HELEN INNES, ANDREW DAWSON & MARTIN INNES

# OSINT VS DISINFORMATION: THE INFORMATION THREATS 'ARMS RACE'

## Exploring the interplay between open-source intelligence (OSINT) and disinformation to illuminate how they drive vital innovations in the organisation and conduct of each other.

Disinformation has emerged as a compelling policy problem over the past decade. Since the discovery that the St. Petersburg based Internet Research Agency attempted to interfere in the 2016 US Presidential election, multiple studies have documented various disinforming, distorting and deceptive communications shaping public understanding and political decision-making across policy domains. These include democratic elections, public health crises, climate change, counterterrorism, and warfare, amongst others. The public 'unmasking' of disinformation often relies upon a range of methods and techniques collectively labelled as 'OSINT', or open-source intelligence.

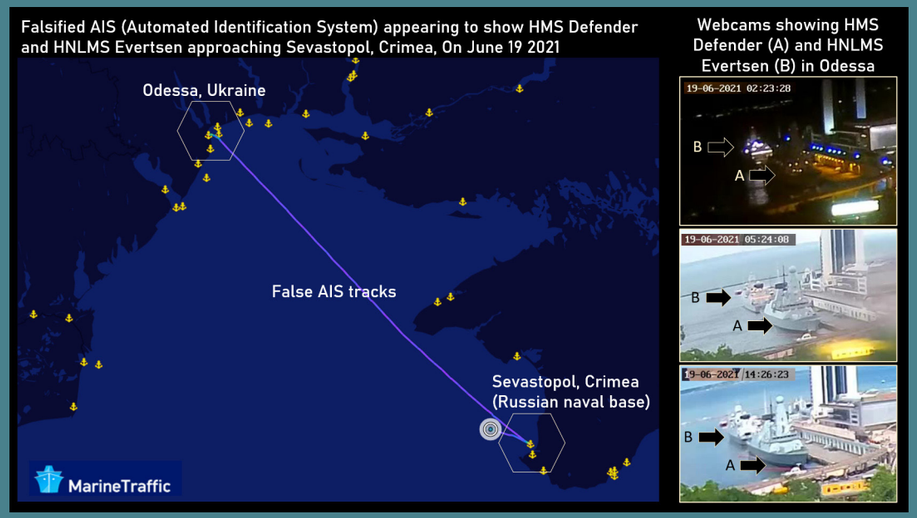
> **The public 'unmasking' of disinformation often relies upon a range of methods and techniques collectively labelled as 'OSINT', or open-source intelligence.**

Disinformation involves communications deliberately designed and delivered to mislead. It is closely aligned with several overlapping concepts, including 'misinformation' (unintentionally misleading messages), propaganda, and conspiracy theories. Concerns about the causes and consequences of misleading public communications are not new – the term misinformation was used in the 17th century in the context of the English Civil War, and George Orwell addressed its influence when writing about the Spanish Civil War in the 1930s. The key difference today is that our information environment enables misleading, yet highly persuasive, communications to be transmitted and received at a previously unimaginable pace and scale. As a result, disinformation is an important component of hostile state information operations and comparable influence campaigns by non-state actors.

Since disinformation and OSINT are both prominent features of the contemporary information environment, it is surprising that more attention has not focused on their interplay. Instead, many empirical accounts (of varying quality and sophistication) now describe various information operations and disinformation campaigns. However, these are largely separate from an increasing number of books on the craft of open-source intelligence collection and analysis.

There are intriguing co-production processes regarding how disinformation and OSINT shape innovations in each other. There is a kind of 'arms race' between OSINT analysts and the authors of disinformation. The purveyors of disinforming, distorting and deceptive communications seek to construct messages that reach and impact their targets, obscuring their origins and circumventing any attempts to intercept them. Meanwhile, the open-source analyst community seek to configure methodologies that maximise the chances of discovering misleading messages and confidently attributing sources.

There is then a continual dance of point and counterpoint as each side seeks to outwit and out-flank the other. The result is that disinformation frequently evolves and adapts, seeking new opportunities for malign influence whilst dodging the defences erected against it. The significance of this is twofold. First, although public and political discourse around disinformation centres on the role of social media, there are other vectors via which it can be transmitted and received. Second, as implied above, crises like the war in Ukraine can act as a crucible of innovation, inducing quick and substantial breakthroughs in deceptive communications. We briefly explore two examples of these dynamics.




## AIS SPOOFING

Automatic Identification System (AIS) is a radio-based system designed to alert ships to other ships in their area, preventing collisions due to poor visibility. An AIS transponder receives data from a GPS to broadcast the ship's position whilst receiving similar messages from AIS transponders on other ships in the area, allowing all the ships and their headings to be plotted on a map. Open-source marine traffic aggregators such as 'MarineTraffic.com' and 'VesselFinder.com' use AIS transponder messages to create global real-time maps of ship movements. To do this, they rely on volunteers erecting antennas on the coastline to receive AIS signals from passing ships, which are decoded by a computer and uploaded to the website. AIS is not encrypted and was not designed with security in mind. As such, AIS signals can be spoofed, resulting in incorrect or missing AIS data.

