The (Non) Deus-Ex Machina: A Realistic Assessment of Machine Learning for Countering Domestic Terrorism[1]

Christopher Wall[2]

Abstract: In light of the January 6 insurrection, the Department of Homeland Security (DHS) and other national security agencies are looking towards using more artificial intelligence (AI) and machine learning (ML) tools to detect and combat extremism in America. AI and ML hold much promise for the domestic CT mission, but the discourse has placed on them unrealistic expectations that do not conform to what is technically possible. This essay seeks to create a baseline conversation about what is ML, how it actually works, and what is a more realistic use case for ML in domestic CT. The core argument is that current ML tools are not optimal for the CT enterprise because terrorism experts are often sidelined in the development and the implementation of these algorithms.

[1] The author of this paper would like to give a special thanks to Gary Shiffman and Harsh Pandya. Both played critical roles in the genesis of this project through their perspicacious critiques of the state-of-the-art of machine learning as it is applied to the counterterrorism enterprise. Their perspectives prompted the initial investigation into the disconnect between the promise of ML and its discomfiting reality. The author would also like to express his gratitude to Patricia Cogswell, Brian Drake, and Shane Quinlan for carving time out of their busy schedules to converse at length about the policy and technical aspects of machine learning for counterterrorism. They read early drafts, provided important feedback, and gave the proper policy context that illuminated the challenges of data and CT. Lastly, the author would like to thank Dylan Marshall and Alexander Meleagrou-Hitchens who served as sounding boards for many of the ideas contained within the while they were in an inchoate state. Machine-learning, in spite the simplicity implied by its name, can lead to analytical labyrinths, and those who supported this project were the proverbial torch bearers that prevented me from getting lost.

[2] War Studies PhD Candidate at King's College London, Adjunct Professor at Georgetown's Security Studies Program, and Social Scientist at Giant Oak. Author can be reached at christopher.s.wall@kcl.ac.uk

The emergence of artificial intelligence (AI) and machine learning (ML) as integral components of the United States' national security toolkit coincided with the inauguration of the global on war on terror. To the American intelligence community, the shock of 9/11 rendered the entire world a potential battlefield, requiring it to orient its full information gathering apparatus to track and target al-Qaeda and its affiliates. In the early years, the volume of information gathered by the sensors in service of the intelligence community was too much for humans to analyze, spurring the development of algorithms capable of rapidly detecting patterns and highlighting anomalies or indicators of risk to the country, improving analysts' understanding of the terrorist threat.[1] This work is now laid bare across battlefields stretching from North Africa to South Asia, where ML algorithms power drones,[2] satellites, and cyber-bots tracking and monitoring terrorists and insurgents alike across cellular networks, the internet, and financial systems, to produce data that is later combined with human-intelligence to offer perspective into extremist activities once thought solely the domain of science fiction.[3] With a near apodictic certainty, the United States can now order kinetic operations while minimizing civilian casualties owing to its mastery of data.[4] None of this would be possible without machine-learning.

True AI does not exist yet.[5] When commentators discuss AI, they refer to ML algorithms trained to discern trends among data to make decisions that reflect human preferences.[6] The power of ML algorithms for counterterrorism resides in their ability to detect patterns at scale, removing some of the fog of war to allow humans to focus their mental resources on higher order tasks like inference and policymaking.[7] This makes ML a force-multiplier, as it enhances the neurocognitive capabilities of humans, allowing individuals to process information more efficiently.[8] Or, as Brian Drake, the Defense Intelligence Agency's former Director for Artificial Intelligence, noted in an interview for this paper, ML flattens information, highlights insights not

discernable through a purely human review, and has a greater chance of reducing cognitive biases for analysts.[9] As the United States changes its security posture to deprioritize counterterrorism in favor of great power competition across Eurasia and the Western Pacific,[10] ML will enable the United States to maintain an economy of force of effort in places like Afghanistan and Iraq, with drones and other sensors collecting and analyzing data in real-time about al-Qaeda and Islamic State, which can then inform decisions to strike at militants before they threaten the homeland or its security interests abroad.[11]

Yet for all the benefits machines powered by algorithms offer for fighting terrorism abroad, domestically, there has been more apprehension towards moving forward with using ML for counterterrorism over fears of its intrusiveness and the risks these tools pose to civil liberties.[12] These discussions have become more urgent to address in response to the January 6 insurrection. As NBC news reported in March 2021, the Department of Homeland Security (DHS) and other agencies are now augmenting their current technological suite with ML-enabled platforms capable of scraping social media content to detect potential terrorists and terrorist plots.[13] This change in posture comes from how brazenly the January 6 insurrectionists coordinated their activities on platforms like Gab, Facebook, and Parler,[14] and the intelligence failures surrounding law enforcement not acting more forcefully on these leads.[15] This is not necessarily a break in policy but rather an acceleration in trends, as before 2021, various national security agencies had invested in the use of big data platforms to collect and analyze biometric data, target cross-border trade and travel to detect illicit activity, to cross-reference DNA samples, and other activities.[16]

ML tools hold much promise for countering domestic terrorism, but caution is warranted. In spite the vast experience the United States has in fighting terrorism abroad, the tools

developed for battlespaces like Afghanistan or Yemen are not necessarily optimal in a domestic context. In these spaces, the Intelligence Community (IC) has a general sense of expressed behaviors of terrorists for as economist Gary Shiffman explains, humans are predictable and rational,[17] and when aggregating individual data, it is possible to identify outliers or individuals engaging in patterns of negative behavior. The IC has greater leeway in collecting information overseas, enabling ML tools to detect patterns-of-life and anomalies. This permits the IC and the military to operate with more certainty and leverage sophisticated war-fighting tools as permitted under international law.

At home, the U.S. government is not at war with anyone, even if extremists routinely declare war against the state.[18] This means that any machine-learning powered counterterrorism policy must align to a constitutional framework, something that will curtail the most ambitious efforts to use ML for threat detection within social media. The country's freedom of speech tradition makes its online milieu a free for all in terms of discourse,[19] as nearly all forms of rhetoric are permissible unless they deal with state secrets or the direct incitement of violence.[20] This raises the bar for law enforcement agencies who use ML who need to differentiate between individuals who express hateful ideas from extremists seeking to radicalize others to commit harm or to commit violence themselves. Indeed, American jurisprudence has sought to create a right to privacy that severely curtails precisely the type of mass surveillance that sifting through volumes of social media data without warrants implies.[21] But the most important consideration is the risk of false-positives. If the United States were to mistakenly target the wrong person using an ML algorithm designed for warfighting, it could incite further violence and radicalize others, as was the case with Waco in the early 1990s.[22] There could also be further negative externalities besides violence. For instance, Congress could respond by passing legislation that prohibits the

use of ML in domestic jurisdiction or it could sow public distrust towards law enforcement and the government.

In this sense, the central ontological problem of terrorism studies, defining terrorism,[23] portends real-world consequences when algorithms are tasked with mapping human meaning of terrorism and extremism among disparate datasets. As mentioned before, ML, in the most simplistic terms, replicates and industrializes human behaviors, as computers make autonomous decisions informed by training data fed to it by developers and researchers. Algorithms succeed or fail based on the training data.[24] In the past ten years, ML has shown its effectiveness in situations with limited discrete outcomes or closed systems with clearly defined rules that offer little variance, such as helping a person purchase goods based on their established preferences or the optimal speed for self-driving cars on roads with clearly demarcated signs.[25] This is not the case for domestic extremism, where extremist groups and individuals constantly evolve and mutate, evading a stable definition, undermining the value of pre-existing datasets.

