Is AI Ready to Help Win Wars?

Lois Hollan and Robert Hummel

STEPS: SCIENCE, TECHNOLOGY, ENGINEERING, AND POLICY STUDIES

ISSUE 9, 2024

STEPS (Print)   ISSN 2158-3854
STEPS (Online) ISSN 2153-3679

Lois Hollan and Robert Hummel. “Is AI Ready to Help Win Wars?” *STEPS* 9 (2024): 17-25.

STEPS
SCIENCE, TECHNOLOGY, ENGINEERING, AND POLICY STUDIES
IN THIS ISSUE
ISSUE 9 2024
Robert (Bob) Hummel, PhD
Editor-in-Chief
Peak China: Personal Observations as a Western Businessperson in China
Patrick Ennis, PhD
Is AI Ready to Help Win Wars?
Lois Hollan and Robert Hummel, PhD
Applying Lessons from the Commercial Innovation System to the National Security Innovation Base
John W. Wilson
Trusted Access to Microelectronics: Addressing DoD's Unique Issues of Accessibility, Integrity, and Confidentiality of Microelectronics
Ted Glum
AI Technology to Increase US Government Transparency
Rindha Sudhini
POTOMAC INSTITUTE PRESS

POTOMAC INSTITUTE PRESS

STEPS: Science, Technology, Engineering, and Policy Studies
is published by Potomac Institute Press of the
Potomac Institute for Policy Studies.

Disclaimers: The Publisher, Institute and Editors cannot be held responsible for errors or any consequences arising from the use of information contained in this publication; the view and opinions expressed do not necessarily reflect those of the Publisher, Institute and Editors. The Potomac Institute is non-partisan and does not take part in partisan political agendas.

Access to *STEPS* is available free online at:
www.potomacinstitute.org/steps.

Is AI Ready to Help Win Wars?

Lois Hollan, Senior Fellow, Potomac Institute for Policy Studies
Robert Hummel, Chief Scientist, Potomac Institute for Policy Studies

INTRODUCTION

Artificial intelligence (AI) is expected to transform the way wars are fought and revolutionize the enterprise of national security. However, it is still unclear how this technology can be successfully leveraged for national security purposes. The problem stems from the ambiguity of the term "intelligence." Intelligence is generally taken to mean: "the ability to learn or understand or to deal with new or trying situations: reason; *also* the skilled use of reason."[1] But current AI systems are "artificial" and neither perform reasoning beyond their training nor adapt to novel situations. The value of AI to national security will be in accessing data to provide relevant, confident, and reliable information to operators, analysts, and commanders in a real and uncertain world. In this article, we examine the kinds of data that AI technology might address, the challenges of exploiting that data, an approach by which AI could enable a new dimension in the recognition of threats, and why we should develop those capabilities now.

Automation technologies are already supplanting human analysis of vast amounts of sensor data to understand "the battlespace." Techniques have been developed to perform "automated (sometimes assisted) target recognition" (ATR) to identify tanks, other military ground vehicles, aircraft, ships, submarines, and objects of significance to military operations. Exquisite sensor systems have been developed to collect data to feed into recognition systems. Such sensor data supply both human and machine recognition systems, with the latter employing both classical and emerging AI techniques to recognize threats.

Yet, these elaborate systems have failed to adapt to two new realities:

1. A massive amount of timely data is available for public consumption, which is considered unconventional and separate from the capabilities of exquisite sensor systems designed to collect (conventional) battlespace information; and
2. The kinds of items, threats, and events that must be recognized are distinctly different from the artifacts of war that have been modeled and taught to existing recognition systems.

Related to (1), there is a great deal of accessible digital data (such as social media, cell phone data with images or videos, Twitter [now X] content, news commentaries, and search engine queries). These timelier sources dominate traditional intelligence-gathering sources.

Regarding (2), it is important to recognize that the battlespace is increasingly shaped by influence operations, psychological techniques, civilian technologies, and economic and political dynamics. These novel operations elude current recognition systems; can be engaged before, during, and after kinetic conflict; and can replace kinetic warfare. Recognizing propaganda, deep fakes, nefarious ideological intent, and foreign influence has become as important as tracking troop movements or detecting tank convoys.

