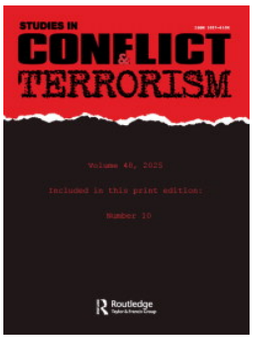

Studies in Conflict & Terrorism

ISSN: 1057-610X (Print) 1521-0731 (Online) Journal homepage: www.tandfonline.com/journals/uter20

The Ghost in the Machine: Counterterrorism in the Age of Artificial Intelligence

Christopher Wall

To cite this article: Christopher Wall (06 Mar 2025): The Ghost in the Machine: Counterterrorism in the Age of Artificial Intelligence, Studies in Conflict & Terrorism, DOI: 10.1080/1057610X.2025.2475850

To link to this article: https://doi.org/10.1080/1057610X.2025.2475850

STUDIES IN CONFLICT & TERRORISM
https://doi.org/10.1080/1057610X.2025.2475850

Routledge
Taylor & Francis Group

The Ghost in the Machine: Counterterrorism in the Age of Artificial Intelligence

Christopher Wall

War Studies Department, King's College London, London, UK

ABSTRACT

In the aftermath of 9/11 security agencies augmented their counterterrorism (CT) apparatuses with advanced analytics, machine learning (ML), and artificial intelligence (AI) to improve their ability to identify and neutralize terrorists. Under this regime, humans remained the central actors, tasked with understanding information and crafting a response. The advent of Generative AI (GenAI) changes this equation. GenAI's ability to mimic humanity's reasoning skills augurs a world where machines assume responsibility for most CT activities. This possibility raises fears of machines outside of human control. These fears are currently unfounded, and to the extent that they're real, they must be weighed against the ability to reduce the victims of terrorism. As this world forms, what will matter more is decision-makers' understanding of AI/ML outputs for counterterrorism, as they will have to make strategic choices around a series of ethical and policy choices that are inherently human. This article explores this subject more in-depth, reviewing the evolution of AI/ML and its impact on different CT domains, exploring the strategic dimensions of AI/ML, and concluding with a series of policy recommendations.

ARTICLE HISTORY
Received 7 November 2024
Accepted 2 March 2025

The New Age of Counterterrorism

The advent and democratization of generative artificial intelligence (GenAI) augurs a fundamental change to the creation of information. The most prominent form of GenAI, large language models (LLMs), excel at the imitation game first described by Alan Turing.[1] LLMs can spoof human conversations, pass American legal exams,[2] pen Shakespearean sonnets,[3] write functional computer code,[4] and summarize text. Other forms of GenAI go a step further, creating original images and videos that seem crafted by humans.[5] GenAI's proliferation has led to myriad articles, essays, and opinion pieces about what this means for humanity, as these tools show reasoning capabilities in subjects once thought solely the domains of humans. The implicit thesis embedded in these discussions is that humanity is losing its central role in the creation and understanding of information. Taking this premise at face value, then all aspects of human life will be affected by this, and counterterrorism (CT) is no exception.

CONTACT Christopher Wall christopher.s.wall@kcl.ac.uk War Studies Department, King's College London, London, UK.

In the aftermath of 9/11, governments turned to advanced analytics, machine learning (ML), and artificial intelligence (AI) to support their CT efforts to complement a global CT effort to beat back al-Qaeda and other forms of Islamic extremism.[6] Security agencies, tasked with disrupting the next major attack, pooled disparate intelligence streams to detect and counter terrorist activity across the world. Governments with access to data and processing power developed statistical tools to filter through the noise produced by human chatter, signals data, and other intelligence sources to find the proverbial needle-in-the-haystack.[7] In parallel, security agencies invested in unmanned platforms for reconnaissance and direct action, which they combined with insights derived from AI/ML to plan strikes against terrorists.[8] The advent of neural networks in the 2010s turbocharged these tools, as AI could extract deeper insights from data and plug directly into manned and unmanned systems.[9] These developments led to economies of scale for counterterrorism that greatly curtailed extremist opportunity to plot and execute mass casualty attacks. In spite these technological innovations, this counterterrorism regime remained an all too human endeavor. Governments could not simply automate CT; humans needed to mediate machine learning parameters, validate designs, translate AI/ML outputs for decision-makers, and quality control outputs against false positives or negatives that could ensnare innocents or overlook credible threats.[10]

GenAI, and LLMs in particular, represent a cognitive leap towards a world where CT is automated. The Economist describes LLMs as reasoning engines that facilitate the acquisition of information.[11] On their own, LLMs cannot reason away the world's problem. When combined with others AI/ML tools however, they offer an evolutionary path where machines consume intelligence, generate insights based on this information, and respond using an assortment of counterterrorism instruments at their disposal.

The emerging literature on this subject is quite pessimistic, much of it focused on the potential for misuse by governmental bodies and extremist groups.[12] Aside from the risks of misuse, this literature warns of unrestrained AI turning against humanity,[13] a common motif in science fiction movies like the Terminator or the Matrix. This scholarship is important in terms of delineating the potential harms with AI/ML. Yet, in some ways, the terrorism studies literature is over-indexed on these theoretical harms at the expense of discussing the value AI/ML offers for combating terrorism. Moreover, this literature is overtly focused on the tactical application of AI/ML tools without placing them in the broader context of why governments engage in CT. The platonic ideal of counterterrorism is disrupting plots to reduce harm to civilians, and AI/ML offers a suite of tools that policymakers can use to serve the public in this regard. This matters.

On most technical and tactical tasks, machines can outwork the most capable human, making certain tasks redundant, a trend set to accelerate owing to cost reductions in production, deployment, and compute. There have been initial efforts to interrogate the technical and tactical capabilities these tools offer the CT enterprise, such as using autonomous vehicles for surveillance or algorithms for predicting extremist behaviors.[14] On their own, these tools are simply that: tools. Their value derives from their use in concert with each other to aggregate the tactical and operational effects to achieve some policy goal, which is the realm of strategy. This is where the conversation is bare. Unlike tactics and operations that involve algorithmic superiority and

optimization, strategy requires ethical decisions that artificial intelligence cannot adjudicate.[15] This merits further scrutiny, as it will shape counterterrorism strategies more so than the mere existence of AI/ML tools.

Successful CT in this new age of reason will depend on how well decision-makers understand how these tools work and how to optimize their use for preventing and countering terrorism without courting the risks they entail. There is much good here, but these tools need to be studied judiciously. Like all security endeavors, unless aligned with a proper strategy, these tools will not fulfil their promise and may instead create additional harms. This article aims to be among the first to discuss artificial intelligence and counterterrorism at the strategic level. It will give an overview of AI and LLMs and how they generate knowledge, as well their limits when it comes to strategy. It then gives an overview of CT during the Global War on Terror (GWOT) to present a counterfactual on modern AI/ML could have prevented some of the policy failures from the era. Next, it gives an overview of how LLMs can work with other AI/ML tools to shape the CT mission before describing how AI/ML reaches it theoretical limits when it comes to shaping CT strategy. This paper finally concludes with a list of policy recommendations for decision-makers and democratic governments to consider.

The Mind-Body Problem in the Modern Age of Counterterrorism

AI/ML tools are programs that learn from data to make predictions and fulfill tasks that that normally require human intelligence. Computers learn through machine learning algorithms that extract knowledge from data using methods like supervised (data with labels) or unsupervised (data without labels) learning.[16] This information then allows them to make predictions based on the probabilities learned from their training data. The mishmash of conditional probabilities, built upon human knowledge, is what gives AI/ML the veneer of intelligence.[17]

The strength of an AI/ML tool does not derive from an algorithm, with many of these being open source and readily available to anyone online.[18] Rather, their abilities derive from the volume and quality of training data alongside the actual design of tool. This scoping allows humans to train AI/ML tools to discern the human world based on the data fed to them and the design principles driving them. For instance, biometric tools can identify certain correlates associated with disease in humans to improve diagnostic capability for doctors while reducing the cognitive burden of needing to do the detection.[19] Machines are therefore limited or strengthened by what information they learn, with all the biases and nuances embedded within them.

Large language models follow the same principles. LLMs are trained on troves of data derived from the internet and are optimized to receive information and respond in a manner that mimics human speech. They can translate text into any number of languages, help consumers shop, summarize content from spreadsheets and PDFs, and help solve logic problems from mathematics, medicine, law, among other fields.[20] At their core, large language models are AI systems that mathematize language and sequence words based on probabilities around the likelihood of being correct.[21] When users type in a string of words, LLMs parameterize and tokenize the phrases and generate probabilities for what word or content correlates best and returns this

to users.[22] The information they can learn is only bounded by what information they are trained upon and any other constraints humans place on them. These capabilities are augmented when combined with other forms of Generative AI that use multi-modal models to analyze language, video, and audio to generate images, videos, and sounds.[23]

Humans learn in a comparable way, consuming information to extract insights and making inferences about the world. Humanity itself has multimodal sensors that absorb information from sight, sound, touch, and other senses and fuse this information in the mind to create meaning. The computer scientist Judea Pearl contends though that human cognitive function is several steps higher in the complexity rung than machines because humans are not entirely bound by data, allowing them to make logical inferences even when they have no training data or data are scarce.[24] In addition, humans use their imaginations and previous experiences to envisage possibilities and counterfactuals to things like the effects of an intervention in an experiment or actions that may-or-may not trigger war.[25] Most importantly, humans are corrigible and can interrogate past results on their own to search for biases and missing information to better themselves. The human mind then represents an alternative configuration for a reasoning engine, with additional structural features that allow understanding in a more complex manner.

This alternative configuration for understanding the world allows people to interact with the world, form likes and dislikes, and express preferences on things both trivial and significant like musical taste or the person they love.[26] These things are not quantifiable and touch upon what it means to be human. Data and algorithms cannot quantify "flan > vanilla ice cream," for all people in all situations, as these are individual preferences. This rationale applies to values and morality as well. While evolutionary biology suggests that nature hardcodes certain moral traits,[27] humans can reason about morality and then change their behavior.[28] The human ability to experience the universe, articulate preferences about intangibles, and make decisions based on these preferences empowers humans to be moral-agents that make determinations about right and wrong and then work towards creating the life that fits their preferences.[29] In contrast, AI/ML systems cannot make choices beyond what they are designed to do.[30] When it comes to machines, humans pass on their preferences based on the data they feed a machine, with all of the biases and selection effects. If a machine fails at its task, it is because humans did not design it nor feed it with the appropriate information.[31]

These ideas might seem self-evident, but they are easy to gloss over when rattling off the harms associated with AI/ML. The British philosopher Gilbert Ryle coined the phrase "the ghost in the machine" to critique Cartesian dualism that treated the mind as a separate entity from the body, which in turn directed the actions of the body.[32] Ryle argued that mental activities were extensions of the body, and their reasoning power was a by-product of the body's design. The conversation around AI flows in an analogous path, where fears abound of runaway AI harming humanity because of the apparent distinctness and separateness of both. Adapting Ryle's critique to the human-machine relationship, the design and creation of artificial intelligence is an extension of human knowledge and values, thereby making all machine activity a by-product of human consciousness.