However, ML algorithms *can* augment American domestic counterterrorism efforts and *can* play important roles in the context of countering domestic terrorism of any kind in the coming decade, especially when it comes to considering extremist activity online. But ML must be calibrated to a specific problem. To do this, there needs to be a discussion of what ML algorithms can accomplish in a practical sense, especially when removing the war-fighting rhetoric used for counterterrorism in a foreign context and thinking about the optimal application of ML for domestic CT. The terrorism literature has not delved into this because it has prioritized answering the normative question of *should* ML be used for counterterrorism, and when answered in the affirmative, *where* should it be used.[26] To truly answer the practical question,

there needs to be greater exploration of *how* ML works, as it sets the conditions for the *should* and *where* questions.

This essay will give a brief overview of the state-of-the-art on machine learning and the social sciences, discuss its limitations for domestic CT and how it can reasonably be applied in ways that are effective, but may not be as immediately apparent as options. The argument advanced is that ML-algorithms are powerful tools that can help in a discrete fashion, but owing to both technological and policy questions, cannot be used as a panacea nor in such a broad fashion as the United States has used them in foreign battlespaces. These limitations, however, present opportunities for the country to use ML-algorithms in creative ways, particularly when it comes to non-kinetic aspects of CT like counter-radicalization and enhancement of end-user decision-making process.

The Business of Machine Learning

The baseline conversation on the use of ML tools for counterterrorism presupposes the existence and efficacy of these tools as a given. The literature cites the capabilities of companies like Google or IBM and the technological space carved out by big data companies like Palantir or Dataminr whose software supports counterterrorism efforts across the world.[27] Analytically, this is problematic. These companies could not be more different in terms of what they do, but they are grouped together as manifestations of the same thing because they use algorithms powered by statistics. Dataminr, for instance, scrapes social media content and builds text models on Twitter data.[28] In contrast, Google touches everything from voice recognition software to tools for combating cancer,[29] and the type of modeling required for these tasks differ in scope, scale, and process. All that these companies have in common is that they use

algorithms to parse out big data, but that is very general statement and does not say much about *how* they parse out data or *how* effective these techniques are. The *how* questions are coming to fore for many of these companies, however. For example, scholars have started questioning IBM and the safety of the recommendations pushed by its ML-powered health initiatives, raising concerns about *how* it has trained some of its algorithms.[30]

Partially, the inability to the answer the *how* comes from the nature of business. Tech companies are incentivized to protect intellectual property and are reluctant to reveal their trade secrets, leaving scholars to infer what these tools do based on what scant information exists in the public domain. To resolve this, the literature works by analogy from war games to hypothesize what machine learning could do for the future of combat.[31] This is not to say there is not public record documentation of ML tools supporting CT. In Israel where there is less compunction collating meta-data on individuals, Israeli security forces have leveraged a wide range of analytics, including Palantir, to counter the lone-wolf challenge.[32] But not all these analytics are examples of algorithms making decision on their own nor much is known about the algorithms themselves, turning the software design component into a black box left untouched by terrorism scholars.

These issues point to analytical problems for the terrorism studies literature regarding how it interrogates these tools. As machine learning becomes synonymous with modern counterterrorism, ML operates like a nebulous concept that *does something* that supports the mission, but that *something* is glossed over in favor of focusing solely on outcomes dictated by technology companies. Whether the United States uses facial recognition software to identify terrorists or to detect IEDs,[33] the focus is on what tools the United States has at its disposal and how it uses them. This is problematic when poorly understood black box algorithms inform

decisions that materially affect the general public's wellbeing. In fact, a concern among data scientists is that the end users for machine-learning outputs often struggle in making sense of the information,[34] putting the onus on technology companies to explain it. Unless some neutral arbitrator validates claims by technology companies, the government depends on those producing these tools to be faithful stewards of the counterterrorism enterprise while balancing their profit-seeking motives.

The emphasis on ML as a given also relegates terrorism scholars to a secondary role where they only respond to technological innovation and its effects on the world rather than exerting agency and driving innovation in terms of developing machine-learning tools that more optimally support the counterterrorism mission. As Patricia Cogswell, who served as the former Deputy Administrator at the Transportation Security Administration and the former Acting Undersecretary for the Office of Intelligence and Analysis at DHS, noted in an interview for this paper, ML companies strive to meet the demands of the consumer, which in this instance are government agencies with security-oriented missions.[35] Companies may not seek out terrorism scholar input absent requirements from their clients that values their input. For government agencies, the traditional success metric for assessing many machine learning tools is whether they lead arrests or disrupt plots; information that is easily quantifiable. But few tools are developed to address the non-kinetic aspects of counterterrorism. As an example, Deputy Administrator Cogswell noted that the economic drive to develop machine-learning tools for government security agencies is significantly stronger than the economic incentives to provide CT tools for civil society groups that are best equipped to intervene in the radicalization process of an individual, which would in the long-term have a more meaningful impact for combatting domestic extremism.[36]

The paramount role that machine-learning plays in the counterterrorism space requires that terrorism scholars think through with greater alacrity whether machine-learning platforms benefit the CT mission and if so, what are the best ways to employ ML for the benefit of counterterrorism. From a conversational perspective, this begins with noting that even though the terms AI and ML are used interchangeably, machine learning is a subdiscipline of artificial intelligence that focuses on how non-sentient computerized platforms mimic human decision-making to detect patterns in data germane to humans.[37] In that sense, defining these concepts and history matters just as much as defining terrorism when it comes to developing intelligent machines for counterterrorism, as it elucidates machine-learning's limitations. Moreover, knowing *how* machine learning works will help scholars and policymakers think through the issues of counterterrorism in an age of algorithms and what are better ways to use them.

Statistics, Artificial Intelligence and Machine Learning, and the Social Sciences

The name "machine learning" is intuitive enough that most people encountering the term can deduce what it entails without much explanation. But to explore why the current state-of-the-art of ML does not herald an age of intelligent machines that can replace humans in the CT space, some background on its origins is necessary. The following summary is general and will read as redundant to those familiar with quantitative research methods but will help non-specialists understand the assumptions that sustain all ML platforms. This section is in no way indicative of the rigor surrounding ML nor is it intended to be a literature review describing how ML can advance the field of political violence.

Machine learning owes its intellectual pedigree to statistics and computer science. Statistics, as a discipline, emerged alongside the nation-state, as governments sought information

to better manage their territories.[38] Statistics uses data to generate estimates and inferences about the world in an unbiased way and from those estimates, "infer associations among variables, estimate beliefs or probabilities of past and future events, as well as update those probabilities in light of new evidence or new measurements.[39]" Through causal inference, statisticians go a step further to look at probabilities and estimates in the context of changing conditions, such as the effect of treatment variables on a phenomenon, making it possible to forecast or predict outcomes,[40] which is the information ML algorithms use to make decisions. These tools have allowed political violence scholars to answer questions such as the effect of paying off rebels rather than mounting a full-scale counterinsurgency campaign or the effect of cellphones on the ability of terrorists to execute plots, which have tactical, strategic, and ultimately, policy implications.[41] When used properly, statistics mitigates human biases in analysis and reveals the world as it is and helps decision-makers have more realistic assessments about possible outcomes for different policy choices.[42]

Computer science, for its part, developed as a means of using electronics to automatically perform a set of logical operations and develop algorithms to solve both theoretical and practical computational challenges like classification of data or executing large-volume statistical analyses.[43] With the advent of computers in the 20th century, researchers hypothesized that computers could become autonomous agents that absorb information from the environment to make autonomous decisions without human intervention, or simply put, develop true artificial intelligence.[44] Computer science's origins in statistics spurred much of the initial discussions on artificial intelligence and how machines learn, with researchers exploring how computers could infer information from patterns in data using statistical techniques, and then using these methods to make decisions, classify objects, or make forecasts.[45] Machine learning, in sum, transformed

statistics into an industrial endeavor that could look at data and learn information relevant to humans.