So maybe we've been doing it wrong or, at least, not keeping up with the times. AI may be the panacea, but likely not in the way that we have been expecting.

ACCESSIBLE DIGITAL INFORMATION

Accessible digital information comes in many forms (see Figure 1), is often unstructured, and requires interpretation. It becomes clear that the use of accessible digital information—including social media—changes the nature of military intelligence, information gathering for national security, and even the role of the "warfighter."

The explosion of available digital information has vastly multiplied the opportunities for and scope of exploitation capabilities. This is especially true for commercial and public sector applications. The US government, however, has only begun to leverage such opportunities for national security and automated exploitation purposes.

The Challenge of Exploiting Accessible Digital Information

Exploitation of open-source digital information has been used in various high-profile criminal and military cases. Often, data comes from video cameras used for surveillance by local businesses or individuals. Still images and videos are also volunteered by individuals using smart phones as cameras. In the Boston Marathon bombing of April 15, 2013, imagery from over 13,000 videos were exploited by professional and crowd sourcing analysts.[2] The massive amount of available video and other data led to a realization of the importance of volunteered, popular footage.

Since then, the government has accelerated efforts to use multimedia data to maximum advantage. The FBI has established the **Multimedia Exploitation Unit,** which employs advanced video processing technology[3] (called the MXU) to use multimedia data for identifying leads in criminal cases. The US Department of Homeland Security (DHS) has established a program called **War Crimes Hunter** to deny

US entry to persons engaged in war crimes and human rights violations. The program collects data from the Human Rights Violators and War Crimes Unit within DHS's Homeland Securities Investigations, and collects online imagery and evidence to publish facial images and other biometric data of perpetrators.[4] The New York Police Department's **Domain Awareness System** (DAS)[5] collects data from cameras and sensors throughout the city to forensically solve crimes. Its work is controversial due to implications of invasion of public privacy.[6] Recently, it has been reported that DAS will integrate the use of Ring surveillance cameras.[7] National fusion centers were established after 9/11 to receive both classified and unclassified data from governmental and open sources, and to share information with state and local government agencies.[8] There are 80 such fusion centers throughout the United States that can provide counterterrorism support to the federal government. Similar to the DAS, their use is also controversial.[9]

Figure 1. Digital Information Sources

Below is a proposed categorization of what we might consider "accessible digital information:"

Owner-disclosed Open-Source Data: Freely volunteered open-source data is any information that is posted, published, or disseminated and is available to anyone for any reason, free of charge. This type of information is typically available to anyone with an Internet connection. The value of exploiting owner-disclosed open-source data lies in its unrestricted usage. However, the veracity of the information can be suspect, and it can be difficult to align with specific applications.

Volunteered Information: Sometimes individuals voluntarily give authorities information that is not publicly available. Such "tips" are received by law enforcement as well as intelligence authorities and news outlets. Individuals with security clearances have a duty to report observations and suspicions. Examples of volunteered information include identifying insider threats or adversarial spies.

Accessible Open-source Data: Information that can be purchased includes newspaper publications and materials available through paid subscriptions or newsstand purchases, and online content behind paywalls. The purchaser is the intended recipient of the content. The intelligence community uses the term "publicly available information" (PAI) to include anything that is available to the public but may be copyrighted, require payment for access, and be subject to end-use agreements. Government use of such information is subject to restrictions.[10] Commercial satellite data fall into this category.

Profiling Information: Online resources use account information or "cookies" to track individuals' activities within and across computer applications, thereby collecting information about them. By clustering information across various dimensions, individuals can be profiled according to their attributes. This information is exchanged and sold, especially to advertisers, political campaigns, and brokers who use it for profit and gain.

National Technical Means Sources: Systems procured by government agencies for government collection of information, for example through the use of satellites, are continually upgraded and improved to provide classified information about activities on Earth.