Values are not typically the first thing that comes to mind when thinking of counterterrorism, but they underpin the entire counterterrorism enterprise. Strategy reflects a series of policy-choices government makes for combatting extremism, representative of the normative preferences of a society expressed through its leaders (assuming there is a choice in this regard).[33] Repressive leaders, with no qualms about killing can use total force to suppress dissent whereas democratic leaders, legitimized by consensus and who are accountable to their electorate, tend to seek more benign options.[34]

As machines assume the paramount role in the fight against terrorism, the intimate nature of killing, brimming with existentialist consequences for humans, requires that values scaffold counterterrorism. *How* societies fight is just as important as the *choice* to fight. In the age of AI/ML, the task is deploying these tools in a manner that conforms to the values that legitimize a state.[35] With government delegating more and more counterterrorism responsibilities to machines, humans must remain cognizant that they are the moral-entity in the human-machine relationship. The choice to use AI/ML for counterterrorism requires that humans assess the whole enterprise, from the use of force to the second-and-third order effects that can arise from their decisions.

In the CT realm, the successful application of AI/ML will hinge on three factors, all laden with ethical choices: policies governing training data, how governments choose to use these tools, and how well policymakers understand these systems and their outputs. Data are the heart of AI/ML systems. Data quality shapes a machine's biases, which in practice may lead to a tool disproportionately targeting specific groups or creating analytical blind spots. Government decisions on whether data are good enough or if there is a need to generate higher quality to minimize biases will shape machine performance. Likewise, AI/ML offers myriad tools for counterterrorism, but each comes with trade-offs, whether it is for intelligence gathering or targeting extremists. For instance, governments can scrape images from social media at the expense of civil liberties.

The most important consideration is whether and how humans understand the former two problems, and from there, what choices they make. Modern AI/ML's ability to play the proverbial imitation game and give the impression of sentient-intelligence may lead policymakers to suspend critical thinking in favor of the immediacy of a machine result. Decision-makers have an obligation to understand outputs, or if failing that, ensuring they have the right support staff that can do the proper analysis for them. Even then, they must consider if their advisors truly understand a system and are not giving into their own biases.

These ethical considerations do not mean that governments ought to shy away from using AI/ML. There is an argument to be made that they should do the opposite; the benefits reaped will be economic and strategic. Governments now have at their disposal computer systems capable of instantaneously processing data from across various information mediums and platforms, regardless of language, to create meaningful results without the same labor costs of humans doing the same.[36] These systems can generate reports or images that summarize information tailored to the need of the moment without exhausting finite mental resources. Well-trained systems can go a step further and guide the use of other AI/ML systems, generating outputs that can automate various functions of the CT enterprise.[37] The ability to process information faster than people positions AI/ML to pierce the veil of noise from intelligence to find credible signals that security agencies can prioritize to disrupt plots.

Reasoning on the Edge of Forever

Recent history offers perspective on how AI/ML, powered by LLMs and GenAI, will shape counterterrorism. The modern age of CT began on the morning of September 12, 2001. For the better part of a decade, the inertia of the Cold War led the United States to undervalue the threat posed by international terrorism,[38] leaving it unprepared for the Global War on Terror (GWOT).

Its intelligence and defense bodies were staffed by Cold Warriors trained to root out spies, run human networks in denied environments to steal secrets from the Soviet Union, track weapons of mass destruction, and assess the risks posed by hostile nation-states.[39] Its military, trained for winning conventional conflict, had little appetite for protracted counterterrorism campaigns.[40] The military's leadership also maintained a "disdain for special operation forces [SOF]," the one element of the military specifically designed for counterterrorism.[41] Similarly, the intelligence community (IC) suffered from disunity, with its members lacking a coordinating body that would centralize information and set intelligence-gathering priorities.[42] Equally problematic, the IC failed to invest in case officers and linguists knowledgeable in the regions where al-Qaeda operated, such as Afghanistan.[43] It continued, instead, to evaluate individuals based on the standards created for the Cold War that over-indexed on the quantity of information they produced rather than long-term strategic analysis of the threat landscape.[44] This, in turn, meant that the intelligence produced pertained to short-term concerns rather than burgeoning trends like international terrorism in the 1990s, creating analytical blind spots for policymakers. In the background, the few analysts dedicated to terrorism warned of an imminent attack - alerts that went unheeded owing to bureaucratic siloing of intelligence and competing national security priorities.[45] Putting aside the very human failings of this era, the United States' struggle could be summarized as lacking the cognitive edifice to capture data around terrorism and engage in sense-making.

Overnight, these organizations had to pivot to tracking and fighting terrorism abroad and at home. The knowledge on how to do this was inchoate, if non-existent. This had to do with structural deficiencies that predated 9/11. Intelligence bodies had to recruit an army of linguists and analysts while the military had to adapt its training manuals for neutralizing terrorists.[46] Yet, even with the urgency of preventing the next major terrorist attack, the fragmentation of the American national security establishment persisted. In the early years of the GWOT, the IC still could not coordinate its activities, with individual agencies still unwilling to share information or to collaborate with one another.[47] There was also uncertainty about who would lead the GWOT; bureaucratic turf wars pitted different agencies, each with their own counterterrorism strategies, against each other. The Department of Defense (DoD) eventually assumed the paramount kinetic role, with the IC assuming responsibility for supplying it information.[48] This caused its own problems. The DoD's early success in rolling back the Taliban and al-Qaeda in Afghanistan and its successful invasion of Iraq were pyrrhic victories that created the impression that violence sufficed in defeating terrorism.[49] These wins delayed America's security establishment from defining a strategy for defeating terrorism beyond neutralizing extremists.[50] By late 2003, the United States faced an insurgency in Iraq, and soon thereafter, the Taliban began

reconstituting in Afghanistan. Meanwhile, the United States expanded the GWOT across the world without a defined strategy dictating an end-state beyond killing-or-capturing terrorists.

This was not only an American problem, even affecting those countries that never deprioritized terrorism. For instance, Spain had developed a sophisticated CT apparatus for pushing back the Basque separatist group ETA. In spite these CT assets, the country's leadership lacked an appreciation for the threat of Islamic extremism that had rooted itself in the country in the 1990s, culminating in the Madrid train bombings in 2004.[51]

By the middle of the 2000s, the United States addressed some of its structural problems at the tactical and operational level. It created the Office of the Director of National Intelligence (ODNI) to coordinate the IC and provide strategic analysis for the United States.[52] ODNI was a relatively malnourished entity at first; the traditional fights within the IC made members reluctant to collaborate. This body took years to get its footing, but when it did, it became an organization that could furnish the United States a strategic view of terrorism.[53]

At the same time, Joint Special Operations Command (JSOC), perfected its kinetic approach for counterterrorism through its network-centric warfare concept, which would become the model for American counterterrorism efforts more broadly. JSOC showed its mettle during the invasion of Afghanistan, but the U.S. military did not have a concept for integrating it with the rest of the force. This changed when General Stanley McChrystal took charge in 2003 and set about flattening the organization, removing the traditional hierarchy that slowed information sharing. He integrated diverse elements of the military, the intelligence community, and even State Department into joint operation centers (JOCs) – or fusion centers.[54] These had real-time access to different information streams, from human intelligence to satellite imagery, creating a more cohesive picture of the battlespace.[55] Further, McChrystal embedded linguists and analysts trained in advance analytics techniques like social network analysis into the JOCs who then worked with JSOC operators, CIA officers, and drone pilots to analyze intelligence and then execute kill-or-capture missions.[56] Through JOCs, the United States addressed many of the structural issues that prevented the defense community from fulfilling its CT mission by centralizing intelligence to develop its understanding of the threat at the operational level. The United State later exported this model outside of Iraq, making it its preferred approach for counterterrorism across the world.[57] By the end of the decade, the United States finally had an intelligence and kinetic body that could prosecute the GWOT in spite an elusive definition of victory.

Even as governments came to understand terrorism's character and how to fight it militarily, its mutability required constant learning. At the start of the GWOT, al-Qaeda's nucleus in Afghanistan was the main threat. Within a decade, al-Qaeda had franchises that adapted their strategy and tactics to local circumstances, with its Iraqi franchise metastasizing into various strains of terrorist-cum-insurgencies that reached their final form with Islamic State. These changes had consequences for domestic societies in North America and Europe, as lone-wolfs inspired by al-Qaeda and its spinoff plotted and executed attacks.[58] In the background, right-wing terrorism percolated across the United States. By the time the U.S. withdrew from Afghanistan in August 2021, western societies had to contend with different flavors of extremism across the world.

The attacks on 9/11 have been described a failure of intelligence and imagination.[59] Another way to view them is as cognitive failures arising from an overworked labor force, hampered by bureaucracy, that rendered it unable to process the available information. As noted in various sources, the United States' IC had the information necessary to disrupt 9/11. It failed for very human reasons. Intelligence and counterterrorism are laborious activities. Humans attempt to streamline intelligence analysis through the creation of specialized bureaucratic structures that enable information capture, information analysis, policy recommendations, and action driven by intelligence. National security becomes a suboptimal endeavor if the attendant bureaucracy is not properly staffed, organized, and structured for the pooling of information and the dispassionate analysis of the world as it is. This was the hard lesson for the United States. It only started to effectively prosecute the GWOT when it complemented its industrial investments in counterterrorism tools like drones with a commensurate intellectual edifice that helped it understand terrorism.

Future cognitive transformations like that which characterized the GWOT could be streamlined at a fraction of the material and intellectual cost with AI/ML. This extends to the kinetic aspects of the GWOT as well. The advent of AI/ML has occurred with advancements in robotics capable of engaging in combat. The combination of all these artificial intelligence systems and robotic platforms will upend the current CT paradigm. In some ways, the end of the GWOT reflects the twilight of the romanticized image of warriors risking their lives in the name of freedom, as machines become the face of the fight against extremism.