Initially, the scarcity of high-quality data blunted ML's wide adoption outside of specialized computer science labs, as algorithms thrive on the information used to train them.[46] At the dawn of the millennium, the world experienced an information revolution through the concomitant democratization of data through the internet, reduction in costs associated with computer processing, and academia's penchant for collaboration,[47] opening big data to virtually any human domain. Along this revolution came the development of powerful free software like Python and R that came prepackaged with statistics library, making the barrier to entry negligible, paving the way for big data companies to offer statistical insights at a much lower cost to people in industry and government.[48]

At a high level, ML algorithms fall under three buckets: supervised learning, unsupervised learning, and semi-supervised learning.[49] In supervised learning, researchers take labeled data for input x and outcome y and design an algorithm that learns how to map the relationship between x and y so that in a real-world scenario it can reliably predict the y variable if given x.[50] These algorithms are useful for generating estimates on prices for houses or classifying images. Unsupervised learning operates in places where researchers only have inputs and they task algorithms with detecting the inherent relationships within data to cluster groups based on innate or non-obvious patterns.[51] This is useful for finding missing trend lines that humans lack the mental capacity to comprehend. Lastly, semi-supervised learning algorithms come into play when researchers have many inputs and only some labels for y, and researchers must train an algorithm to detect for y with limited data.[52] Semi-supervised learning helps in

generating additional y labels for classification problems like classifying websites as extremist or not extremist.

What ties all three approaches together is that an algorithm depends on access to data, and this starts elucidating ML's current limitations for combating terrorism. Data questions always bring with them the specters that haunt all research enterprises: selection and sampling biases. Yet, to address these concerns with fidelity, those engaged in a particular ML task need to understand the very nature of the problem they are trying to address to make sure that data available are adequate, which starts getting at the *how* machine learning accomplishes what it does. The obstacle though is that many of those designing ML algorithms for combating terrorism and extremism have backgrounds in fields not tied directly to the study of terrorism or the social sciences, such as computer science, statistics, and the natural sciences,[53] and may not know *how* to create an algorithm for real world settings. The latter point was highlighted by Shane Quinlan, the Chief Product Officer for the data science firm Certilytics.[54] In an interview for this paper, Mr. Quinlan explained that some tech firms assume that an arbitrary large number of observations in a laboratory setting will predict behaviors or outcomes in a real-world setting, but never contextualize those observations sufficiently to account for real-world variation.[55] As an example, he described how in the medical field, where records are more standardized than the usual terrorist dataset, firms train models with upwards of 20,000 observations and still struggle to account for unforeseen variations across tens of billions of events that occur in production settings. This inability to comprehend variation means that many algorithms are ill-suited for detecting the type of black swan events that characterize terrorism.[56] The only way to build an algorithm that has both internal and external validity is to understand a problem, which is the ultimate *how* of a successful machine-learning platform.

Mr. Quinlan noted that the answer to the *how* comes in the form of research design.[57] Research design refers to the investigatory strategy used to answer research questions, and social science scholars over the years have outlined steps for sound research projects.[58] King, Keohane, and Verba note, in the absence of proper research design, many interesting research projects or experiments in the social sciences lose predictive power or generalizability because they ask the wrong questions,[59] or because they do not take into account the data generating process, how this affects whether the employed methods truly answer the question being asked, and ultimately what data communicates.[60] This form of training is often lacking even among quantitative social science researchers more accustomed to deriving inferences from data as a given.[61] The most famous example of improper research design in security studies is Abraham Wald's bomber study during World War II.[62] After analyzing damage received by airplanes that survived combat, the US military determined that to increase the bomber survival rate, its engineers needed to place more armor on those areas of an airplane that showed the most damage. Wald, in contrast, discerned that the data only showed planes that survived even after receiving massive damage. What was missing in the data were the places on an aircraft where planes that did not survive were hit. Based on the data generation process, Wald determined it was better to place more armor on those places where bombers showed little damage, as it meant that those were the areas that if hit would destroy an aircraft.[63]

These considerations weigh more for domestic counterterrorism. Terrorism is a heterogenous concept that has no defined battlefield, thrives on secrecy and surprise, and changes in meaning. Consider the collective understanding of terrorism at the beginning of the millennium, with al-Qaeda and its desire to commit mass casualty attacks as representative of the threat, to that of the early 2020s, where security officials must consider both al-Qaeda and lone-

wolves affiliated to a myriad of ideologies but not necessarily attached to any organization. In addition, modern extremists, as demonstrated by the January 6 insurrection, belie expectations of what makes a person an "extremist," with many of those participating coming from well off backgrounds and no previous connection to extremist groups.[64] This makes the challenge of separating out potential terrorists much harder, as often, the telltale signs of an extremist are indistinguishable from highly motivated political citizens expressing their first amendment rights.[65]

The attendant question once the subject is defined is whether there are available data that captures a phenomenon under study. Without deviating too much in epistemology, data are records of events as understood by humans,[66] which means that researchers must go beyond understanding humans but also the type of behaviors associated with humans engaged in a particular activity like terrorism and how data defines and captures them. Even in data rich environments like the counter-IED efforts in Iraq and Afghanistan proved insufficient for allowing data scientist to properly answer the question, as the data could not capture what made IEDs successful.[67] Without this grounding in research design, scholars can perform rigorous statistical analyses on data and stumble upon spurious correlations that are statistically valid but analytically meaningless or wrong in practical terms.[68]

Where an algorithm shows its mettle is in its accuracy. One of the key considerations as ML-algorithms move from university settings into critical national security programs is their ability to successfully identify terrorist in environments where there is less tolerance for mistakes. Two critical measures of an algorithm's efficacy are the precision and recall scores.[69] Both are closely related measures of whether an algorithm made the right prediction. Rarely does a model map with 100% accuracy the relationship between variables X and Y, creating the

distinct possibility of producing false-positives (variables mistakenly categorized as Y) and false-negatives (variables mistakenly classified as not-Y).[70] Depending on what an algorithm is tasked with doing, a low recall or precision score could make the difference between life and death. If an algorithm is designed to indicate flags for risk within large datasets, relatively low precision and recall rates are tolerable as they at least help analysts prioritizes entities to investigate. But if an algorithm guides a drone with the power to fire at will, even moderately high precision and recall scores might lead to the loss of innocent lives. The key to improving precision and recall are data. The better the data, whether in quantity or quality, the better these scores will be.

These points all help ground the core of the challenges behind the *how* when it comes to using ML-algorithms for domestic CT. Yet to get at the *how* does not mean that only those with quantitative methods training are adequately skilled to think critically about ML. For policymakers and terrorism scholars alike, what matters is not necessarily whether they are sufficiently versed in statistics and machine learning to design their own classifiers, but rather if they are intelligent consumers of data.[71] This goes beyond envisioning what relationship is an algorithm supposed to capture, and instead making sure that researchers and developers are asking the right questions, offering valid hypotheses between relationships, asking about the data generating process, and ultimately, are the data appropriate and sufficient for any algorithm.