Purloined Information: Government intelligence services engage in the business of pilfering secrets from foreign entities. When the information is not intended to be shared but has been obtained through nefarious means—which can include illegal hacking or espionage—then the information has been purloined.

In 2005, based on recommendations of the 9/11 Commission and the Robb-Silverman Commission to counter weapons of mass destruction,[11] the US established a branch of the Office of the Director of National Intelligence (ODNI) called the Open-Source Center for exploiting information of overseas activities. The Center succeeded the Foreign Broadcast Information Service (FBIS), which had focused on intercepted foreign language messages and publications. Congress had long recommended that the intelligence community (IC) make greater use of open sources, but codified these recommendations in the "Intelligence Reform Act" of 2004.[12] Today, the renamed Open Source Enterprise (OSE) is part of the CIA's Directorate of Digital Information (DDI). However, there are continuing concerns that open-source intelligence is underutilized.[13]

The OSE gleans open-source data from newspapers, internet postings, publications, and other sources, which are collectively labeled open-source intelligence (OSINT). When combined with classified sources (e.g., SIGINT or IMINT) it is "all-source intelligence," which can be exploited by other elements of the IC. Like all intelligence activities, the output may be useful for military operations, but is generally aimed at national decision-making activities.

An example of OSINT is the geolocation of adversarial activity that can be acquired from posted imagery such as selfies and terrorist recruitment videos. DARPA and IARPA co-sponsored the development of a software system using a semi-automated process to geolocate imagery for which the metadata have been stripped (as is customary for posted imagery).[14] The techniques have been adopted by news organizations and private companies to assist analyses such as forensic analysis of war crimes in Ukraine.

Within the US IC, the Defense Intelligence Agency (DIA) leads the National Media Exploitation Center (NMEC), which recently has been refocused on analysis of Chinese military actions. The DIA practices all-source intelligence analysis to understand installation and movements of foreign military assets and their capabilities, including exploitation of OSINT and social media.

The Dutch firm Bellingcat is famous for using open-source information for its forensic investigation of Russian involvement in the downing of flight MH17 in July 2014.[15] Bellingcat has continued to leverage open-source information in ongoing investigations of atrocities in Ukraine. Because their independent findings are not classified, intelligence agencies can openly discuss their work.

The Ukrainian company Molfar performs open-source investigations, publishing findings in English, in support of Ukraine's defense against Russia.[16] For example, Molfar identified a missile factory near Moscow as a source of weapons being used against Ukraine.[17]

