ACICE
ADMM Cybersecurity and
Information Centre of Excellence

UPDATE ON THE INFORMATION DOMAIN

Issue 07/24 (July)

The Role of Pre-Bunking in Combating Disinformation and Misinformation

What is Pre-Bunking?

1.   Pre-bunking is the process of debunking falsehoods before individuals encounter them. It involves pre-emptively exposing the community to a weakened dose of false information and the techniques used to spread it, along with accurate information and critical thinking skills, to provide a degree of cognitive immunity. There are three main types of pre-bunking strategies: (a) fact-based, aims to correct a false claim or narrative; (b) logic-based, which explains manipulation techniques; and (c) source-based, which points out bad sources of information.

2.   By taking a proactive approach to combating disinformation and misinformation, pre-bunking builds more lasting resilience against falsehoods compared to debunking, which is a more reactive response that refutes false claims only after they have been circulated. Pre-bunking also enables "cross-protection" against multiple misleading narratives. For example, a study by the University of Bristol published in 2022 showed that subjects who were shown fake-expert disinformation (in this case, misleading advertising in the tobacco industry) were better able to identify similar techniques in another domain (climate denial narratives).

3.   If implemented early, pre-bunking allows governments to seize the initiative from malicious actors and prevent false narratives from taking root, thereby reducing the spread and impact of disinformation in a more sustainable manner. Pre-bunking can also complement more reactive strategies such as fact-checking, content moderation of social media platforms, or government legislation, which are less agile to keep pace with the increasing speed and scale of disinformation.

Limitations of Pre-bunking

4.   However, pre-bunking has its limitations. It requires considerable implementation time before the population builds up cognitive immunity. For example, in 2014 Finland began its pre-bunking efforts against Russian disinformation following the annexation of Crimea and is considered the most resilient country in Europe to misinformation. In contrast, Poland, the Czech Republic, and Slovakia implemented such measures much later with less success. Observing the state of disinformation in these three countries, Google launched a pre-bunking campaign in 2022 to combat the disinformation threat. In this campaign, Google created a series of short videos highlighting techniques that were seen in false claims about Ukrainian refugees. Studies have also shown that the effects of pre-bunking can wear off over time, with significant decay occurring two to three months after the initial programme implementation. In a 2020 study by the University of Cambridge, researchers tested the effects of pre-bunking by having participants watch pre-bunking videos on the tobacco industry and climate change, then tested the participants' responses on these topics at weekly intervals. They found that the benefits of pre-bunking could wear off by the second week, with a significant decline after two months. Even with an additional viewing of a video during this period, the effect was only extended for up to three months.

5.   In addition, pre-bunking of narratives may be less effective if the public's prior viewpoints related to a topic have already been fixed due to polarisation or radicalisation. Research by Google Jigsaw and American University found that pre-bunking white supremacist narratives among Americans was effective in reducing support for white supremacist messengers and their narratives among most of those tested. However, the pre-bunking videos had no effect on those who showed strong pre-existing white supremacist beliefs.

Types of Pre-bunking

Fact-based Pre-bunking

6.   In addition, Pre-bunking typically focuses on pre-empting false narratives, as seen in efforts by governments to address fake news about COVID-19. Proactively informing the population about fake COVID-19 stories allows them to identify falsehoods. It is especially useful when the disinformation being tackled requires a refutation grounded in specific facts. However, studies suggest that the effectiveness of fact-based pre-bunking is limited due to the challenge of accurately predicting false narratives that malicious actors target to spread. This could be seen during the COVID-19 pandemic, where Meta and Google sought to pre-bunk against specific falsehoods on topics such as anti-vaccination, and the effectiveness of mask protection. However, they were unable to predict the rise of misinformation related to the resurgence of the Wuhan lab leak theory in 2022.

Logic/Source-based Pre-Bunking

7.   Logic and source-based pre-bunking focuses on predicting and countering the techniques used by malicious actors. If implemented sufficiently early, it prevents false narratives from taking root. In the previously mentioned campaign by Google in Poland, the Czech Republic, and Slovakia, researchers found that people who watched the videos describing techniques and malicious actors were able to better identify misinformation and less likely to spread false claims.

Active Pre-Bunking

8.   Active pre-bunking engages an individual in the process of critical thinking and debunking misinformation. It is more interactive and immersive, potentially enhancing the learning experience. This can be carried out through online games which require players to respond to questions or prompts to learn about misinformation techniques. For example, Go Viral!, a COVID-19 counter-misinformation game developed by the University of Cambridge, teaches players to identify misinformation techniques such as the use of emotionally-manipulative language, was found to be effective in reducing their belief in false information for up to three months.