The Machine Learning Status Quo

Instances of machine learning operating in suboptimal ways are legion, from an Amazon algorithm designed to make gender hiring more equitable being prejudiced to bring in more males to automated soap dispensers that struggle in identifying people with darker skin.[72] These

examples are products of improper research design that did not consider data biases. In the past, the weak interaction between the social sciences and the policy community limited the effects of improper research design or data science.[73] This is no longer the case, as the field of political violence has come into vogue with the desire to answer critical questions with policy implications and tech companies promising to bring sophisticated machine-learning techniques to the forefront. The use of quantitative techniques without properly understanding what data communicates has already resulted in poor outcomes for counterterrorism.

A prime example is Robert Pape's study that found a correlation between occupation and suicide terrorism.[74] Pape framed this paper as offering policy prescriptions derived entirely from empirical evidence.[75] The emphasis on statistical robustness and the parsimony of the theory, that suicide terrorism was caused by wars of national liberation, proved attractive to those wanting to make sense of the threat posed by al-Qaeda in the aftermath of 9/11 and later during the Iraq War. A cursory look throughout the historical literature would have shown the obvious spuriousness of the posited hypothesis,[76] but these weren't leading to research grants from the Department of Defense.[77] Later, social scientists with an eye to the actual nature of terrorism and with a background in research design identified the key statistical flaw within Pape's paper that rendered its findings unfounded: Pape tested on the dependent variable, failing to take into account all instances of occupation that did not lead to suicide terrorism.[78] Pape published his findings on the eve of the big data revolution and did not use a modern ML algorithm for his study. He also followed political science's open access standards, publishing his data and methods online, facilitating assessments of his paper that helped correct its methodological flaws.

Examples like Pape, where the academic community managed to address flaws in a study, will be uncommon in industry. As noted previously, this is combination of both the black box nature of many sophisticated ML-algorithms where the underlining relationship mapping is inscrutable to humans and the premium private firms place on protecting intellectual property and data.[79] In other instances, when technology firms perform work in service of government, it is possible that the training data will be classified and controlled entirely by the government, raising questions as to what the data may or may not contain. Together, this reduces the ability for the public to scrutinize any deployed ML-algorithm.

Returning to the subject of this essay, using ML-algorithms for domestic CT, it is worth considering what is in the realm of the possible. The nature of terrorism makes training an algorithm to predict and prevent a terrorist attack unlikely. Terrorism, despite its unfortunate frequency, is still a relatively rare event, bringing up data issues, as the activity bucketed under the label covers the gamut from mass shootings to airliner attacks, which further complicates data questions.[80] Some scholars have worked by analogy, looking at applied uses of machine-learning for policing or for suicide prevention and repurposing these tools to predict terrorist incidents.[81] There are limits though. Crimes like car-jackings occur within confined geographic regions, with the targets not necessarily evolving, which creates a rich data source lacking in noise.[82] In contrast, terrorism's unpredictably and variance makes it difficult to create datasets capable of predicting actual terrorist incidents.[83]

When breaking apart CT into component parts, there has been more success. The status quo for machine-learning and domestic CT, however, has centered on detecting and disrupting terrorist plots by examining social media content, using machine vision, and detecting terrorist financing through money transfers.[84] Intuitively, these seem like valid approaches, as these were

valuable data streams for combat operations abroad, and domestically, these data are readily available. There are thousands of online forums for extremists, radicals, and would-be terrorist, and not all of them are secure messaging platforms, offering ample opportunity for data collection and analysis in data rich environments. Many of these platforms contain images of extremists, which when married to text, can provide insight into potential terrorists. But, as Mr. Quinlan noted, these uses at best represent marginal updates to the profiling and predictive analytical techniques employed by law enforcement for decades except they come with data issues many algorithms struggle in overcoming.[85]

To make sense of this argument, it helps to tease apart the nature of these data, starting with social media and the government's aspirations for machine-learning tools on these platforms. If the government is looking to social media to conclusively determine who is a terrorist, this runs the risk of repeating Pape's mistake of testing on a dependent variable owing to the available data by either modeling only known terrorists and their online profiles or prioritizing only websites where extremists congregate. If modelers use the profiles of the January 6 insurrectionists, data imbalance issues will arise that will limit an algorithm's predictive power,[86] as those who stormed the capitol are a fraction of those who discussed the insurrection online. This would create other biases as well. The far-right, the form of extremism that is currently dominating policy-discussions, is not a monolith, and historically has been notorious for internecine fighting and disputes.[87] Focusing solely on the January 6 insurrectionists would not capture the biographical variation of the other major far-right event of the last five years, the 2017 Charlottesville rally. January 6 attracted many former law enforcement and military personnel while Charlottesville involved more extreme neo-Nazi groups.[88]

Another issue with social media is the amount of ephemeral noise produced by anonymous figures. The United States does not criminalize most forms of speech unless there is direct incitement to violence.[89] An improperly calibrated algorithm runs the risk of mistakenly ensnaring many highly politicized individuals that at no point intended to commit violence and would likely violate civil liberties and the individual right to privacy, as well as opening up the government agency to criticism for targeting political opponents. The field of ML that addresses rhetoric and language, natural language processing (NLP), still struggles with surfacing intent and context, and generates many false-positives.[90] Nor does NLP offer answers for de-anonymizing individuals, especially when they use aliases and likely obfuscate their technical signatures through VPNs, secure browsers, and other operational security measures.[91] NLP at best gives flags on where to focus analysis, but it does not tell you much about the human behind the screen and whether their other behaviors might be indicative of them being a risk.

But most importantly, as Brian Drake explained, the government has restricted authorities to look at social media, which limits data availability and therefore the predictive power of most algorithms.[92] What little data government might have is complicated further by the heterogeneity of these data. The revealed behaviors for individuals differ from platform to platform owing to their contrasting functionality like whether users must use real names, if users can share videos, and the length of the messages users can share.[93] Drake noted from a CT perspective, a more realistic approach is for the government to bolster the status quo where the social media companies detect extremist content themselves using ML-tools to deplatform extremists.[94]

Until now, this process has been self-regulated, with companies outlining their own terms-of-services and their enforcement mechanisms.[95] Facebook, Twitter, and others have automated algorithms that screen for the most obvious forms of extremist propaganda,[96] although

these can be gamed through the generation of new rhetoric to avoid censorship.[97] This approach succeeded in deplatforming Islamic State and other Islamic groups from most mainstream platforms, limiting their ability to spread propaganda and to recruit,[98] and this model seems most promising in the American context. Jigsaw and various civil society organizations are also pioneering counter-messaging work when individuals search for extremist content online and retooling their algorithms to hide propaganda videos.[99]

There is an inherent tension with depending on social media platforms to police themselves. Social media succeeds because it provides meaningful connections and interactions that provide dopamine to the mind.[100] The entire business model behind social media companies is to show people more of what they want to see, and this greatly drives the propagation of extremist content on Facebook, Twitter, and other platforms.[101] To prevent this from happening, these companies would have to completely upend the algorithms and business offerings that made them successful in the first place.[102] This creates a cynical argument that the incentive behind sponsoring counter-radicalization policies and research by these technology companies is to give the veneer of action without necessarily addressing the radicalizing nature of their algorithms and to prevent government from trying to regulate them.[103] Moral suasion by the government and the public could force social media platforms to strengthen their terms-of-service and enforce more stringently their internal policies, but the United States could support these companies further by sponsoring R&D projects that center on how to improve their ML algorithms for combating extremism.