The Challenge of Too Much Data

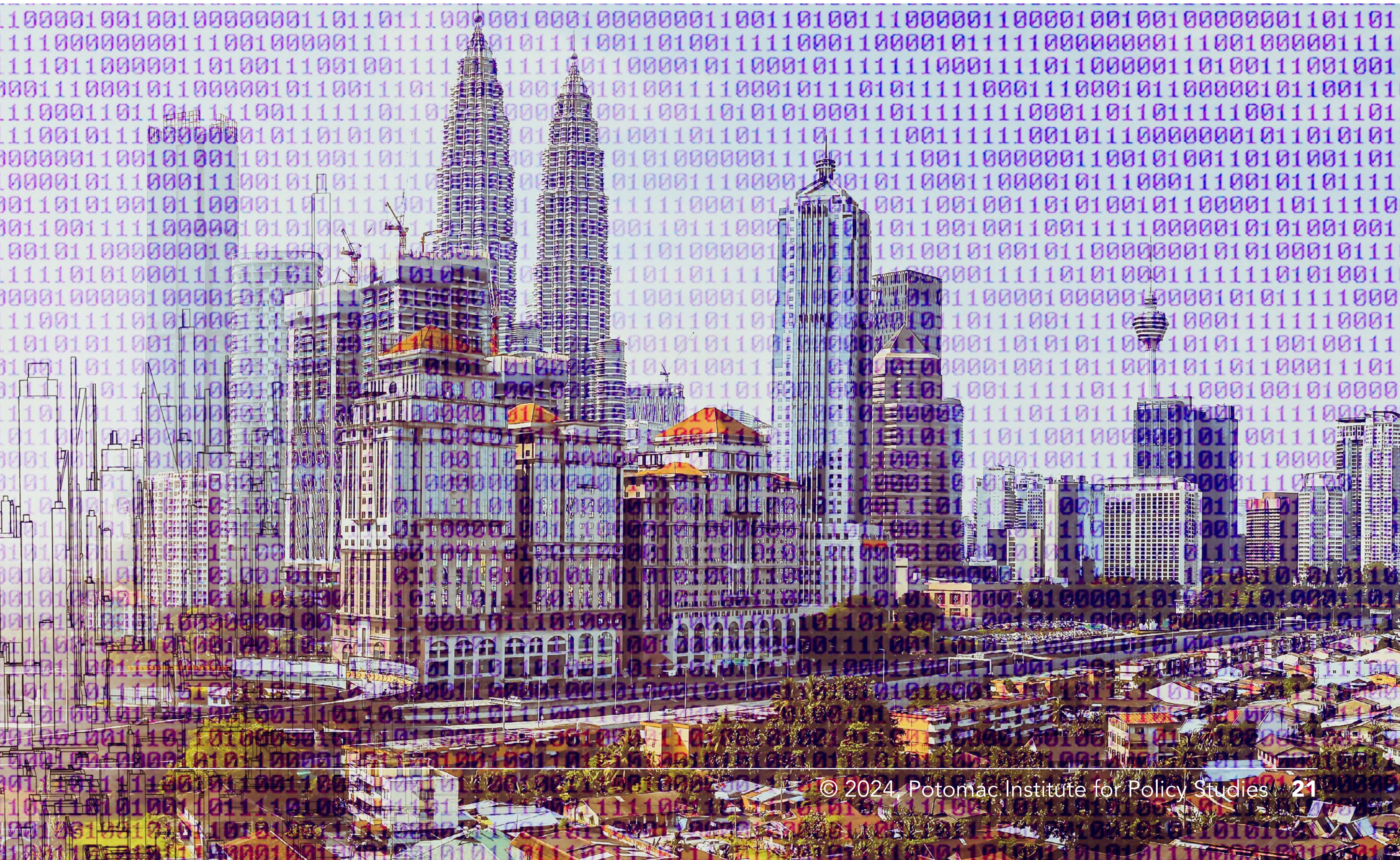
Such examples demonstrate the power of exploiting accessible digital information. At present, however, there is relatively little automation beyond the formatting and dissemination phases of data processing. Much of the analysis is performed by human analysts who are inundated by the sheer volume of available data. Analysts must comb, interpret, correlate, and productize data from multiple sources, often operating within a compressed operational timeframe.

These techniques are labor intensive and require specialized analysts trained in image processing methods, text filtering, and object recognition software. Still, it is human analysis that generates useful intelligence derived from multiple sources.

Decades ago, researchers bemoaned the "pixels to pupils" ratio, wherein the number of pixels that had to be analyzed far exceeded the capacity of the number of human pupils available to attend to those images. Thus, many images and pixels were left unobserved. Today, the situation is far worse. In addition to imagery deliberately collected by specialized sensor systems, all media in accessible digital data—combined with commercial and national collection systems—confront yet fewer analysts. Thus, the challenge is to choose which data to view and analyze.

Moreover, there is only incipient use of novel data types. Despite concerns over US civil liberties and individual privacy, new data sources can provide greater security by affording defensive and intelligence-gathering measures without impinging civil rights. The fact that adversaries are using these sources and technologies against the United States only emphasizes the urgency to recognize and defend against nontraditional combat operations using all available sources of information. Simply, valuable data cannot remain unobserved and unused.

Many hope that we can supplement the number of analysts by making use of AI to create virtual analysts. But AI is not truly "intelligent" in ways that human intelligence reasons about threats. If AI is to be used, it will not be to reason about data, but rather to assist human analysts extract relevant data from large volumes of incoming data.