All things being equal, the transition to automation is likely to occur in stages. Counterterrorism encompasses a spectrum of state activities aimed at combating extremism and terrorism that range from non-kinetic to kinetic. No single activity defeats terrorism, but when deployed in an ensemble fashion, they limit the possibilities for terrorist to harm or kill people. These actions must be done continuously to proscribe extremism given its evolutionary character. AI/ML tools can assist or assume responsibility for many of these activities, reducing the human element and thereby enhancing state power.

Counterterrorism in the Age of Intelligent Machines

How might counterterrorism look in the age of intelligent machines? There is not a one-size-fits-all approach to counterterrorism, with local circumstances dictating requirements, whether they be focusing on counter-radicalization or violent countermeasures. AI-enabled CT follows the same logic. In theory, AI/ML can insert itself into any CT activity, but for the sake of simplicity, this section focuses on three tactical activities that have evident operational effects: counter-radicalization, intelligence, and direct action. Nothing within this section is meant to be predictive or prescriptive; it only seeks to inform and provide areas for further research.

Counter-Radicalization and de-Radicalization

Terrorism at its root emerges from extremist beliefs in violence's power to shape political outcomes. From this statement flows a syllogism that offers the pathway for stopping terrorism: the best way to stop terrorism is by preventing people from

adopting extremist views that endorse violence. This syllogism masks the complexity of successful counter-radicalization. Quassim Cassam notes that extremism is best thought of as occurring within a multidimensional plane that encompasses ideological, motivational, and psychological extremism.[60] These respectively capture an individual's political beliefs, openness to violence, and the fervor with which they cling onto their beliefs.[61] This adds difficulty beyond countering alternative interpretations of reality.[62] Permutations in these dimensions create different flavors of extremism, not all of which merit action. For example, Cassam notes that a person may be ideologically extreme but may reject violence as a means for advancing a political agenda while being open to surrendering their beliefs.[63] Similarly, a person may hold politically moderate ideas but hold an extremist mindset that precludes them from updating their priors and still not believe in the necessity of violence.

Particularly for those with extremist mindsets, the human mind is quite adept at rationalizing contradicting facts, whether it is by elevating the credibility of unreliable sources to challenge mainstream narratives or through selective media diets that exclude any data that disconfirms a worldview.[64] In other situations, the grievances are real, and the challenge for democratic societies is to cauterize hostile beliefs to prevent them from festering into justifications for violence.[65] The internet's decentralized nature worsens the problem, as there are many vectors for misinformation and propaganda, with content mutating faster than the counter-messaging governments design.

AI tools like LLMs offer an opportunity to automate counter-radicalization efforts using many of the previously discussed techniques.[66] For instance, using a two-stage process, security agencies can train LLMs on the counter-radicalization canon to identify the cognitive interventions that halt or reverse extremist indoctrination. Armed with this information, LLMs can then consume extremist propaganda, identify weaknesses in messaging, and then generate credible counter-messaging at volume and scale. This approach has limitations, namely the lingering "uncanny valley effect" within LLMs that triggers discomfort in people when they interact with objects that seem human but are not quite human.[67] The uncanny valley though is neutralized by either achieving the initial peak of the valley and staying there, or creating objects that pass the threshold into resembling flesh-and-blood people.[68] Some forms of LLMs and GenAI can accomplish the latter,[69] but in many situations, the former suffices.[70]

As governments train models for counter-radicalization, they will open a suite of tools, such as counter-measuring. Extremists tend to distrust government sources, relying instead on non-traditional sources of information. Counter-messaging must also address variegated worldviews encompassed within the extremism lens. Each person's radicalization journey is unique and non-linear and individuals may hold slightly different beliefs while belonging to a similar movement.[71] LLMs could impersonate an extremist and generate on the spot counter-narratives on forums, chatrooms, and social media platforms, and do so in a dynamic way, adjusting to the content seen online in real-time. A model would not be able to counteract the full constellation of extremist belief on its own, but it could inject enough uncertainty online to sow doubt among believers and overwhelm extremist channels with benign content to drown out the more dangerous rhetoric. At minimum, deradicalizing even a handful of extremists would represent a victory.

There is already evidence suggesting this approach can help counter-radicalization efforts. In 2024, Thomas Costello, Gordon Pennycook, and David G. Rand ran a

large-N experiment where participants who believed in conspiracy theories interacted with a chatbot designed to counter their beliefs.[72] As conversations between participants and the chatbot flowed, the latter adapted its messaging, creating a dynamic and credible response for each point raised by individuals.[73] The researchers found that these interactions reduced individual belief in a conspiracy theory, an effect that lasted for months after the initial interaction.[74]

This type of chatbot can be further specialized to address different and emerging types of extremism. One of the common criticisms levied against LLMs is that they tend to respond to queries with prosaic and generic answers, if not outright nonsense. This has to do with many open-source models being trained on a generic corpus of information rather than specialized datasets that contain more germane information. One of the more promising methods for reducing hallucinations and making LLMs more adept at responding to mutations within extremism is through Retrieval Augmentation Generation (RAG) models.[75] With RAG models, LLMs augment their learning through a steady diet of specialized information, refining their results from these data. This approach currently does not guarantee that an LLM will no longer hallucinate or generate generic responses, but research in improving model accuracy continues apace.[76] There are other methods that are being pursued, such as knowledge refinement through knowledge graphs or the use of model self-correction to improve outputs.[77] In all these approaches, the aim remains the same: creating LLMs that fit-for-purpose through model and data refinement. Indeed, the American military has adapted LLMs for internal use, training them on its own corpus of data for specialized results.[78]

Although the internet plays a major role in radicalization, ties of kith-and-kin play an important role in the journey people take to becoming extremists.[79] The same counter-messaging could be adapted by LLMs for in-person scenarios. Governments could collaborate with civil society groups, such as religious bodies, and use LLMs to help them craft messages that help break the radicalization cycle.[80] In places where the right to privacy is not expected, such as jails, AI/ML tools could process audio-visual data from security cameras to study the pattern-of-life of individuals and use this to enrich counterpropaganda or design interventions by trained psychologists.

In the fight against IS, one of the most effective tools used was deplatforming extremists. Deplatforming limits the reach of extremist content, reducing the chances that someone might stumble upon it and radicalize.[81] After IS used their platforms to spread propaganda, Facebook and Twitter deployed ML algorithms to detect and remove extremist content.[82] This forced IS and other like-minded groups to move their activities onto more obscure parts of the internet that offered stronger privacy controls where their propaganda had less organic reach.[83] A suite of AI/ML tools could scan the internet for these channels, monitor the content, use an LLM to derail conversations online, and automate cyber-operations to take down the servers hosting content, helping the counter-radicalization process in a more forceful way.[84]

There are caveats here. These uses represent generalized views without considerations for different legal regimes and the differences in counter-radicalization during peace and war times. Determining the success rate will be challenging as well. Terrorism, for all the fear it inspires, is a relatively rare event, and the data available are limited.

For every extremist detected, there are an untold number that hide their beliefs and remain undetectable until they finally act out on their beliefs. There are others that likely deradicalize after encountering counter-messaging but are never flagged or identified, creating an entire sample of censored data that algorithms never detect.

This is not to downplay the propaganda potential for states as well. Nation-state disinformation is a problem and this is likely radicalizing individuals. This subject merits interrogation, but this sits outside the scope of this article.

Intelligence and Anomaly Detection

The hallmark of successful counterterrorism strategies is the agile and effective use of intelligence. This is an area where AI/ML can make an immediate impact. Security agencies already employ AI/ML for intelligence gathering,[85] such as geospatial analysts training models to interpret topographical features from satellite and drone images.[86] However, the current approach relies heavily on human interpretation of machine-generated information, which can be cognitively taxing.

LLMs, capable of processing vast amounts of diverse data, can offer much-needed support. Although LLMs like ChatGPT are trained to process text, researchers have developed generative AI systems that use multimodal algorithms to analyze text, images, videos, and other communication forms.[87] This is valuable in virtually any online and offline setting: geospatial imagery recognition, video analysis, social media content exploration, and other intelligence sources. In each of these examples, AI/ML could augment current capabilities with pattern-of-life analyses that alert on deviations.[88] Models deployed on forums can automate social network analysis, identifying changes in leadership,[89] or they can focus on anomaly detection on topographical features from imagery.[90]

This function does not necessarily have to focus on individuals either. In terrorist-cum-insurgencies, public opinion plays a big role in the success of an extremist organization. For example, the success of David Petraeus' Surge came in part from Sunni disenchantment with al Qaeda in Iraq, leading to Sunni leaders cooperating with coalition forces.[91] In more recent years, researchers have used AI/ML for large sentiment analysis to measure public sentiments towards diseases.[92] Governments can adapt algorithms to analyze public chatter online and other mediums of communication to gauge public sentiment towards extremists and the attendant counterterrorism policies. This ability to analyze information in such broad fashion positions AI/ML systems to detect signals indicative of impending plot and to alert security agencies to respond with more alacrity.

What makes these models so adept at this task is that they are language agnostic. A frequent impediment slowing intelligence analysis are language barriers. For some languages, finding capable linguists is not a struggle, such as Spanish speakers in the United States. For other languages, like Arabic, the pipeline is tighter because of the difficulty of acquiring these languages for non-native speakers and the necessary background checks to vet linguists.[93] This is not a problem for LLMs, as they can understand multiple languages, accounting for dialectic differences, in near real-time without being a counterintelligence risk.

An area ripe with potential is audio-visual analysis. The use of satellite imagery has been mentioned already, but this intelligence function can go much further. One of the

most important tasks after a terrorist incident is intelligence gathering from crime scenes, searching for explosive residues or biological markers that might give clues on the perpetrators or victims. Crime scene investigators could use robotic sensors to search for DNA and chemical remnants and use AI/ML systems to reconstruct incidents using principles derived from mechanical engineering and physics with less risk of contaminating evidence.[94] In warzones, this is of great value. At the height of the Global War on Terror, JSOC revolutionized intelligence by collecting, analyzing, disseminating, and acting upon intelligence without pause.[95] As discussed earlier, JSOC would conduct operations, gather information, and use this intelligence to execute other operation within the same night. AI/ML-enabled robots would do this mission with less risk to operators who could enter a booby-trapped house and not worry about explosives.