Figure 1: Screenshots from *Go Viral!* COVID-19 Counter-Misinformation Game
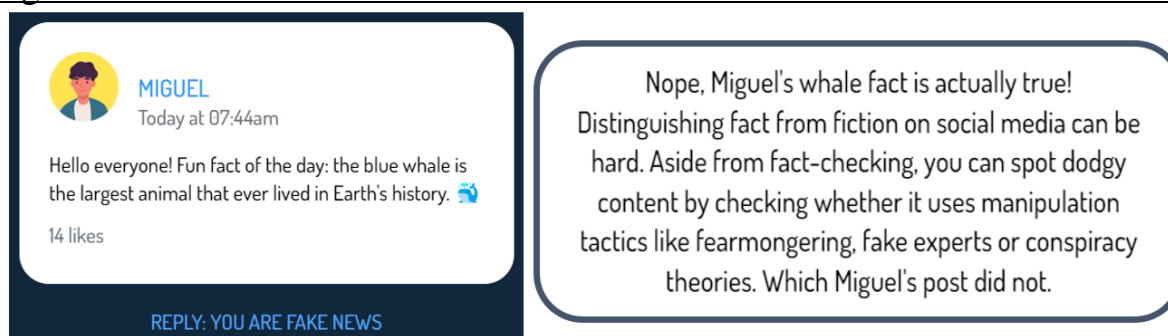
Passive Pre-bunking

9.   Passive pre-bunking involves exposing individuals to debunked or weakened forms of misinformation without requiring their active participation, and can include initiatives such as fact-checking articles, social media posts, or educational videos. For instance, prior to the circulation of a fake video in March 2022 of Ukrainian President Volodymyr Zelensky calling on his soldiers to lay down their arms and surrender in the fight against Russia, the Ukrainian government had warned its citizens about the possibility of deepfake videos as part of its passive pre-bunking efforts. Ukraine's military intelligence agency tweeted on 3 Mar 2022 about how state-sponsored deepfakes could be used to sow panic and confusion, with a link to a YouTube video elaborating on this disinformation technique. While passive pre-bunking does not provide the same level of engagement and might be less impactful, it can be employed on a larger scale through media or public awareness campaigns, thereby complementing active and more targeted pre-bunking efforts.

Best Practices in Implementing Pre-bunking

Customising to Platforms

10.   Pre-bunking formats should be customised to different social media platforms in both mobile and computer versions. For example, text and graphic formats appear better suited to Facebook and Instagram, while video formats suit YouTube and TikTok. Online games and short animated videos of about two minutes each have also been proven to be impactful in engaging audiences. In addition, pre-bunking interventions should be customised based on the target audience's language and local context, and be easily shareable on the same and across online platforms as well as closed messaging apps, particularly on mobile phones.

Collaborate with Trusted and Influential Actors

11.   As target audiences are more likely to be receptive if they trust the source of the information, one can partner social media influencers with a large following for pre-bunking programmes, particularly among the younger generations. For example, the White House carried out a pre-bunking programme in March 2022 via a virtual briefing for 30 top TikTok influencers to share the United States' strategy in dealing with Russia in the Ukraine conflict. The briefing to TikTok's most influential individuals, some with millions of followers, took into consideration TikTok's rise as a dominant source of news on the Russia-Ukraine war, and hence running the risk of being a major source of disinformation. Following the briefing, influencers began posting messages and videos to their followers, with some videos being viewed more than 300,000 times.

12.   Pre-bunking is most effective when it is implemented before an information campaign has begun, and declines in effectiveness over time. Studies suggests that a multi-pronged approach that incorporates pre-bunking along with debunking and legislative levers will be more effective than disparate, uncoordinated efforts.

Conclusion

13.   Pre-bunking is a useful tool to deal with disinformation campaigns, if it is implemented early and allowed to run for extended periods of time, with regular revisits by the population. It has the potential to inoculate the population and mitigate the negative impact when malicious actors spread false narratives. It is best implemented alongside other counter-disinformation methods, such as debunking. As part of a broader effort to counter disinformation, pre-bunking can shore up shortfalls in traditional debunking methods while also limiting the spread of disinformation through unwitting civilians. Finally, enhancing information sharing and building digital literacy are important supporting factors for the development of a successful pre-bunking campaign.

CONTACT DETAILS

All reports can be retrieved from our website at www.acice-asean.org/resource/.

For any queries and/or clarifications, please contact ACICE at ACICE@defence.gov.sg.

Prepared by:
ADMM Cybersecurity and Information Centre of Excellence

• • • •

REFERENCES
News Articles

1 Countering Online Manipulation in Europe
   [Link: https://prebunking.withgoogle.com/eu-prebunking/]

2 Meta goes public with its efforts to tackle climate change
   [Link: https://www.digitalinformationworld.com/2022/11/meta-goes-public-with-its-efforts-to.html]

3 Debunking misinformation failed. Welcome to "pre-bunking"
   [Link: https://www.washingtonpost.com/technology/2024/05/26/us-election-misinformation-prebunking/]

4 EU Media and News Survey 2022
   [Link: https://europa.eu/eurobarometer/surveys/detail/2832]

5 Play Go Viral!
   [Link: https://www.goviralgame.com/books/go-viral/]

6 The White House is briefing TikTok stars about Ukraine
   [Link: https://www.washingtonpost.com/technology/2022/03/11/tik-tok-ukraine-white-house/]