The Challenge of Over-specification

Current approaches to analyzing data (whether government sensor data or other accessible digital data) largely focus on finding specific targets that are known and/or well characterized. Targets might be military vehicles, missiles, radar sets, or other well-defined objects that present specific signatures. In more complex situations, recognition of events or intentions relies on detecting specific indicators in sufficient numbers; but those indicators, in turn, rely on recognition of well-specified objects or activities. Automating the process of recognition (e.g., automatic target recognition) accelerates the search for indicators.

Regardless of how precisely an object is characterized, it remains that increasing amounts of data can lead to false alarms. False alarms must be recognized and negated by human analysis, especially given that false positives can lead to adverse consequences. The propensity for excessive false alarms renders automated recognition systems worthless.

Further, recognition techniques based on detailed modeling fail to account for new types of targets. Rarely do techniques use context and higher-level reasoning that are implicit to human thinking. Machine-learning approaches attempt to overcome this impediment but can lead to overtraining and a narrowed understanding of targets. Such systems often fail in real-world, evolving, and unknown situations. Thus, a different approach is needed to enable exploitation of massive amounts of available data.

A NEW DIMENSION IN AUTOMATED RECOGNITION OF THREATS

A viable solution involves discerning between mundane data, normal data, and data that need careful attention. This requires a more abstract view of the world. The questions is not "What kind of tank is this?" but rather, "Is this a normal event or scene?" If we can focus analysts' attention on locations and events that require attention, we can liberate the time and effort required to check on normal situations.

The central construct is to perform automated screening of data to filter out normality and to detect anomalous situations that require further analysis. Instead of trying to detect precisely modeled objects, automation should present human analysts with small and highly relevant portions of data that can assist in their assessment and understanding of the situation. By discarding the mundane, we vastly increase the breadth of data that is effectively processed. The data that should be discerned for normal versus abnormal situations entails the joint use of imagery, text, audio, and all accessible digital information.

The technical challenge is to define normality for the system to properly filter the data. Normality is a statistical phenomenon, and in multimedia environments of different data sources, it is defined by complex and highly interrelated multivariate distributions. Recent advances in AI have demonstrated an ability to parameterize complex multimodal distributions. At issue is whether such models can sufficiently characterize normality to automate sifting and analysis of accessible digital information.

Large Language Models

The technology of large language models (LLMs) represents a breakthrough in AI, which has demonstrated that generative techniques can create realistic text and images. Evidence shows that the statistics of normal text and images can be encoded in a "model" with a (mere) few billion parameters[18] within the framework of a graphical network. The statistics can be modeled so accurately that generative methods are able to produce text and images that appear normal (as opposed to nonsensical noise).

Since normality can be effectively modeled, it should be possible to detect what is "not normal." Statistical parameters of normality might need to be dependent on location, or categorization of location type. For example, these systems could model normal activity in an urban environment or normal tweets in the Middle East. Developing a model of normality, from all kinds of accessible digital data will likely require careful curation of data, so as not to pollute the model with unusual occurrences (that perhaps should draw attention). The development of a model that parameterizes "normality" should be based on multiple data sources so that dependencies and correlations can be modeled across multiple dimensions of features.

Moreover, it may be necessary to train systems to recognize the kinds of "not normal" circumstances that are of interest. Because these are (presumably) rare events, it will be useful to simulate patterns that should be flagged by a recognition system. Of course, simulations will use generative models trained in an adversarial fashion, which then may be used to bootstrap a recognition system capable of detecting targeted anomalous activities. Such simulations would involve multiple modalities to mimic an abnormal situation that warrants attention.

Therefore, it is not the precise "form of a tank" in an urban setting that is a cause for concern, but rather the movements of a set of tanks through a downtown area where tanks are not normally present. A screening tool should detect such unusual circumstances by combining images, texts, "tweets," search engine queries, and metadata about the locale and environment at large. Available digital information will presage concerns by locals that can be indicative of early stages of conflict or disasters. The mix of different information sources provides confirmation of abnormal conditions.