On 19 June 2021, two NATO warships were recorded on MarineTraffic.com leaving Odesa in the middle of the night and sailing to Crimea, coming within miles of the strategically vital Russian naval base of Sevastopol. This caused a flurry of social media activity as webcams from Odesa showed the two ships never left the port, meaning that someone had created false AIS tracks to trick OSINT users of MarineTraffic into believing NATO had violated Russia's security.

> **There is a kind of 'arms race' between OSINT analysts and the authors of disinformation.**

Stories in mainstream media outlets about this episode claimed Russia had spoofed AIS data; launched a GPS cyberattack; placed a nefarious AIS transmitter nearby; or interfered with the GPS. However, many of these misunderstood how AIS works and how it relates to open-source tracking websites as a form of socio-technical system. MarineTraffic.com makes it easy for volunteers to submit a data report so that anyone, anywhere in the world, can submit data. However, such reports are not verified, and just because something is displayed on the website does not mean it is happening in the physical world. Although many of the experts cited in the media discussed the technical sophistication of the systems involved, they missed the relatively easy-to-manipulate vulnerabilities to seed disinformation on them at time-sensitive moments.

The broader point, however, is that influential sources of disinformation in the contemporary information environment are not confined only to media and social media. Consequently, the OSINT community needs to widen their radar and toolkit for detecting potential vulnerabilities and exploits.

Image source: https://news.usni.org/2021/06/21/positions-of-two-nato-ships-were-falsified-near-russian-black-sea-naval-base


Image credit: (top) © President Zelensky, Handout/Anadolu Agency via Getty Images
(bottom) a screenshot from the now deleted deep fake video of President Zelensky



## ARTIFICIAL INTELLIGENCE AND DEEP FAKES

Following Russia's invasion, Ukrainian officials publicly warned that adversaries might be preparing a deepfake video of President Zelensky announcing his surrender. At the time, it was unclear if this was speculation or based on credible intelligence. However, less than two weeks later, a video was circulating on multiple social media platforms showing 'Zelensky' speaking directly to the camera. Although the manipulation was relatively unsophisticated and easy to spot, it is believed to be the first weaponised use of a deepfake during an armed conflict.

Social media platforms removed the video in violation of policies on the deceptive use of synthetic media, and Zelensky quickly debunked it. The early timing of its release, its central message, "lay down your arms and return to your families… I am going to do the same" was clearly intended to disorient and cause panic and doubt. It co-existed with disinformation coming from Russian officials that Zelensky had fled the country, contrary to Zelensky's own highly effective use of social media to broadcast 'proof of life' videos from the centre of Kyiv, the day after Russia attacked.

Deepfakes are at the cutting edge of artificial intelligence (AI) and machine learning algorithms can digitally forge a manipulated image of an individual using material sourced online. Also, in March last year, a Putin deepfake used clips from his televised Presidential address, adding new audio to make him appear to be surrendering to Ukraine. It was such poor quality and Putin's words so incongruous that audiences widely regarded it as satire, but it is certain that technological capability and expertise will rapidly advance to challenge human capacity to discern what is real and what is fake. For a lesson in how rapidly AI technologies evolve and become widely accessible and multi-purpose, look no further than ChatGPT. This large language model chatbot was only launched at the end of last year, but over 100 million users have queried it for many different purposes, some more nefarious than others.

In the hands of malign state actors, such AI-assisted technologies can create a high-volume stream of potent disinformation. An AI-assisted writing tool was recently used to generate misleading citations in a news website article about Russian opposition leader Alexei Navalny, for example. Automated text generation will facilitate the mass creation of social media accounts that look more authentic to users, whilst it appears inevitable that visual disinformation in the form of deepfakes will be deployed with other techniques of information warfare, such as hacking. The outcomes will exacerbate social tensions at critical moments of war or elections, damaging the credibility of its targets. Even a growing volume of poor-quality, more readily accessible media manipulation techniques ('shallowfakes' require only basic editing software) will erode public trust in news media.

**In the hands of malign state actors, such AI-assisted technologies can create a high-volume stream of potent disinformation.**

## CONCLUSION

The methods via which disinformation is authored and amplified are rapidly evolving and adapting. There is understandable concern that new tools and technologies will enable false and misleading messaging to be produced at a pace and scale that will overwhelm our capacity for information defence. It is also worrying that increasing numbers of actors, both state and non-state, appear to be seeing information manipulation as a key tactic and technique for achieving digital influence in the information age. In 2024 there are important elections scheduled across the UK, US, Russia and the European Union, amongst others. It is vital and urgent to consider how OSINT methods can be re-tooled and 're-armed' against future disinformation threats, to mitigate or slow this advancement.

*Helen Innes is a Research Fellow at Cardiff University Security Crime and Intelligence Innovation Institute. Her work identifying and analysing disinformation campaigns and foreign state information operations contributes to the Disinformation, Strategic Communications and Open Source Research Programme.*

*Andrew Dawson is a Research Associate at Cardiff University Security Crime and Intelligence Innovation Institute. His work spans topics such as Automated Facial Recognition, terrorism, and the Internet Research Agency. His recent work focuses on exploiting Open Source Intelligence for the Disinformation, Strategic Communications and Open Source Research Programme.*

*Martin Innes is a professor at Cardiff University and Co-Director of the Security, Crime and Intelligence Innovation Institute. His work on policing, counterterrorism, and disinformation has been internationally influential across the academic, policy and practice communities.*