Perhaps the one area where computer systems could not replace humans is in the collection of human-intelligence. Human-intelligence, or HUMINT, is often the best source of information for combatting extremists, as terrorists tend to avoid means of communication that can inadvertently signal their location or intent.[96] In fact, this was one of the deficiencies of the IC prior to 9/11 where case officers lacked training for collecting HUMINT for terrorist networks that operated in remote rural areas, caves, or in urban environments not accessible to westerners. There is also more basic HUMINT that comes from the relationship between civilians and security forces. For example, effective policing for CT involves police officers developing rapport with locals in a community to incentivize people to share information about their community.[97] Until androids emerge that overcome the uncanny valley effect, this task will remain human. Police officers and other HUMINT collectors could still use AI/ML tools to record, centralize, and disseminate information. Using AI/ML powered wearables like headsets, they can verify people's identities or speak with anyone with generative AI systems specializing in translation.

The examples described represent platonic ideals. How governments integrate LLMs with other artificial intelligence systems will depend on the context. The logic guiding AI/ML use differs between war zones and domestic scenarios and must navigate privacy guidelines that restrict surveillance and monitoring, international humanitarian principles, and the laws of war that make certain actions permissible. Regardless of how these tools are used, humans are likely to play a diminishing role but will not be removed from the process.

Countermeasures

The most controversial application of AI/ML lies in their kinetic potential. The use of unmanned platforms has escalated worldwide ever since the United States carried out the first drone strike in 2001.[98] In August 2018, Venezuelan dissidents nearly assassinated President Nicolas Maduro using two drones carrying explosives while he gave a speech.[99] Robots need not be costly or complex; the Ukrainian military has used modified commercial drones with great success against Russian invaders.[100] A parallel development that is set to accelerate the creation of cheap drones is the increasing availability of 3D printers.[101]

There is no reason to expect a slow-down in the mechanization and robotization of combat scenarios in CT. AI/ML systems have already demonstrated their ability to

outperform humans in aerial dogfighting and to outmaneuver even the most skilled individuals in strategic games.[102] Governments are actively integrating these AI/ML algorithms into drones and other military assets to minimize risks to human soldiers. In May 2024, the U.S. Air Force displayed a fully autonomous F-16, showing that AI/ML is no longer limited to just drones.[103] Importantly, access to these tools is not limited to governments, as numerous AI/ML models are open source and freely available online,[104] which means extremists can acquire similar tools and develop their own violent instruments.

The robotization of violence occurred without LLMs and would have continued developing organically independent of ChatGPT. What is different now is the notion of a reasoning engine furnishing the cognitive framework for the application of force and this is perhaps the most underappreciated change for direct action in counterterrorism. For thousands of years, the judicious use of force has been the sole domain of humans. Societies organized armies and chose leaders to guide their warriors. Samuel Huntington regarded those skilled at the martial arts as professional soldiers; experts at managing violence for political ends.[105] Soldiers designed tactics, operations, and strategies, and contributed to doctrine that distilled best practices for combat. With LLMs, many intermediary professional soldiers that operate between generals and machines will lose importance.

Consider the Iraq War. The United States initially misunderstood al-Qaeda in Iraq, leading to improper tactics and strategies that necessitated a significant intellectual effort to overhaul. The United States required three years and figures like David Petraeus to articulate the strengths and weaknesses of al-Qaeda in Iraq, culminating in the Surge.[106] This experience informed how the country responded to IS when this group emerged in 2013 and 2014. Instead of a large-scale military deployment like 2003, the U.S. and its allies adopted an unconventional warfare strategy, embedding military advisors with host-nation units to train them and enhance their capabilities while providing aerial support to rollback IS.[107]

What AI/ML offers is the ability to examine counterterrorism scenarios with less biases beyond those embedded in algorithms. Large language models, trained on relevant intelligence, can become a source of expert knowledge augmentation, providing insights on tactics best suited for a particular threat. In combat, LLMs can provide operatives automated decision-support, processing data from different sensors, to adjust fire or efforts to detain terrorists. This can escalate up to operational levels. During the GWOT, commanders used drones for situational awareness in combat.[108] LLMs incorporated into AI/ML platforms overseeing engagements can make dispassionate recommendations to commanders suffering from stress and fatigue, ensuring more effective decision-making in high-pressure situations.[109] None of these cases are the exclusive domain of combat elements, as law enforcement bodies can adapt some of these innovations for their specific needs in policing and apprehending extremists.

The idea most discussed and feared is that of killer robots. In one sense, this concern is an update of age-old fears from the Romantic era of humanity's scientific experimentation creating a super-intelligence that rebels against its creators.[110] These fears are palpable, as governments have built the skeletal framework for such thing to emerge with drones and robots, with the taboo around their use in non-combat scenarios eroding. In 2016, the Dallas Police Department used a bomb-defusal robot to

deliver an explosive to a heavily armed shooter that killed several police officers and had pinned down the rest of the force.[111]

These concerns must be weighed against the benefits. Terrorist innovation revolves around finding new ways to harm or threaten harm against civilians.[112] Terrorists can excel at this task, as was the case with IS when it conquered large swaths of Northern Iraq.[113] Government-controlled robots can increase the survival chances for security forces and bystanders alike in scenarios like this, as they could absorb attacks without faltering during hostile engagements with terrorists.

Robots for CT would not just focus on tip-of-the-spear activities, and probably would not assume primary responsibility in this sphere for several years. In the more immediate future, they can operate like dynamic shields that escort police officers or soldiers while overhead drones, powered with multimodal sensors, collect environmental data that give security forces better situational awareness. These type activities would be akin to R2-D2 giving insights during firefights, reducing the fog-of-war, and improving decision-making in high stress situations.[114] With time, CT operations are likely to evolve into situations where a handful of humans deploy AI/ML tools concert, with each AI system specializing in a particular task.[115]

None of this is to say that the risks involved with kinetic activities should be dismissed for the expediency of tracking and eliminating terrorists. AI/ML systems in their current state are prone to errors, some more benign than others, such as ChatGPT's tendency to fabricate facts and sources.[116] While these are humorous examples, poorly designed AI/ML systems can be pernicious, hurting people in systemic ways, without developers being aware. In August 2023, a black woman sued the Detroit Police Department after being arrested by police officers using poorly calibrated facial recognition software.[117] This problem is prevalent in facial recognition software, which tends to generate higher rate of false-positives for non-white people.[118]

This type of unconscious bias exists in all data, and governments must be vigilant. Unlike facial recognition software used to arrest someone, a wrongly trained CT algorithm can kill civilians, an action that cannot be undone. These risks should not lead to moratoriums on the use of AI or prohibition on certain platforms. Rather, they should be calls to action for governments to demand better design for their AI/ML tools. This means acquiring better data and creating strenuous testing regimes that limit risk and expose shortcomings as much as possible. A way to assist this is by enforcing explainable AI standards and transparency policies that give governments insights to the rationale for every action a system takes. These are not easy things to solve and show the immense intellectual edifice governments must develop before the public can trust AI/ML tools to handle most CT activities.

We Think, Therefore It is

This section sought to explore how governments could leverage the reasoning potential of LLMs and combine these with known AI/ML applications for the tactical and operational elements of counterterrorism. Everything described until now has described how AI/ML *can* change counterterrorism. The word "*can,*" however, implies choice. In the final analysis, LLMs and other AI/ML applications, like any other instrument

humans have created for war, continue being that: instruments. No AI/ML program has inherent value-systems or consciousness, and their actions only take on meaning by the way they affect humanity.[119] The character of counterterrorism, even if it becomes robotized, is a human choice.

The potential applications described in this section will largely be the domain of governments and with time will grant them new means for defeating extremists. Certainly, the reduction in costs associated with 3D printing and the fact that many AI/ML algorithms are open-source and available online lead commentators to speculate about the misuse and abuse of this technology in the wrong hands.[120] Most of the proposed uses described in this article though require vast amounts of compute and data to be effective.[121] These are industrial resources that belong to the richest tech companies or nation-states, limiting the scope of actors who can exploit the most sophisticated tools.

This means that governments will always have the upper hand in resources and intelligence, and as LLMs come online designed to mitigate and overcome cognitive blind spots among security analysts, the terrorist element of surprise diminishes.[122] Assuming governments choose to increase their dependency on AI/ML for counterterrorism, in the long run, this can translate into less people suffering from the trauma of violence. Yet, these must be evaluated on their own terms. As discussed in the previous sections, there are a suite of tools for counterterrorism, and governments have choices on which to use and how they are constructed. The harms present in one tool do not necessarily cross-apply to others.

In the counterterrorism context, decision-makers have the ability to interrogate these tools with a series of questions that do not require advanced mathematics degrees. If data creates models and reveals what is important to humans, then what is the breadth and depth of the data? Are there trade-offs in some data that might lead to violation of privacy but improve accuracy, such as zip codes proxying for race and ethnicity? Are there alternative data that help build an equally suitable model or does this model provide intolerable costs in performance? If decision-makers are satisfied with these questions, then they can perform tests that compare humans (incumbent systems) against machines (challengers) and measure their performance. If governments or society are not satisfied with the results, they can scrap a system, revise it, or produce another design. Nothing about the use of AI/ML is deterministic. Governments always have a choice to use or not use a tool, or to find ways to improve a tool. The flipside is the raw power accrued by states, with less humans in the bureaucracy to challenge and veto policies. The word "*can*," unfortunately, opens the door to abuse emerging from both accidental or deliberate misuse.

There is already evidence of how the use of AI/ML for CT can go wrong. Hamas' October 7 attack has been equated to Israel's 9/11. Initial reporting suggested that Israel was caught off guard from a failure of imagination and a leadership dismissive of warnings coming from analysts.[123] These were the same accusations levied against the American security establishment after 9/11. A surprising factor about October 7 is that Israel was actively using AI/ML to identify terrorist threats, something not widely known at the time. In the years leading up to October 7, Unit 8200 – Israel's signal intelligence (SIGINT) body – reorganized itself, moving away from relying on traditional human analysts in favor of AI/ML tools designed to predict terrorist and

insurgent activity.[124] This system failed to alert Israeli leadership of the pending attack even as a human analyst detected the plot and raised alarms. This false-negative was not the only problem with Israel's use of AI/ML.[125] As it began targeting Hamas, it deployed an ensemble of tools to detect and anticipate militants, which had a high false-positive rate, contributing to the death toll in Gaza.[126]

This example is necessary to keep in mind as governments acquire and deploy AI/ML tools. In spite the dizzying pace of innovation, many of these AI/ML-enabled tools are in their infancy. October 7 reflects a military that pivoted too far in the direction of AI/ML while eschewing the very practices that made it successful in the first place. If anything, it suggests a government that has not fully interrogated what AI/ML is supposed to do and how to use it. As has been stated previously, everything about the use of AI/ML is a human choice. Yet, the wishful fantasy of pausing the weaponization of AI/ML is not likely because terrorists and states alike will exploit any tool that offers an advantage in combat. The next best thing then is the assertion of moral agency over CT strategy.