A RACE FOR INTELLIGENCE

The technology of LLMs has rapidly developed over the past decade, yet to date has been limited in application to generative models. Those models have now become commercially available, if not fully monetized. This is an opportune time to explore the use of technologies of complex models to screen for abnormality in available digital data to be employed for national security purposes.

We propose a program that would develop techniques to screen all forms of digital information for anomalous patterns that might be of interest to analysts. The system would sift massive amounts of available data, in real time, to find unusual patterns that can provide warning of military plans or activity. This information would be filtered by geographic regions of interest and used to alert teams of expert analysts about significant findings.

The process of building models for national security purposes will be labor and cost expensive. The number of "tokens" that must be extracted as "features" in the data will be large when compared to today's LLMs. Processing training data will necessitate considerable computer resources. Curation of training data will need to ensure that the corpus of data to be searched is relevant to each domain chosen for modeling. The generation of target scenarios will require complex scripts and production.

If successful, such an alerting system, built on large language modeling technology, would provide a powerful cutting-edge capability for national security by providing early warning and attributional evidence for adversarial activity. The first to acquire this capability will have a major global intelligence and defense advantage, which will enable countering disruption and/or aggression before it becomes critical. **The LLM breakthrough that gave rise to surprisingly good generative models would now be leveraged for important defense capabilities.**

The technology and computational power exist to build a system that ingests streams of accessible digital data, correlates these data with normality as modeled by the system, examines anomalous patterns to recognize the kinds of non-normal situations that should be flagged, and rapidly brings relevant data to the attention of analysts who can easily corroborate or deny the concern.

While this challenge is not easy, the technological advances in AI and LLMs point to viable solutions. The United States currently has an advantage over other nations' development and experience with AI information technologies. But there is no guarantee that the development of such a screening system will happen first in the US. The race to develop systems that leverage new sources of accessible digital data and screen for relevant defense and intelligence information has already begun.

SUMMARY

A principal hope for enterprises in AI is to develop capability to manage and discover intelligence from massive amounts of available data. The intent is that AI systems will supplant much of the human labor currently needed to cull data and replace current methods that can only access a fraction of the available data. With the ever-expanding availability of data (particularly open-source data), the need for such AI tools is rapidly increasing.

In the past, AI techniques have been used to assist in tracking objects; identifying vehicle types; correlating "tweets" and other online postings with events and geolocation; and alerting of changes in scenes. These techniques have very specific applications and provide utility, but fall short of accommodating the deluge of multiple dependent data sources or novel data types. Furthermore, these techniques neither address nor consider the changing nature of threats.

It should be possible to train an AI system to recognize anomalies of military significance. LLMs have surprised the technology world with their ability use billions of parameters in deep networks to model complex statistical patterns of language and imagery, when provided sufficient training examples. The generative aspects of existing models (e.g., in ChatGPT®) demonstrate the ability to model normality and might be useful for generating examples of anomalous activities that need to be recognized in text and imagery.

Training and development processes require access to massive amounts of prior data—groupings of data that have been labeled according to whether the instance is "normal" or "relevant," wherein "relevant" might be "of military significance," or might (in other applications) be categorized by other criteria. The key is that recognition of events must be broad-based, as opposed to specifically focused on a set of target vehicles, patterns, images, and/or words.

The proposed research program would require full participation of and collaboration with the US government to access requisite training data and effectively guide the development process. It will be crucial to train systems with curated data sets that intentionally either include or do not include unusual military activity.

Importantly, the system is not requesting that an AI system do any reasoning or apply actual intelligence to the analysis of situations. Instead, the program would apply techniques that have demonstrated value, namely, the ability to model statistical patterns that result in "nothing significant to report." It is the parameterization of statistics that can differentiate between "normal" and "not normal" activities that leverage breakthroughs in AI for the benefit of national security.

ACKNOWLEDGEMENTS

The authors are grateful for discussions with the Honorable H. Lee Buchanan, whose ideas on the use of social media and LLMs directly led to this paper.