The challenges posed by social media exist in almost equal measure when it comes to facial recognition software or other aspects of machine vision such as augmented reality. In early 2020, Clearview AI drew scrutiny when the New York Times profiled the company and noted

how it scraped millions of images from social media platforms over several years and now sells data to the government and commercial entities.[104] Clearview scraped these images without user consent from Facebook, Twitter, and others,[105] and will likely find access to these websites more difficult, limiting data availability.[106] From the government's perspective, there are legal questions as to whether various agencies are authorized to use the tool.[107] Further, Clearview AI's data should not be indicative of the entire universe of images in the world. Something that is true both about social media content and facial recognition software is that they only capture the universe of individuals that post their photographs online and speak little of analog content. This might work in terms of detecting amateurs that broadcast an insurrection on Facebook and later brag about their participation on dating websites, but truly professional terrorists are likely to operate more securely and seek to limit their online technical signature to not be detected by these tools.

Machine-vision, the subfield of ML that detects patterns in images, is also easy to compromise and is laden with potential for data biases. Hackers in the past have fooled cameras on Tesla with simple stickers to mistakenly switch lanes or speed up at stop signs.[108] Researchers have also noted that the current state-of-the-art in machine-vision struggles in successfully classifying young people, black people, and women, creating a large number of false positives.[109] The National Institute of Standards and Technology (NIST) analyzed several facial-recognition algorithms and found that many have high false-positive rates when it comes to identifying people from East and West Africa and East Asia.[110] Focusing solely on domestic images, these algorithms maintained high levels of false-positive rates for African Americans and Asians, but had their highest false-positive rates when looking at American Indians.[111] NIST noted the issue of selection bias, as it found algorithms designed in China to have a reduced false-positive rate

when it came to identifying individuals from East Asia.[112] The government itself has a wide-swath of mugshots it could use for training an algorithm, but black people are overly represented in these data.[113] Moreover, if the majority of the data available derives from American citizens, facial recognition software could miss individuals traveling from abroad to the United States to commit acts of terrorism.

These concerns should weigh heavily when thinking through near-future technologies like augmented reality or simulations. With the advent of smartphones and headsets, U.S. security personnel and law enforcement will have access to new sensors that can absorb information from the environment and provide near instantaneous updates on individuals or objects these devices detect. These technologies promise real-time translation capabilities, instantaneous crime updates, and immediate updates on biometric sensors.[114] But their precision and recall rate will depend entirely on what corpus of data companies use for training purposes, and the information returned may not prove dispositive for the end users of these tools. This matters, as Deputy Administrator Cogswell noted that in the American context, much of the counter-CT and counter-radicalization work is done by local law enforcement.[115] While DOJ, technology vendors, and other entities provide training on various tools, if end-users fail to understand what is an acceptable false positive / false negative rate, or if an entity does not set a policy for what an appropriate action is to take based on a match, there is potential for harm to the domestic public.[116] For their part, simulations used for training or predictive analytics will depend on the quality and quantity of observations. Returning to Mr. Quinlan's example that models built around 20,000 high quality observations struggle in predicting medical outcomes, machine-vision simulations may not provide real predictive value when domestic terrorist data are relatively scarce.

Machine-vision does have its uses. The state of Maryland used facial-recognition software to identify the Capital-Gazette shooter in 2018,[117] and according to publicly available sources, the FBI has had some success with identifying the January 6 insurrectionists.[118] But machine-vision still needs better calibration with the right data, and given the range of false-positives, such programs will likely require extensive human-in-the-loop supervision to assure an algorithm does not identify the wrong person. This is the case for any domestic ML tool, as there is little tolerance for mistakes.

Perhaps the status-quo area where ML has shown its mettle is in finance. For the past decade, ML algorithms have been used extensively to detect risky actors transacting through banks.[119] Such systems are well-equipped to generate flags about money-launderers and drug-traffickers seeking to raise funds for terrorist financing. These data work by looking anomalies in transaction patterns and matching them with know-your-customer (KYC) requirements, giving perspective on who might be using the financial system, why might they be transacting, and if a transaction might be something more nefarious that requires the attention of the government.[120] Regulatory pressure and the contained nature of the financial system makes this possible, and permit for greater insights than more open-ended systems allow. For instance, when screening individuals, an algorithm can map-out networks of known associates and identify connections that might indicate risk.[121] It might be the case where an individual applying for a bank loan might not have any evident ties to extremism, but a family member might have connections to the Oath Keepers or to Hezbollah, and an ML algorithm can generate a flag that draws an analyst's attention. The reason this is possible is due to the vast trove of homogenous data financial institutions hold and government pressure, which compels banks to cooperate with the government in transmitting data.[122] These systems, more importantly, give an escape valve for

false-positives by focusing on generating a comprehensive number of flags that the government must then review that gives small room for evasion to illicit actors.[123]

Not everything is perfect with the financial industry, however. Even with regulatory pressure, financial institutions and banks tend to use a rules-based approach to terrorism financing that, much like social media norms, is easily evaded.[124] But banks are aware of this and are increasingly adopting ML platforms that are better able to detect anomalies in data.[125] These successes speak more to the clarity in expectations government has about ML in finance, and the fact that banks as tightly regulated entities must share data with government. Compare this with social media companies where firms are directing policy innovation themselves and government must be responsive or else engage in costly lawsuits.

What are realistic uses for ML and CT?

The tone of this essay until now has been skeptical about the ongoing uses of machine learning for domestic counterterrorism. The examples cited represent the status quo where tools developed in the context of the global war on terror are now being suggested to be cross-applied to domestic use-cases where data scarcity blunts their usefulness, where the risk of inappropriate use present risk of harm to the public, or where the government lacks authorities to use the technology. What is critical is that these tools were developed without direct inputs from terrorism scholars that could speak on *how* to better develop ML tools for the domestic CT mission within the frameworks of democratic societies with strong civil liberties. The nature of modern CT means this is an important research area for the terrorism studies literature writ large. By engaging in it, the scholarship will transition from that of a bystander, where it observes the development of ML-tools to critique them or to think through their applications for the CT-

enterprise to one where the scholarship is participatory, playing the paramount role of bringing the vanguard of research to the development of ML-platforms that combat extremism and terrorism.

This paper has already explained that ML-algorithms are tools that excel at overcoming human bias and augmenting a person's cognitive abilities by revealing the world more rapidly through data. But whether an algorithm does harm or if it succeeds depends both on the training data and the scoping of an algorithm to a problem. This more nuanced assessment transforms machine-learning from a nebulous panacea for CT to that of a tool whose success is driven by how its crafted and how it is used.

The policymakers and developers interviewed for this paper offered perspective about how this could look. Brian Drake, who oversaw numerous ML projects for DIA, gave a very practical answer. While explaining that DIA does not operate domestically, Mr. Drake noted that his agency leverages ML to understand information warfare campaigns being waged by hostile actors.[126] Because ML flattens information to reveal non-obvious patterns and limit human biases, DIA deploys various means to detect coordinated messaging campaigns that spread misinformation and various conspiracies.[127] He noted that one of the most troubling things that the United States' nation-state rivals do is sow discontent and attack the sinews that make American society strong.[128] From a counterterrorism perspective, this plays to ML's strengths, detecting anomalies at scale that humans would struggle in piecing together, finding recurrent patterns, and providing analysts various indicators of risk. It is worth noting that this adds another layer to the domestic extremism challenge. If individuals radicalize into terrorism by consuming propaganda broadcast by adversaries, this blurs the distinction between domestic and foreign terrorism. Given the ongoing debate about the need for a domestic terrorism statute,[129]

Congress might need to change authorities for FBI and DHS to better combat these terrorist incidences facilitated by information warfare.