Creating Strategy From the Sum of All Our Wants

Strategy, in its purest form, enables humans to make normative determinations about how to create the world in which they want to live. This elevates war away from mere banal savagery in the name of survival to an aggregation of moral choices to secure the good life.[127] To prevent amoral killings, humans rely upon strategy to guide and control the use of force when actors must confront the inevitable moral dilemmas that arise from conflict.[128] Strategy distilled is an expression of the normative values of society expressing its preferences for the ordering-of-things and the legitimate approach for safeguarding this way of life.[129]

As AI/ML assumes more combat responsibilities, to include targeting and striking, leaders must consider how they fit into their strategies' theories of victory. Victory is a political idea that humans must decide upon that is reflective of the desired end state in which humans want to live.[130] The ability to excel at human tasks does not presuppose victory for the belligerent party deploying the most intelligent machines because machines themselves have no inherent understanding of the concept of victory. Only humans can determine what victory means.[131]

How do these theoretical concepts translate into practical considerations for strategy? First, is that AI/ML cannot define strategy despite the fact it will become the focal point for intelligence gathering and the production of knowledge. Algorithms can provide optimizing functions that might seek to maximize some arbitrary measure of utility with attendant recommended courses-of-action, but they will struggle with disputes about values, as these are not universal.[132] A 2018 study found that western and eastern societies have different outlooks on whether, in the course of reducing accidental deaths, autonomous vehicles should prioritize the lives of the young or the elderly.[133] An algorithm trained to obey human rights cannot adjudicate these disputes to answer if the life of a child is worth more than that of an elderly person. These types of ethical quandaries await governments, as they update their counterterrorism strategies to account for the platforms coming online in the coming years.

The logical consequence of this assertion is the need for governments to assiduously evaluate the design of AI/ML systems: from purpose, scope, and training data to the implementation of safeguards. This idea, which permeates this paper, must go beyond questions of unbalanced or bad data for counterterrorism. Governments need to create strategies for what types of data are necessary and how to evaluate the data received given the moral-strategic frameworks societies construct.[134] These are not do-and-forget activities. These decisions must be continuously re-evaluated to assure their harmony with societal values. Ways to facilitate this can come from efforts to make ML algorithms explainable. There has been advances in this regard by organizations like DARPA (Defense Advance Research Project Agency),[135] but readability will become more salient with the reduction in workforce LLMs are likely to entail. With less humans involved, governments must be certain those still in positions to oversee the deployment of CT applications understand the creation of artificial intelligences and the meaning of their outputs. These stewards must also have the power to override any commands to assure their conformity with counterterrorism strategy. At the time of writing, this seems to have been the failing of Unit 8200 on October 7.

Going a step further, the risk of abuse is potent enough that governments should contemplate additional safeguards around the authorization and control of certain AI/ML platforms. Possibilities include dual validation systems, timed sunsetting of machines to avoid corrosion of their metallic systems that might hamper their functioning, mandatory audits of AI systems and their algorithms, and the creation of kill switches.[136] Systems that disregard policy or strategy owing to failure in design or faulty data should be shut off immediately; humanity should never lose authority or control over its strategic tools. This becomes particularly pertinent at the strategic level where governments must coordinate the deployment of thousands of complex algorithms working together to fulfil a particular policy. Complexity built upon complexity begets entropy. In these situations, governments might find themselves building additional AI applications that help humans manage these systems of systems, which adds to the complexity, creating a vicious cycle.

Second, in spite the lack of a comprehensive legal regime guiding the use of AI/ML, the principles that regulated the use of technology during the GWOT remain in effect in many countries. These laws limit surveillance and intelligence gathering to protect speech, privacy, and the right to life. Just like American law enforcement cannot look at American social media data without warrants,[137] intelligence collection, whether done by humans or algorithm, in a lawless fashion risks making evidence inadmissible in court cases. Likewise, the laws of war and the law of armed conflict do not cease to exist simply because war becomes robotized.

American tech companies in the 2010s defined themselves by breaking rules to force regulators to adapt to their innovations. Security agencies do not have that freedom, and governments across the world will need to formulate laws that guide the use of modern AI/ML tools in such a way that protects constitutional rights without compromising their ability to provision security. What this might look like across different jurisdictions is outside the scope of this article and merits exploration.

Third, the power afforded to governments with AI/ML seems infinite, but counterterrorism through AI/ML should not forget the lessons learned from the GWOT. History is littered with terrorist groups and insurgencies emerging from misrule and

poor governance.[138] States that abuse power have the tendency of begetting reprisals from aggrieved victims and spectators. Timothy McVeigh saw government overreach in Waco Siege and responded by bombing the Oklahoma Federal building.[139]

Neither should governments become overconfident that their technological supremacy will serve as a bulwark against terrorism. Clear strategic goals are necessary. Time after time, technologically superior governments have failed to beat back terrorist groups and insurgencies- a result of governments failing to articulate realistic strategic goals, or for that matter, a strategy in the first place.[140] Twenty years of the GWOT resulted in the Taliban reasserting control over Afghanistan and the persistence of Islamic extremists across North Africa and the broader Middle East despite overwhelming victories by the United States and its allies across many theaters. This feeds into Richard English's commentary about whether counterterrorism works. Between absolute victory and resounding defeat is a constellation of outcomes from partial victories where the public is secured but terrorism remains present to situations where a government wins every tactical engagement but ultimately loses a counterterrorism campaign.[141] AI/ML will offer many victories: it will accelerate intelligence analysis, reduce the number of humans engaged in actual counterterrorism, and harden civilian populations against potential attacks. Yet, unless the use of AI/ML for counterterrorism is aligned to strategic goals, their effects can be self-defeating.

Indeed, a corollary requirement is calibrating the use of force otherwise the tactical use of AI/ML can be self-defeating. Terrorist campaigns often seek to bait governments to overreact to attacks in such a way they alienate the public,[142] whether it is the suspension of certain rights, aggressive use of policing power and surveillance instruments, or indiscriminate violence. Counterterrorism policy, if not grounded with necessity, can validate terrorist grievances, or weaken the social ties between individuals and the government, giving more political room for extremists to operate. An analog example is France's experience fighting the National Liberation Front (FLN) where it won the tactical and operational fight using torture but lost the war owing to the moral disconnect between its tactics and strategy.[143] There are myriad examples of counterterrorism policy succeeding tactically through brutal measures, which then preclude the political solutions needed to achieve some strategic end. The newest generation of AI/ML tools increases this danger because of computers' reaction speed and agnosticism towards actions that affect humans depending on the optimizing function at the heart of a system. Governments must choose wisely if the benefits of increasing surveillance or harming a few by accident with AI/ML tools outweigh the social harms. AI/ML at the tactical level – whether it is for counter-radicalization or countermeasures – can prove counterproductive in the long-run if misused.

These considerations should also feed into decisions surrounding technology sharing. The ascendance of AI/ML is happening in the context of democratic decline across the world and authoritarians consolidating their power. Outside of the Afghan-Iraqi theatres, the GWOT involved western societies capacitating various states to fight terrorism. Many of these governments were imperfect democracies, with some being outright hostile to democracy and human rights.[144] Exporting AI/ML platforms without safeguards can empower autocrats, perhaps making permanent their rule with little possibility for reform owing to asymmetries in technological power to repress.

The focus of this article is AI/ML for counterterrorism strategy, but choices made in this sphere will ripple across society. Like the proverbial sword of Damocles,

governments are now offered what seems like divine power, and with it comes the power to destroy and save lives. This is a responsibility rife with peril. There is nothing deterministic about the future and there is time to proactively engage with these legal ambiguities, policy considerations, and political debates.

Giving Up The Ghost?

Artificial intelligence is a peculiar technology that developed through cyclical spurts of creativity leading to intellectual cliffs where its concepts lay dormant and moribund until being revived again.[145] Throughout these cycles of enthusiasm-and-disappointment, the discourse on how AI/ML could transform humanity never ceased, oscillating from polar extremes of risk and reward. The proliferation of LLMs seems to mark an inflection point affecting the future history of artificial intelligence where this technology seems inextricably linked to modern living. While past conversations should not be dismissed as mere parables about technology, the new wave of AI systems coming online should be interrogated on their own terms to understand both their promise and peril. This type of individual examination provides better insights into how CT can or will change with innovations in artificial intelligence compared to the broad brushstrokes of utopia and doom.

This article argued that the value-proposition of emergent artificial intelligence systems is the reduction of the cognitive load associated with the tactical and operational levels of counterterrorism. The ability to consume, process, and summarize volumes of information, regardless of data type, will streamline decision-making and better prepare governments to respond to terrorism. With time, most of the CT enterprise can be automated with an ensemble of AI platforms. The challenge for governments is figuring out how to build strategies with these new capabilities. To accomplish this, they must first understand the strengths and weaknesses of artificial intelligence, realizing there are certain functions that only humans can execute, which primarily reside at the strategic level.

Strategy is more than ends-ways-means. Strategy encapsulates societal values about the ideal way of life and the means for acquiring and preserving it. These values must inform all the choices surrounding AI/ML's role in counterterrorism, from design, data, and to the actual use. There is no universal answer, and different governments will follow different approaches based on what they perceive as legitimate. A way to explore these questions is by reassessing the lessons of the GWOT and the importance of intelligence, the calibration of violence, and the risks over-reacting. None of what is proposed is easy and perhaps will not be fully settled for decades. But when it comes to artificial intelligence, there is no ghost separate from the machine; there is only human choice, and that is the only way to navigate the future.

