including surveys, measures of information behaviors, quantitative measures of news coverage and keyword-based search returns, and scientific journal articles that measure censorship. Lastly, she considered how all these measurements and methodologies vary based on individual searching, such as language, geography and cultural relevance.

Question and Answer

How do we spot data poisoning and synthetic data? How do we know?

If we are skeptical of certain data, there are individual tools and platforms available for us to check to see if our data is synthetic or not. However, there is no great method of detecting synthetic data at scale. This is the reason that these information dark spaces are such a vulnerability: it is because we do not have great methods yet to mitigate these spaces from being exploited.

KEY POINTS OF DISCUSSION

- There is a void between demand and supply of information. Individual keyword searches on topics such as disasters and emergencies result in

harmful content through connections of seemingly unrelated keyword searches where vulnerabilities in information environments are exploited.

- Information dark spaces is defined as: the combination of search queries that return no relevant results (*Type 1*); inability of supply of credible information to keep up with demand (*Type 2*); information that exists but is unknown, concealed or hasn't been used or analyzed (*Type 3*); and information that appears to exist, but is AI-generated or is synthetic data (*Type 4*).
- Actors that exploit IDS include: data brokers that sell data through reverse engineering individuals through anonymized data; extremists that utilize tools such as copycat websites, document dumps and information laundering; and China that utilizes SEO manipulation, URL and keyword squatting to exploit search platforms.
- We need to be able to identify IDS through appropriate measurements and methodologies before malicious actors identify IDS in order to prevent exploitation of it.

FURTHER READING

Orr Bueno, C. (2023). Russia's role in the far-right truck convoy. *The Journal of Intelligence, Conflict, and Warfare*, *5*(3), 1–22. https://doi.org/10.21810/jicw.v5i3.5101

Orr Bueno, C. (2023a). Fractures: The impact of discord, disinformation, and damaged democracy. *The Journal of Intelligence, Conflict, and Warfare*, *5*(3), 183–186. https://doi.org/10.21810/jicw.v5i3.5194

Published by the Journal of Intelligence, Conflict, and Warfare and Simon Fraser University