On a domestic level, as noted previously, similar work is being carried out by technology companies, such as Jigsaw and Moonshot CVE providing counter-messaging.[130] These tools, however, presume that radicalization occurs solely via the internet, and do nothing to address other sources of radicalization, such as ties of kith and kin.[131] Deputy Administrator Cogswell suggested that some of the productive potential uses for the results from a ML tool in a domestic counterterrorism context could be in providing information to support to non-profits and civil society entities, to include religious organizations, that are better able to interact with individuals and communities at higher risk, in order to intervene in the radicalization process.[132] Research into this type of work is limited because there is no holistic effort to capture data on extremist and radicalization activity happening in the off-line environment. Mr. Quinlan explained that this should be regarded as opportunity.[133] When data are lacking, organizations can get creative in finding less obvious data that captures non-internet behavior in a non-intrusive manner, which with the help of a machine learning company, could then build tools for detecting the incidence of terrorist radicalization in a manner that protects civil liberties.[134]

Mr. Quinlan noted that ML serves an important role for empowering national security professionals.[135] For instance, even though for many people the internet has become synonymous with social media activity, the world wide web contains vast amounts of publicly available information, from adverse media and other websites with germane information. National security has more authorities to look at this data and if collated, can be used to screen-and-vet individuals seeking to join the military, police, or others in sensitive positions of trust for extremist or violent prior activity. Currently, the process for screening and vetting potential recruits is by doing

labor-intensive manual google searches.[136] Even though terrorists adopt strategies that magnify their capabilities by exploiting the media, [137]the percentage of Americans that might be terrorists is small. The government cares more about directing its limited resources on those that pose a threat to the homeland. A machine learning algorithm that can look at entire populations, like enlistees, in bulk to identify those individuals whose behavior online best matches that of an extremist would greatly help the government surface threats on the front-end rather than downstream when an extremist has joined the military.

ML can also detect extremist websites early on to give law enforcement clues about where individuals might be gathering online.[138] The internet is a big messy place, with new websites sprouting daily. Staying abreast of every page, blog, and platform is laborious and intensive. The government could easily deploy a scraper that monitors the emergence of new websites that share propaganda and identify whether these sites have a foreign provenance. This could serve the dual purpose of keeping a pulse of emerging threats and also understanding how extremist propaganda and rhetoric evolves, which it can then share with social media platforms, fostering that necessary public-private cooperation.

An area that remains underexplored in a tactical scenario is the use of unmanned vehicles and robots powered by ML to gather sensory data for tactical engagements. Law enforcement in the United States broke the taboo of using robots to eliminate targets in Dallas in 2016 when the local police department used a robot to deliver a bomb against a mass shooter.[139] In calibrated high-risk situations, such as those involving hostage-taking or bomb threats, drones and other unmanned vehicles might serve as sensors for data collection and augment the government's knowledge about a target without risking personnel and obfuscating its presence. But the emphasis in this example is calibrated and focused on information gathering rather than giving

autonomous vehicles decision-making power. As noted previously, the recall and precision rates associated with machine-vision is not high enough to instill confidence in these vehicles taking the optimal decision in these scenarios. Nonetheless, their ability to parse data in scenarios where humans would be overwhelmed by emotions such as fear or anxiety would provide a more accurate perspective in these tense moments.

All the described settings are hypothetical because they are limited by what data are available now. But as these previous examples show, ML operates best when the goal is centered on automating the difficult and the labor-intensive work that produces mental fatigue so humans can focus on the higher order tasks of inference and decision-making. Certainly, the previous sentence can be rendered untrue in coming years. For this to happen, researchers and government need to produce better data for the full spectrum of the CT enterprise. The important caveat is that current uses of machine-learning for counterterrorism must be tempered to not risk societal backlash that could stigmatize development and use.

Counterterrorism in the Age of Intelligent Machines

There is a universe where ML programs for CT become a force of oppression. This is the model postulated by China where it uses the pretext of counterterrorism to control its domestic Uyghur population and commit genocide,[140] and to become the very manifestation of big brother. It controls access to the internet, uses facial recognition software against its entire population, assigns reputational scores to each individual based on their public postings on social media, and controls the movement of its population.[141] This model evinces no sympathy for civil liberties and the democratic spirit.

In contrast, the rationale for American CT has been to safeguard democratic liberties and secure public liberty. This means using CT to create a crucible so that Americans can freely assert their rights guaranteed by the constitution and American jurisprudence. The balance between security and liberty is perennially in tension, and machine-learning can either ease or exacerbate this condition. Policymakers can ameliorate this challenge through data literacy and becoming intelligent consumers of data to understand what algorithms actually do. Algorithms are wonderful things that can recommend books on Amazon, but they can mistakenly identify people based on racial or ethnic biases. When policymakers consider using ML domestically or when terrorism scholars think about the future of counterterrorism using ML, they need to think beyond the morality of using these tools and instead focus on *how.* As discussed in this essay, that *how* does not mean learning how to code or how to perform statistics. The *how* centers on knowing what questions to ask when exploring these technologies: what exactly is the tool supposed to accomplish, what is the research framework that gives guardrails and validity to an approach, what is the data generating process, and what do data ultimately communicate. It means going a step further as well and thinking of more creative uses of ML for CT that do not correspond to a warfighting posture. There is also an obligation to educate decision-makers and the public about what machine-learning outputs signify and mean in real world examples.

There is no better moment than the present for policymakers and terrorism scholars to delve into these issues. In the long-term, well used ML platforms can encourage public confidence in these tools, as they reduce human biases in counterterrorism, and prevent potential terrorist attacks. This is not solely a task for the United States government. It must leverage partnerships with the technology industry, with finance, and with civil society groups that understand data and can provide clear insight into ML's benefits for CT.