Notes

1. A.M. Turing, "Computing Machinery and Intelligence," *Mind: A Quarterly Review of Psychology and Philosophy*, Vol. LIX, No. 236 (October 1950).
2. Samantha Murphy Kelly, "ChatGPT passes exams from law and business schools," *CNN*, January 6, 2023, https://www.cnn.com/2023/01/26/tech/chatgpt-passes-exams/index.html

3. "How good is ChatGPT," *The Economist,* December 8, 2022, https://www.economist.com/business/2022/12/08/how-good-is-chatgpt
4. Sean McManus, "Friend or foe: Can computer coders trust ChatGPT?," *BBC,* March 21, 2023, https://www.bbc.com/news/business-65086798
5. Nuha Aldausari et al, "Video Generative Adversarial Networks: A Review," *ACM Computing Surveys* 5, no. 2 (2022): 7-8.
6. Boaz Ganor, "Artificial or Human: A New Era of Counterterrorism Intelligence?," *Studies in Conflict & Terrorism* 44, no. 7 (2021): 605.
7. Paul Scharre, *Army of None: Autonomous Weapons and the Future of* War (W.W. Norton, 2018), 95; Damien Van Puyvelde, Stephen Culthart, M. Shahriar Hossain, "Beyond the buzzword: big data and national security decision-making," International Affairs 93, no. 6 (2017): 1398–9.
8. Scharre, *Army of None,* 14.
9. Ben Buchanan and Andrew Imbrie, *The New Fire: War, Peace and Democracy in the Age of AI* (The MIT Press, 2022): 17-21.
10. Christopher Wall, "The (Non) Deus-Ex Machina: A Realistic Assessment of Machine Learning for Counter Domestic Terrorism," *Studies in Conflict & Terrorism* 2021: 6-8.
11. "Large language models' ability to generate text also lets them plan and reason: what will come next?" *The Economist*, April 19, 2023, https://www.economist.com/science-and-technology/2023/04/19/large-language-models-ability-to-generate-text-also-lets-them-plan-and-reason
12. "Algorithms and Terrorism: The Malicious Use of Artificial Intelligence for Terrorist Purposes," *United Nations Office of Counter-Terrorism*, 2021, https://www.un.org/counterterrorism/sites/www.un.org.counterterrorism/files/malicious-use-of-ai-uncct-unicri-report-hd.pdf; "Emerging Technologies May Heighten Terrorist Threats," *National Counterterrorism Center,* October 14, 2022, https://www.odni.gov/files/NCTC/documents/jcat/firstresponderstoolbox/134s_-_First_Responders_Toolbox_-_Emerging_Technologies_May_Heighten_Terrorist_Threats.pdf; Gabriel Weimann et al." Generating Terror: The Risks of Generative AI Exploitation," *CTC Sentinel* 17, no. 1 (January 2024): 17-24; Allie Funk, Adrian Shahbaz, and Kian Vesteinsson, "The Repressive Power of Artificial Intelligence," *Freedom House: Freedom on the Net 2023*: https://freedomhouse.org/sites/default/files/2023-11/FOTN2023Final.pdf; Steven Feldstein, "The Road to Digital Unfreedom: How Artificial Intelligence is Reshaping Repression," *Journal of Democracy* 30, no. 1 (January 2019): 40-52.
13. Zachary Kallenborn, "Policy makers should plan for superintelligent AI, even if it never happens," *Bulletin of the Atomic Scientists,* December 21, 2023: https://thebulletin.org/2023/12/policy-makers-should-plan-for-superintelligent-ai-even-if-it-never-happens/; Tamlyn Hunt, "AI Safety Research Only Enables the Dangers of Runaway Superintelligence," *Scientific American,* January 9, 2024: https://www.scientificamerican.com/article/ai-safety-research-only-enables-the-dangers-of-runaway-superintelligence/; Kevin Roose, A.I. Poses 'Risk of Extinction,' Industry Leaders Warn," *New York Times,* May 30, 2023: https://www.nytimes.com/2023/05/30/technology/ai-threat-warning.html
14. Marie Schroeter, "Artificial Intelligence and Countering Violent Extremism: A Primer," *Global Network on Extremism & Technology,* September 28, 2020, https://gnet-research.org/2020/09/28/artificial-intelligence-and-countering-violent-extremism-a-primer/; Ben Buchanan, "Artificial Intelligence and Counterterrorism," *Prepared Testimony and Statement for the Record of Ben Buchanan for the House Homeland Security Committee, Subcommittee on Intelligence and Counterterrorism,* June 24, 2019: https://www.congress.gov/116/chrg/CHRG-116hhrg38781/CHRG-116hhrg38781.pdf; "NATO demonstrates new technology to counter terrorism in crowded venues," *NATO,* May 25, 2022: https://www.nato.int/cps/en/natohq/news_195801.htm; Ganor, "Artificial or Human," 606-609.
15. Michael Walzer, *Just and Unjust Wars: A Moral Argument with Historical Illustrations, Fifth Edition* (Basic Books, 2015): 14-15.
16. Kevin P. Murphy, *Machine Learning: A Probabilistic Perspective* (The MIT Press, 2012), 2-16.

17. For an effort treating AI on its own terms, rather than through myth, see: Jaron Lanier, "There Is No A.I.," *The New Yorker,* April 20, 2023, https://www.newyorker.com/science/annals-of-artificial-intelligence/there-is-no-ai
18. A notable example are Meta's Llama models. See Alex Heath, "Meta releases the biggest and best open-source AI model yet," *The Verge,* July 23, 2024, https://www.theverge.com/2024/7/23/24204055/meta-ai-llama-3-1-open-source-assistant-openai-chatgpt
19. United States Government Accountability Office, "Artificial Intelligence in Health Care: Benefits and Challenges of Machine Learning Technologies for Medical Diagnostics," *Technology Assessment,* September 2022, https://www.gao.gov/assets/gao-22-104629.pdf
20. Sebastien Bubeck et al, "Sparks of Artificial General Intelligence: Early experiments with GPT-4," April 13, 2023, retrieved from arXiv database on May 22, 2023, https://arxiv.org/pdf/2303.12712.pdf
21. The Economist provided a non-technical overview of LLMs work. See "Large, creative AI models will transform lives and labour markets," *The Economist* April 22, 2023, https://www.economist.com/interactive/science-and-technology/2023/04/22/large-creative-ai-models-will-transform-how-we-live-and-work
22. Ibid.
23. Xi Chen and Xiao Wang, "PaLI: Scaling Language-Image Learning in 100+ Languages," *Google Research*, September 15, 2022, https://ai.googleblog.com/2022/09/pali-scaling-language-image-learning-in.html; Arsha Nagrani and Chen Sun, "Multimodal Bottleneck Transformer (MBT): A New Model for Modality Fusion," *Google Research,* March 15, 2022, https://ai.googleblog.com/2022/03/multimodal-bottleneck-transformer-mbt.html
24. Judea Pearl and Dana Mackenzie, The Book of Why: The New Science of Cause and Effect (New York: Basic Books, 2018): 23-36.
25. Ibid.
26. Darcia Narvaez, "Moral Complexity: The Fatal Attraction of Truthiness and the Importance of Mature Moral Functioning," *Perspectives on Psychological Science* 5, no. 2 (March 2010): 163-181; Paul Rozin, "The Process of Moralization," *Psychological Review* 10, no. 3 (May 1999): 219-220.
27. Michael Ruse and Edward O. Wilson, "Moral Philosophy as Applied Science," *Philosophy* 61, no. 236 (April 1986): 173-192.
28. Anthony T. Kronman, "The Value of Moral Philosophy," *Harvard Law Review,* no. 7 (May 1998): 1756-1758
29. Francis Fukuyama, *The End of History and the Last Man* (Free Press, 2006): 141-149.
30. Gary M. Shiffman and Christopher Wall, "It's not the algorithm, it's the ethics," *Journal of Financial Compliance* 6, no. 3 (2023): 227.
31. Ibid.,.
32. Gilbert Ryle, *The Concept of the Mind* (Chicago: The University of Chicago Press, 2000).
33. Philip Bobbitt, The Shield of Achilles: War, Peace, and the Course of History (Anchor Books, 2003): 5-8.
34. Matthew C. Wilson and James A. Piazza, "Autocracies and Terrorism: Conditioning Effects of Authoritarian Regime Type on Terrorist Attacks," *American Journal of Political Science* 57, no. 4 (October 2013): 941-955.
35. Bobbitt, *Shield of Achilles,* 207.
36. LLMs have not been studied in-depth until recently, creating an empirical gap in terms of measuring the productivity effects of LLMs and generative AIs. At the time of writing, one of the first studies to examine this subject is: Shakked Noy and Whitney Zhang, "Experimental Evidence on the Productivity Effects of Generative Artificial Intelligence," *Working Paper,* March 2, 2023, available at: https://economics.mit.edu/sites/default/files/inline-files/Noy_Zhang_1.pdf; It's important to consider that this increased in productivity might lead to loss of employment. Unlike previous waves of automation, this would likely target professions in the knowledge economy like writing or lawyers and this will have consequences that can eventually destabilize societies. This is all speculative and outside the scope of this paper. See: David Rotman, "ChatGPT is about to revolutionize the economy. We need to decide what

that looks like." *MIT Technology Review,* March 25, 2023, https://www.technologyreview.com/2023/03/25/1070275/chatgpt-revolutionize-economy-decide-what-looks-like/.

37. Matthew Sparkes, "Microsoft uses ChatGPT AI to control flying drones and robot arms," *New Scientist,* March 5, 2023, https://www.newscientist.com/article/2361382-microsoft-uses-chatgpt-ai-to-control-flying-drones-and-robot-arms/.
38. National Commission on Terrorist Attacks, *The 9/11 Commission Report: Final Report of the National Commission on Terrorist Attacks Upon the United States* (Official Government Edition, 2004): 86-98.
39. Ibid.
40. Dan Caldwell, *Vortex of Conflict: U.S. Policy Toward Afghanistan, Pakistan, and Iraq* (Stanford Security Studies, 2011), 51.
41. Sean Naylor, *Relentless Strike: The Secret History of Joint Special Operations Command* (St. Martin's Press, 2015), 80.
42. Amy B. Zegart, *Spies, Lies, and Algorithms: The History and Future of American Intelligence* (Princeton University Press, 2022), 66-69.
43. Amy B. Zegart, *Spying Blind: The CIA, the FBI, and the Origins of 9/11* (Princeton University Press, 2007), 40-41.
44. Ibid., 68-69.
45. Amy Zegart explored the issue of information sharing by the CIA and the FBI in two separate articles published in 2007. See Amy B. Zegart, "9/11 and the FBI: The Organizational roots of failure," *Intelligence and National Security* 22, no. 2 (2007): 165-184; Amy. B Zegart, "'CNN with Secrets:' 9/11, the CIA, and the Organizational Roots of Failure," *International Journal of Intelligence and Counterintelligence* 20, no. 1 (2007): 18-49.
46. This problem lasted for more than a decade. As late as 2011, US intelligence agencies could not find enough linguists that could pass a background investigation. Tabassum Zakaria, "U.S. spy agencies struggle with post-9/11 languages," *Reuters,* September 19, 2011, https://www.reuters.com/article/us-usa-intelligence-language/u-s-spy-agencies-struggle-with-post-9-11-languages-idUSTRE78I4P820110919
47. Eric Schmitt and Tom Shanker, *Counterstrike: The Untold Story of America's Secret Campaign Against Al Qaeda* (Henry Holt and Company, 2011), 25-28.
48. Steve Niva, "Disappearing violence: JSOC and the Pentagon's new cartography of networked wafare," *Security Dialog* 44, no. 3 (2013), 191; Schmitt and Shanker, *Counterstrike,* 33.
49. See Schmitt and Shanker, *Counterstrike,* 28-42. They give an overview of the national security establishment's thinking regarding counterterrorism, with President Bush dismissing any approach besides "killing or capturing" terrorists.
50. Ibid., 32-33.
51. Fernando Reinares, "Analisis y evaluación de la política antiterrorista en España," *Jornada sobre terrorismos en el siglo XXI: Su persistencia y su declive* (2004): 1-16, available at: https://dialnet.unirioja.es/servlet/articulo?codigo=5774718
52. Zegart, *Spies, Lies, and Algorithms,* 70-71.
53. Ibid.,72.
54. Niva, "Disappearing Violence," .
55. Niva, "Disappearing violence, 185-202.
56. Ibid.
57. Ibid.
58. Raffaello Pantucci, "A Typology of Lone Wolves: Preliminary Analysis of Lone Islamist Terrorists," *The International Centre for the Study of Radicalisation and Political Violence,* March 2011, 6-7, https://icsr.info/wp-content/uploads/2011/04/1302002992ICSRPaper_ATypologyofLoneWolves_Pantucci.pdf
59. National Commission, *9/11 Commission Report,* 339-344.
60. Quassim Cassam, *Extremism: A Philosophical Analysis* (Routledge, 2022), 14-22.
61. Ibid.,
62. Mattias Gardell, *Gods of the Blood: The Pagan Revival and White Separatism* (Duke University Press, 2003): 5-7.
63. Cassam, *Extremism,* 4-5.