ENDNOTES

[1] "intelligence," in Merrriam Webster, 2024

[2] Johnny Nhan, Laura Huey, Ryan Broll, "Digilantism: An Analysis of Crowdsourcing and the Boston Marathon Bombings," published Dec, 2015, in *The British Journal of Criminology*, Volume 57, Issue 2, 1 March 2017, Pages 341–361, https://doi.org/10.1093/bjc/azv118.

[3] Thomas Brewster, "This Secret $35 Million FBI Unit Mixes Facial Recognition With Big Data To Investigate America's Most Horrific Crimes," Forbes Magazine, July 2020, This Secret $35 Million FBI Unit Mixes Facial Recognition With Big Data To Investigate America's Most Horrific Crimes (forbes.com)

[4] Alyssa Erichs, *Privacy Impact Assessment for the War Crimes Hunter*, Department of Homeland Security, 28 May 2020, https://www.dhs.gov/sites/default/files/2022-03/privacy-pia-ice056-warcrimeshunterappendixcupdate-march2022_0.pdf

[5] E.S. Levine, et al., "The New York City Police Department's Domain Awareness System", *Informational Journal on Applied Analytics*, Volume 47(1), 18 January 2017, Pages 70-84, https://pubsonline.informs.org/doi/10.1287/inte.2016.0860

[6] Ayyan Zubiar, *Domain Awareness System*, Surveillance Technology Oversight Project, 26 September 2019, Domain Awareness System – S.T.O.P. - The Surveillance Technology Oversight Project (stopspying.org)

[7] Daniel Schwarz, et al., *The NYPD is Teaming Up with Amazon Ring. New Yorkers Should be Worried*, NYCLU, 11 January 2023 https://www.nyclu.org/en/news/nypd-teaming-amazon-ring-new-yorkers-should-be-worried

[8] Department of Homeland Security, *National Network of Fusion Centers Fact Sheet*, 09 January 2023, https://www.dhs.gov/national-network-fusion-centers-fact-sheet

[9] Michael German et al., *Ending Fusion Center Abuses*, Brennan Center for Justice, 15 December 2022, https://www.brennancenter.org/our-work/policy-solutions/ending-fusion-center-abuses

[10] The Civil Liberties and Privacy Office, *Civil Liberties and Privacy Guidance for Intelligence Community Professionals*, Office of the Director of National Intelligence, July 2011, https://www.dni.gov/files/documents/CLPO/CLPO%20Publication_Publicly%20Available%20Information_July%202011%20-%20Public%20Release%20Version.pdf

[11] National Commission on Terrorist Attacks upon the United States, Thomas H. Kean and Lee Hamilton, "The 9/11 Commission Report," 2004; Commission on the Intelligence Capabilities of the United States Regarding Weapons of Mass Destruction". United States Department of State. Washington, D.C.: Federal Government of the United States. February 6, 2004

[12] *The Intelligence Reform and Terrorism Prevention Act of 2004*, Civil Liberties Privacy Office, Office of the Director of National Intelligence, December 2004, https://www.dni.gov/index.php/ic-legal-reference-book/intelligence-reform-and-terrorism-prevention-act-of-2004

[13] Chris Rasmussen, *Avoiding the Secrecy Trap in Open Source Intelligence*, The Cipher Brief, 21 March 2023, Avoiding the Secrecy Trap in Open Source Intelligence (thecipherbrief.com)

[14] *Intelligence Advanced Research Projects Activity, Finder*, Office of the Director of National Intelligence, IARPA - Finder

[15] *MH17: The Open Source Evidence, A Bellingcat Investigation*, https://www.bellingcat.com/app/uploads/2015/10/MH17-The-Open-Source-Evidence-EN.pdf

[16] *Molfar*, Molfar Limited, https://molfar.com/en

[17] *Kh101, Kh555, Kh69: Where does Russia make its missiles?*, Molfar Limited, https://molfar.com/en/blog/fabryka-raketnogo-teroru-de-i-yak-rosiyany-buduyut-rakety,-yakymy-obstrilyuyut-ukrainu

[18] However, GPT-4 is said to have more than a trillion parameters; see https://the-decoder.com/gpt-4-has-a-trillion-parameters/