[1] Damien Van Puyvelde, Stephen Culthart, M. Shahriar Hossain, "Beyond the buzzword: big data and national security decision-making," *International Affairs* 93, no. 6 (2017), 1398-1399
[2] Allegra Harpootlian and Emily Manna, "The new Face of American War Is a Robot," The Nation, April 29, 2019, https://www.thenation.com/article/archive/tom-dispatch-american-warfare-drones-military-tech-robot/
[3] Homeland Security and Public Safety Division, NGA Center for Best Practices, National Governors Association, "Artificial Intelligence (AI) in Homeland Security and Emergency Management," National Geospatial-Agency, September 2018, https://www.nga.org/wp-content/uploads/2019/08/AI-in-Homeland-Security-and-Emergency-Management-Memo.pdf
[4] Jon Harper, "VSOFIC News: SOCOM All-in on Artificial Intelligence," National Defense Magazine, May 12, 2020, https://www.nationaldefensemagazine.org/articles/2020/5/12/socom-all-in-on-artificial-intelligence
[5] Katja Grace, John Salvatier, Allan Dafoe, Baobao Zhang, Owain Evans, "Viewpoint: When Will AI Exceed Human Performance? Evidence from AI Experts," *Journal of Artificial Intelligence Research* 62 (2018)
[6] Kevin P. Murphy, *Machine Learning: A Probabilistic Perspective* (Cambridge, MA: The MIT Press, 2012), 1
[7] Guilong Yan, "The impact of Artificial Intelligence on hybrid warfare," *Small Wars & Insurgencies* 31, no.4 (2020), 905
[8] James Giordano, "Neurotechnology in National Security and Defense: Practical Considerations, Neuroethical Concerns," in *Neurotechnology in National Security and Defense: Practical Considerations, Neuroethical Concerns* ed. James Giordano (Boca Raton, FL: CRC Press, 2015), 1-10
[9] Brian Drake in discussion with the author, July 2021
[10] Jeff Goodson, "Irregular Warfare in a New Era of Great-Power Competition," Modern War Institute at West Point, May 20, 2020, https://mwi.usma.edu/irregular-warfare-new-era-great-power-competition/
[11] Eric Schmitt and Helene Cooper, "How the U.S. Plans to Fight From Afar After Troops Exit Afghanistan," The New York Times, April 15, 2021, https://www.nytimes.com/2021/04/15/us/politics/united-states-al-qaeda-afghanistan.html
[12] Ken Dilanian and Julia Ainsley, "DHS weighing major changes to fight domestic violent extremism, say officials," NBC News, March 25, 2021, https://www.nbcnews.com/politics/national-security/dhs-weighing-huge-changes-fight-domestic-violent-extremism-say-officials-n1262047
[13] Ibid.
[14] Ibid.
[15] Rohini Kurup and Benjamin Wittes, "Was Jan. 6 an Intelligence Failure, a Police Failure or both?" *Lawfare,* March 1, 2021, https://www.lawfareblog.com/was-jan-6-intelligence-failure-police-failure-or-both
[16] Aaron Boyd, "An Inside Look at All the Data CBP Collects About Everyone Crossing U.S. Borders, "Nextgov, September 18, 2019, https://www.nextgov.com/emerging-tech/2019/09/inside-look-all-data-cbp-collects-about-everyone-crossing-us-borders/159946/; Chris Baraniuk, "Machine Minds: The new weapon in the fight against crime," BBC, March 3, 2019, https://www.bbc.com/future/article/20190228-how-ai-is-helping-to-fight-crime
[17] Gary Shiffman, *The Economics of Violence* (New York: Cambridge University Press, 2020)
[18] Kathleen Belew, *Bring the War Home: The Power Movement and Paramilitary America* (Cambridge, MA: University Press, 2018)
[19] Victoria L. Killion, "Terrorism, Violent Extremism, and the Internet: Free Speech Considerations," Congressional Research Service, May 6, 2019, https://fas.org/sgp/crs/terror/R45713.pdf
[20] Zachary Leibowitz, "Terror on Your Timeline: Criminalizing Terrorist Incitement on Social Media Through Doctrinal Shift," *Fordham Law Review* 86, no. 2 (2017), 802-804
[21] Laura K. Donohue, "Anglo-American Privacy and Surveillance," *The Journal of Criminal Law & Criminology* 96, no.3 (2006), 1064-1136; Richard A. Posner, "Privacy, Surveillance, and Law," *The University of Chicago Law Review*
[22] Richard Abanes, *American Militias: Rebellion, Racism, & Religion* (Downers Grove, Illinois: InterVarsity Press, 1996),45-49
[23] Bruce Hoffman, *Inside Terrorism, 3rd Edition* (New York: Columbia University Press, 2017,), 32-44
[24] Alon Halevy, Peter Norvig, and Fernando Pereira, "The unreasonable effectiveness of data," IEEE Intelligent Systems, 24, no. 2 (2009), 8-12
[25] Paul R. Milgrom and Steven Tadelis, "How Artificial Intelligence and Machine Learning Can Impact Market Design," NBER Working Paper 24282 (2018), https://www.nber.org/system/files/working_papers/w24282/w24282.pdf; Abishek Gupta, Alagan Anpalagan, Ling

Guan, Ahmed Shaharyar Khawaja, "Deep learning for object detection and scene perception in self-driving cars: Survey, challenges, and open issues," *Array* 10 (2021)
[26] Kathleen McKendrick, *Artificial Intelligence Prediction and Counterterrorism* (London, UK: Chatam House, The Royal Institute of International Affairs, 2019), https://www.chathamhouse.org/sites/default/files/2019-08-07-AICounterterrorism.pdf
[27] Boaz Ganor, "Artificial or Human: A New Era of Counterterrorism Intelligence," *Studies in Conflict & Terrorism* 41, 2019, https://www-tandfonline-com.proxy.library.georgetown.edu/doi/full/10.1080/1057610X.2019.1568815
[28] Sam Biddle, "Police Surveilled George Floyd Protests With Help From Twitter-Affiliated Startup Dataminr," The Intercept, July 9, 2020, https://theintercept.com/2020/07/09/twitter-dataminr-police-spy-surveillance-black-lives-matter-protests/
[29] Timothy B. Lee, "Why Google Believes machine learning is its future," Ars Technica, May 10, 2019, https://arstechnica.com/gadgets/2019/05/googles-machine-learning-strategy-hardware-software-and-lots-of-data/
[30] Casey Ross and Ike Swetlitz, "IBM's Watson supercomputer recommended 'unsafe and incorrect' cancer treatments, internal documents show," Stat News, July 25, 2018, https://www.statnews.com/2018/07/25/ibm-watson-recommended-unsafe-incorrect-treatments/
[31] Benjamin Jensen, Scott Cuomo, and Chris Whyte, "Wargaming with Athena: How to Make Militaries Smarter, Faster, and More Efficient with Artificial Intelligence," War On the Rocks, June 5, 2018, https://warontherocks.com/2018/06/wargaming-with-athena-how-to-make-militaries-smarter-faster-and-more-efficient-with-artificial-intelligence/
[32] Ganor, "Artificial or Human,"
[33] McKendrick, *Artificial Intelligence Prediction;* John Harper, "Artificial Intelligence Could Help Neutralize Enemy Bombs," National Defense Magazine, September 18, 2017, https://www.nationaldefensemagazine.org/articles/2017/9/18/artificial-intelligence-could-help-neutralize-enemy-bombs
[34] Pantelis Linardatos, Vasilis Papastefanopoulos, and Sotiris Kotsiantis, "Explianable AI: A Review of Machine Learning Interpretability Methods," *Entropy* 23, no. 18 (2020)
[35] Patricia Cogswell in discussion with the author, July 2021
[36] Ibid.
[37] Eda Kavlakoglu, "AI vs. Machine Learning vs. Deep Learning vs. Neural Networks; What's the Difference," IBM, May 27, 2020, https://www.ibm.com/cloud/blog/ai-vs-machine-learning-vs-deep-learning-vs-neural-networks
[38] Martin van Creveld, *The Rise and Decline of the State* (Cambridge, UK: Cambridge University Press, 1999), 145-147
[39] Judea Pearl, "Causal inference in statistics: An overview," *Statistics Survey* 3 (2009), 99
[40] Ibid.
[41] Eli Berman, Jacob N. Shapiro, and Joseph H. Felter, "Can Hearts and Minds Be Bought? The Economics of Counterinsurgency in Iraq," *Journal of Political Economy* 119, no.4 (2011); Jacob N. Shapiro and Nils B. Weidmann, "Is the Phone Mightier Than the Sword? Cellphones and Insurgent Violence in Iraq," *International Organization* 69, no. 2 (2015)
[42] Daniel Kahneman, *Thinking Fast and Slow* (New York: Farrar, Straus and Giroux, 2011), 3-12
[43] David Gunning and David W. Aha, "DARPA's Explainable Artificial Intelligence Program," *AI Magazine* 40, no.2 (2019), 45
[44] Manuel DeLanda, *War in the Age of Intelligent Machines* (New York: Zone Books, 1991)
[45] Pat Langley, "The Changing science of machine learning," *Machine Learning* 82 (2011), 275-279
[46] Stuart Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach, 3rd Edition* (Boston, MA: Prentice Hall, 2010), 27
[47] Kosuke Imai, *Quantitative Social Science: An Introduction* (Princeton, NJ: Princeton University Press, 2017), 1-2
[48] Russell and Norvig, *Artificial Intelligence,* 27-28
[49] Murphy, *Machine Learning*, 2-16
[50] Ibid.
[51] Ibid.
[52] Ibid.
[53] Justin Grimmer notes that most of those using big data in industry don't have social scientific training. See Justin Grimmer, "We Are All Social Scientists Now: How Big Data, Machine Learning, and Causal Inference Work Together," *PS: Political Science & Politics* 48, no.1 (2014), 80