64. Isaac Kfir, "How Cognitive Dissonance Plays Into Violent Extremism," *European Eye on Radicalization,* December 11, 2020, https://eeradicalization.com/how-cognitive-dissonance-plays-into-violent-extremism/.
65. Jamie Bartlett and Carl Miller, "The Edge of Violence: Towards Telling the Difference Between Violent and Non-Violent Radicalization," Terrorism and Political Violence 24, no.1 (2012): 1-21.
66. Priyank Mathur, Clara Broekaert, and Colin P. Clarke, "The Radicalization (and Counter-radicalization) Potential of Artificial Intelligence," International Centre for Counter-Terrorism, May 1, 2024, https://www.icct.nl/publication/radicalization-and-counter-radicalization-potential-artificial-intelligence
67. The "uncanny valley effect" has not been overcome completely at the time of writing but it is something that can be overcome with specialized data and training. See Kristina Radivojevic et al., "Human Perception of LLM-generated Text Content in Social Media Environments," September 10, 2024, retrieved from arXiv database on November 1, 2024, https://arxiv.org/pdf/2409.06653v1; With diligent training, it is possible to overcome this effect. See: Kevin Roose, "A Conversation With Bing's Chatbot Left Me Deeply Unsettled," *New York Times,* February 16, 2023, https://www.nytimes.com/2023/02/16/technology/bing-chatbot-microsoft-chatgpt.html
68. Masahiro Mori, the person that conceptualized the "uncanny valley" effect argued that rather than trying to move into the second peak of the valley, engineers should aspire to reach the first valley and remain there. Masahiro Mori, "The Uncanny Valley: The Original Essay by Masahiro Mori," *IEEE Spectrum,* June 12, 2012, https://spectrum.ieee.org/the-uncanny-valley
69. For a recent example of GenAI overcoming the "uncanny valley" see Alexander Diel and Michael Lewis, "Deviations from typical organic voices best explains a vocal uncanny valley," *Computers in Humans Behaviors* 14 (May 2024), https://doi.org/10.1016/j.chbr.2024.100430
70. Mathur, Broekaert, and Clarke in the "The Radicalization Potential of Artificial Intelligence" describe a situation where an extremist self-radicalized with his own fine-tuned model. Even if models can create discomfort, they still offer vectors for propaganda and counter-messaging. The desired effect is dependent on what a government wishes: full mimicry or a tool that is good enough at counteracting extremism messaging. Good enough would lessen the importance of the uncanny valley effect.
71. John Horgan, "From Profiles to Pathways and Roots to Routes: Perspectives from Psychology on Radicalization into Terrorism," Annals of the American Academy of Political and Social Science 618, no. 1 (2008): 80–94.
72. Thomas H. Costello, Gordon Pennycook, and David G. Rand, "Durably reducing conspiracy beliefs through dialogues with AI," *Science* 385, no. 6714 (September 13, 2024), DOI: 10.1126/science.adq1814
73. Ibid.,.
74. Ibid.,.
75. Yuning Mao, Penchen He, Xiaodong Liu, Yelong Shen, Jianfeng Gao, Jiawei Han, and Weizhu Chen, "Generatoin-Augmented Retrieval for Open Domain Question Answering," September 17, 2020. Retrievied from *arxiv database on December 31, 2024,* https://arxiv.org/abs/2009.08553
76. Kevin Wu, Eric Wu, and James Zou, "*ClashEval:* Quantifying the tug-of-war between an LLM's internal prior and external evidence," June 9, 2024, retried *arxiv database* on December 31, 2024, https://arxiv.org/pdf/2404.10198
77. Mengjia Niu, Hao Li, Jie Shi, Hamed Haddadi, Fan Mo, "Mitigating Hallucinations in Large Language Models via Self-Refinement-Enhanced Knowledge Retrieval," May 10, 2024, retrieved from arxiv database on December 31, 2024, https://arxiv.org/pdf/2405.06545v1; Sajad Mousavi, Ricardo Luna Gutierrez, Desik Rengarajan, Vineet Gundecha, Ashwin Ramesh Babu, Avisek Naug, Antonio Guillen, and Soumyendo Sarkar, "N-Critics: Self-Refinement of Large Language Models with Ensemble of Critics," October 28, 2023, retried from arxiv database on December 31, 2024, https://doi.org/10.48550/arXiv.2310.18679

78. Sydney J. Freedberg Jr., "How the Pentagon domesticated the chatbot: 2024 in review," *Breaking Defense,* December 30, 2024, https://breakingdefense.com/2024/12/how-the-pentagon-domesticated-the-chatbot-2024-in-review/.
79. Mohammed M. Hafez, "The Ties that Bind: How Terrorists Exploit Family Bonds," *CTC Sentinel* 9, no.2 (2016), https://ctc.usma.edu/the-ties-that-bind-how-terrorists-exploit-family-bonds/.
80. See Patricia Cogswell recommendation in Wall, "Ex Machina,"15.
81. Bennett Clifford and Helen Christy Powell, "De-platforming and the Online Extremist's Dilemma," Lawfare, June 6, 2019, https://www.lawfareblog.com/de-platforming-and-online-extremists-dilemma
82. Sam Schenchner, "Facebook Boosts AI to Block Terrorist Propaganda," The Wall Street Journal, June 15, 2017, https://www.wsj.com/articles/facebook-boosts-a-i-to-block-terrorist-propaganda-1497546000; Natasha Lomas, "Twitter claims more progress on squeezing terrorist content," Tech Crunch, April 5, 2018, https://techcrunch.com/2018/04/05/twitter-transparency-report-12/.
83. Clifford and Powell, "De-platforming".
84. Laura Hanu, James Thewlis, and Sasha Haco, "How AI Is Learning to Identify Toxic Online Content," Scientific American, February 8, 2021, https://www.scientificamerican.com/article/can-ai-identify-toxic-online-content/; Wall, "Deus-Ex Machina," 15.
85. C. Todd Lopez, "National Geospatial-Intelligence Agency in Midst of Revolution," *DOD News,* December 15, 2020, https://www.defense.gov/News/News-Stories/Article/Article/2447871/national-geospatial-intelligence-agency-in-midst-of-revolution/.
86. Juergen Dold and Jessica Groopman, "The future of geospatial intelligence," *Geo-Spatial Information Science* 20, no. 2 (2017): 151-162.
87. Jiquan Ngiam, Aditya Khosla, Mingyu Kim, Juhan Nam, Honglak Lee, and Andrew Y. Ng, "Multimodal Deep Learning," in Proceedings of the 28th International Conference on Machine Learning (ICML-11), 2011, 689-696, https://ai.stanford.edu/~ang/papers/icml11-MultimodalDeepLearning.pdf.
88. Paul K. Davis et al, *Using Behavorial Indicators to Help Detect Potential Violent Acts* (Rand Corporation, 2013): 69-75.
89. Arie Perliger and Ami Pedahzur, "Social Network Analysis in the Study of Terrorism and Political Violence," *PS: Political Science and Politics* 44, no. 1 (January 2011): 46.
90. Dold and Groopman, "geospatial intelligence," 159.
91. Stephen Biddle, Jeffrey A. Friedman, and Jacob N. Shapiro, "Testing the Surge: Why Did Violence Decline in Iraq in 2007," *International Security* 37, no. 1 (Summer 2012): 7-40.
92. Staphord Bengesis et al, "A Machine Learning-Sentiment Analysis on Monkeypox Outbreak: An Extensive Dataset to Show the Polarity of Public Opinion From Twitter Tweets," *IEEE Access* 11 (2023): 11811-11826.
93. For an overview of the difficulties, see: "A national security crisis: foreign language capabilities in the federal government:" hearing before the Oversight of Government Management, the Federal Workforce, and the District of Columbia Subcommittee of the Committee on Homeland Security and Governmental Affairs, United States Senate, 112th Congress, second session, May 21, 2012 (Statement of Tracey A. North, Deputy Assistant Director, Directorate of Intelligence, FBI) available at https://archives.fbi.gov/archives/news/testimony/a-national-security-crisis-foreign-language-capabilities-in-the-federal-government;
94. Partridge et al, "Enhanced detection of threat materials by dark-field ex-ray imaging combined with deep neural networks." *Nature Communications* 13 (2022): 1-12; Ostrinskaya et al, "Rapid Quantitative Analysis of Multiple Explosive Compound Classes on a Single Instrument via Flow-Injection Analysis Tandem Mass Spectrometry," *Journal of Forensic Sciences* 64, no. 1 (January 2019): 1-8.
95. Liam Collins, "Rapid And Radical Adaptation in Counterinsurgency: Task Force 714 in Iraq," *Modern War Institute at West Point,* September 28, 2021, https://mwi.usma.edu/rapid-and-radical-adaptation-in-counterinsurgency-task-force-714-in-iraq/.
96. Matt A. Mayer, "Enhanced Human Intelligence Is Key to Defeating Terrorists," *American Enterprise Institute,* June 16, 2016, https://www.aei.org/research-products/report/enhanced-human-intelligence-is-key-to-defeating-terrorists/.

97. Adrian Cherney and Jason Hartley, "Community Engagement to tackle terrorism and violent extremism: challenges, tensions, and pitfalls," *Policing And Society* 27, no. 7 (2017): 750-763.
98. Chris Woods, "The Story of America's Very First Drone Strike," *The Atlantic,* May 30, 2015, https://www.theatlantic.com/international/archive/2015/05/america-first-drone-strike-afghanistan/394463/.
99. Ana Vanessa Herrero and Nicholas Casey, "2 Blasts, a Stampede and a 'Flying Thing': Witnesses Tell of Attack on Maduro," *The New York Times,* August 5, 2018, https://www.nytimes.com/2018/08/05/world/americas/venezuela-drone-attack-nicolas-maduro.html
100. Greg Myre, "A Chinese drone for hobbyists plays a crucial role in the Russia-Ukraine war," *NPR,* March 28, 2023, https://www.npr.org/2023/03/21/1164977056/a-chinese-drone-for-hobbyists-plays-a-crucial-role-in-the-russia-ukraine-war
101. Pranshu Verma, "How the 3D-printing community worldwide is aiding Ukraine," *The Washington Post,* June 12, 2022, https://www.washingtonpost.com/technology/2022/06/12/3d-printers-ukraine-war-supplies/.
102. Patrick Tucker, "An AI Just Beat a Human F-16 Pilot In a Dogfight – Again," *Defense One,* August 20, 2020, https://www.defenseone.com/technology/2020/08/ai-just-beat-human-f-16-pilot-dogfight-again/167872/; Matthew Hutson, "AI Learns the art of Diplomacy," *Science,* November 22, 2022, https://www.science.org/content/article/ai-learns-art-diplomacy-game
103. Tara Copp, "An AI-controlled fighter jet took the Air Force leader for a historic ride. What that means for war," *Associated Press,* May 3, 2024, https://apnews.com/article/artificial-intelligence-fighter-jets-air-force-6a1100c96a73ca9b7f41cbd6a2753fda
104. Meta is a prime example of this. See: "Models and libraries, *Meta,* https://ai.meta.com/resources/models-and-libraries/(accessed December 31, 2024).
105. Samuel P. Huntington, The Soldier and the State: The Theory and Politics of Civil-Military Relations (The Belknap Press of Harvard University Press, 1957): 11-18.
106. Thomas E. Ricks, *The Gamble: General Petraeus and the American Military Adventure in Iraq* (Penguin Books, 2010).
107. Michael Gordon, "Explainer: U.S. Strategy to Defeat ISIS," *Wilson Center,* September 30, 2022, https://www.wilsoncenter.org/article/explainer-us-strategy-defeat-isis
108. Pablo Chovil, "Air Superiority Under 2000 Feet: Lessons from Waging Drone Warfare Against ISIL," *War On the Rocks,* May 11, 2018, https://warontherocks.com/2018/05/air-superiority-under-2000-feet-lessons-from-waging-drone-warfare-against-isil/.
109. Wall, "Ex Machina," 15-16.
110. Scharre, *Army of None,* 234.
111. Sam Thielman, "Use of police robot to kill Dallas shooting suspect believed to be first in US history," *The Guardian,* July 8, 2016, https://www.theguardian.com/technology/2016/jul/08/police-bomb-robot-explosive-killed-suspect-dallas
112. Audrey Kurth Cronin, *Power to the People: How Open Technological Innovation is Arming Tomorrow's Terrorists* (Oxford University Press, 2020): 61-83; Bruce Hoffman, *Inside Terrorism, Third Edition* (Columbia University Press, 2017): 265-267
113. Suadad Al-Salhy and Tim Arango, "Sunni Militants Drive Iraqi Army Out of Mosul," *The New York Times,* June 10, 2014, https://www.nytimes.com/2014/06/11/world/middleeast/militants-in-mosul.html.
114. Mike Pietrucha, "Building R2-D2," *War on the Rocks,* March 29, 2023, https://warontherocks.com/2023/03/building-r2-d2/.
115. Zachary Kallenborn, "The Era of the Drone Swarm Is Coming, And We Need To Be Ready For It," *Modern War Institute at West Point,* October 25, 2018, https://mwi.usma.edu/era-drone-swarm-coming-need-ready/.
116. Karen Weise and Cade Metz, "When A.I. Chatbots Hallucinate," *The New York Times,* May 1, 2023, https://www.nytimes.com/2023/05/01/business/ai-chatbots-hallucination.html.
117. Jennifer Henderson, "Black mom sues city of Detroit claiming she was falsely arrested while 8 months pregnant by officers using facial recognition software," *CNN,* August 8, 2023, https://www.cnn.com/2023/08/07/us/detroit-facial-recognition-technology-false-arrest-lawsuit/index.html

118. The National Institute of Standards and Technology did a comprehensive survey of facial recognition software in 2019. For an overview, see "NIST Study Evaluates Effects of Race, Age, Sex on Face Recognition Software," *NIST,* December 19, 2019, https://www.nist.gov/news-events/news/2019/12/nist-study-evaluates-effects-race-age-sex-face-recognition-software; For the full study, see Patrick Grother, Mei Ngan, Kayee Hanaoka, Face Recognition Vendor Test FRVT) Part 3: Demographic Effects, (National Institute of Standards and Technology,
December 2019), https://nvlpubs.nist.gov/nistpubs/ir/2019/nist.ir.8280.pdf
119. Shiffman and Wall, "It's not the algorithm," 227.
120. Dan Sabbagh, "Terrorists could try to exploit artificial intelligence, MI5 and FBI chiefs warn," *The Guardian,* October 18, 2023, https://www.theguardian.com/technology/2023/oct/18/terrorists-exploit-artificial-intelligence-ai-mi5-fbi-chiefs-warn; David Gilbert, "Here's How Violent Extremists Are Exploiting Generative AI Tools," *Wired,* November 9, 2023, https://www.wired.com/story/generative-ai-terrorism-content/;
121. Buchanan and Imbrie, *The New Fire,* 59.
122. Wall, "Ex Machina," 2.
123. Ronen Bergman and Adam Goldman, "Israel Knew Hamas's Attack Plan More Than a Year Ago," *New York Times,* November 30, 2023, https://www.nytimes.com/2023/11/30/world/middleeast/israel-hamas-attack-intelligence.html
124. Elizabeth Dwoskin, "Israel built an 'AI factory' for war. It unleashed it in Gaza.," *Washington Post,* December 29, 2024, https://www.washingtonpost.com/technology/2024/12/29/ai-israel-war-gaza-idf/.
125. Ibid.
126. Ibid.
127. Walzer, *Just and Unjust Wars*
128. Ibid., 225-232.
129. Bobbitt, The Shield of Achilles: War, 5-8.
130. J. Boone Bartholomees, "Theory of Victory," *Parameters* 38, no. 2 (2008): 26-29.
131. Ibid.
132. Shiffman and Wall, "Not the algorithm," 224-225.
133. Edmon Awad et al, "The Moral Machine Experiment," *Nature* 563 (November 2018) 59-64, https://www.nature.com/articles/s41586-018-0637-6. The researchers behind this study in Nature used the Trolly Car Problem to frame their investigation. Theoretically, the notion of one-off ethical scenarios are unrealistic. Heather M. Roff ("The folly of trolleys," *Brookings, December 17, 2018,* https://www.brookings.edu/articles/the-folly-of-trolleys-ethical-challenges-and-autonomous-vehicles/) makes the case that with self-driving cars, what can be conceived of the scope of decision - ethical space – is a dynamic system that changes with each movement a vehicle makes. The present author agrees with Roff's framing that at the tactical and operational level, the trolly framing is not useful because every decision opens new opportunities for action. This differs at the national security level where Government principles make policy choices and then use the instruments of power to execute them. These choices are often either/or situations. For a related example, see Stephen B. Wicker, "The Ethics of Zero-Day Exploits – The NSA Meets the Trolley Car," *Communication of the ACM* 64, no. 1, January 2021: 97-103.
134. The EU has made strides in this direction. For an overview, see Spencer Feingold, "The European Union's Artificial Intelligence Act, explained," *World Economic Forum,* March 28, 2023, https://www.weforum.org/agenda/2023/03/the-european-union-s-ai-act-explained/#:~:text=The%20Artificial%20Intelligence%20Act%20aims,of%20AI%20for%20industrial%20use.%E2%80%9D
135. David Gunning and David W. Aha, "DARPA's Explainable Artificial Intelligence Program," AI Magazine 40, no.2 (2019), 45.
136. The EU has also proposed kill switches. Kelly Fiveash, "Treat robots as "electronic persons" but with kill switches, argue MEPs, "*Ars Technica,* January 13, 2017, https://arstechnica.com/tech-policy/2017/01/robots-electronic-persons-ai-kill-switches-eu-committee/.
137. Wall, "Ex-Machina," 11.

138. Nazli Avdan and Gary Uzonyi, "V for Vendetta: Government Mass Killing and Domestic Terrorism," *Studies in Conflict & Terrorism* 40, no. 11 (2017): 934-965.
139. Richard Abanes, American Militias: Rebellion, Racism, & Religion (InterVarsity Press, 1996), 45–9.
140. Richard English, *Does Counter-Terrorism Work?,* (Oxford University Press, 2024), 163-164.
141. Ibid., 4-5.
142. David Kilcullen, *The Accidental Guerrilla: Fighting Small Wars in the Midst of a Big One* (Oxford University Press, 2009), 28-38.
143. Alistair Horne, *A Savage War of Peace: Algeria 1954-1962* (New York Review of Books, 2006),.
144. Daniel L Byman, "The U.S.-Saudi Arabia counterterrorism relationship," *Brookings Institute,* May 24, 2016, https://www.brookings.edu/testimonies/the-u-s-saudi-arabia-counterterrorism-relationship/.
145. Marc Losito and John Anderson, "The Department of Defense's Looming AI Winter," *War on the Rocks,* May 10, 2021, https://warontherocks.com/2021/05/the-department-of-defenses-looming-ai-winter/.

Acknowledgements

The author would like to acknowledge the use of Microsoft's Copilot, embedded within Office 365, for some minor language refinement to make the document more readable. The author did not use Generative AI for any form of idea generation or exploration, nor did the author use it to write content beyond the noted efforts to refine the language for clarity.

Disclosure statement

No potential conflict of interest was reported by the author(s).