[54] Shane Quinlan in discussion with the author, August 2021
[55] Ibid.
[56] Ibid.
[57] Ibid.
[58] Gary King, Robert O. Keohane, and Sidney Verba, *Designing Social Inquiry: Scientific Inference in Qualitative Research* (Princeton, NJ: Princeton University Press, 1994)
[59] Ibid.
[60] Kosuke Imai, Gary King, and Elizabeth Stuart, "Misunderstandings between Experimentalists and Observationalists about Causal Inference," *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 171, no.2 (2008), 481-502
[61] Gary King, Robert O. Keohane, and Sidney Verba, "The Importance of Research Design in Political Science," *The American Political Science Review,* 89, no.2 (1995), 475
[62] Marc Mangel and Francisco J. Samaniego, "Abraham Wald's Work on Aircraft Survivability," *Journal of the American Statistical Association* 89, no. 386 (1984)
[63] Ibid.
[64] Robert A. Paper and Keven Ruby, "The Capitol Rioters Aren't Like Other Extremists," The Atlantic, February 2, 2021, https://www.theatlantic.com/ideas/archive/2021/02/the-capitol-rioters-arent-like-other-extremists/617895/
[65] Jerusalem Demsas, "The online far right is angry, exultant, and ready for more," Vox, January 11, 2021, https://www.vox.com/2021/1/9/22220716/antifa-capitol-storming-far-right-trump-biden-election-stop-the-steal-hawley-cruz
[66] For an in-depth exploration see John Cheney-Lippold, *We Are Data: Algorithms and the Making of Our Digital Selves* (New York: New York University Press, 2017)
[67] Kelsey Atheron, "When big data went to war – and lost," Politico, October 11, 2017, https://www.politico.com/agenda/story/2017/10/11/counter-ied-warfare-data-project-000541/
[68] Herbert A. Simon, "Spurious Correlation: A Causal Interpretation," *Journal of the American Statistical Associations* 49, no. 267 (1954),
[69] D.M.W. Powers, "Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation," *Journal of Machine Learning Technologies* 2, no.1 (2011), 37-39
[70] Blakeley B. McShane, David Gal, Andrew Gelman, Christian Robert, and Jennifer L. Tackett, "Abandon Statistical Significance," *The American Statistician* 73 (2019), 235-238
[71] Mick Ryan, "Intellectual Preparation for Future War: How Artificial Intelligence Will Change Professional Military Education," War on the Rocks, July 3, 2018, https://warontherocks.com/2018/07/intellectual-preparation-for-future-war-how-artificial-intelligence-will-change-professional-military-education/
[72] Jeffrey Dastin, "Amazon scraps secret AI recruiting tool that showed bias against women," Reuters, October 10, 2018, https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G; Sidney Fussell, "Why Can't This Soap Dispenser Identify Dark Skin?" Gizmodo, August 17, 2017, https://gizmodo.com/why-cant-this-soap-dispenser-identify-dark-skin-1797931773
[73] Gary J. Andres and Janice A. Beecher, "Applied Political Science: Bridging the Gap or a Bridge Too Far?" *PS: Political Science and Politics* 22, no. 3 (1989), 636-639
[74] Robert A. Pape, "The Strategic Logic of Suicide Terrorism," *American Political Science Review* 97, no.3 (2003)
[75] Ibid.
[76] Walter Laqueur and Christopher Wall, *The Future of Terrorism* (New York: Thomas Dunne Books, 2018), 11
[77] See Laura Rozen, "Researcher: Suicide terrorism linked to military occupation," Politico, October 11, 2010, https://www.politico.com/blogs/laurarozen/1010/Researcher_Suicide_terrorism_linked_to_military_occupation.html
[78] Scott Ashworth, Joshua D. Clinton, Adam Meirowitz, and Kristopher W. Ramsay, "Design, Inference, and the Strategic Logic of Suicide Terrorism," *American Political Science Review* 49, no. 2 (2008)
[79] David Gunning and David W. Aha, "DARPA's Explainable Artificial Intelligence Program," *AI Magazine* 40, no. 2 (2019)
[80] Jeff Jonas and Jim Harper, *Effective Counterterrorism and the Limited Role of Predictive Data Mining* (Washington, D.C.: Cato, Policy Analysis no. 584, 2006), 6-10
[81] McKendrick, *Artificial Intelligence,* 9-10
[82] Irina Matijosaitiene, Anthony McDowald, and Vishal Juneja, "Predicting Safe Parking Spaces: A Machine Learning Approach to Geospatial Urban and Crime Data," *Sustainability* 11, no.10 (2019)

[83] Hugo M. Verheist, Alexander Stannat, and Giulio Mecacci, "Machine learning against terrorism: how big data collection and analysis influences the privacy-security dilemma," *Science and Engineering Ethics* (2020), 2978
[84] Tarleton Gillepsie, "Content moderation, AI, and the question of scale," *Big Data & Society* (July-December 2020), 1-5
[85] Shane Quinlan in discussion with the author, August 2021
[86] Veheist, Stannat, and Mecacci, "Machine learning against terrorism," 2978-2979
[87] Vegas Tenold, *Everything You Love Will Burn: Inside the Rebirth of White Nationalism in America,* (New York, Nation Books, 2008)
[88] Tom Dreisbach and Meg Anderson, "Nearly 1 in 5 Defendants in Capitol Riot Cases Served in the Military," NPR, January 21, 2021, https://www.npr.org/2021/01/21/958915267/nearly-one-in-five-defendants-in-capitol-riot-cases-served-in-the-military; Center on Extremism, *New Hate and Old: The Changing Face of American White Supremacy* (New York: Anti-Defamation League, 2017), https://www.adl.org/media/11894/download
[89] Killion, "Terrorism, Violent Extremism,"
[90] Melanie Tory and Vidya Setlur, "Do What I Mean, Not What I Say! Design Considerations for Supporting Intent and Context in Analytical Conversation," *2019 IEEE Conference on Visual Analytics Science and Technology* (2019), https://ieeexplore.ieee.org/document/8986918
[91] On secure messaging platforms, see Brian Fishman, "Crossroads: Counter-Terrorism and the Internet," *Texas National Security Review* 2, no. 2 (2019), 85-86
[92] Brian Drake in discussion with the author, July 2021
[93] Bang Hui Lim, Dongyuan Lu, Tao Chen, and Min-Yen Kan, "#mytweet via Instagram: Exploring User Hbeaviour across Multiple Social Networks," *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining* (2015), 113-120
[94] Brian Drake in discussion with the author, July 2021
[95] Dipayan Ghosh, "Are We Entering a New Era of Social Media Regulation?" Harvard Business Review, January 14, 2021,
[96] 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/
[97] Caroline Orr, "How far-right extremsits rebrand to evade Facebook's ban," National Observer, May 10, 2019, https://www.nationalobserver.com/2019/05/10/analysis/how-far-right-extremists-rebrand-evade-facebooks-ban
[98] 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
[99] Jigsaw, *Countermasures in Practice,* Google, https://jigsaw.google.com/the-current/white-supremacy/countermeasures/
[100] Rasan Burhan and Jalal Moradzadeh, "Neurotransmitter Dopamine (DA) and its Role in the Development of Social Media Addiction," *Journal of Neurology & Neurophysiology* 11, no. 7 (2020), 507-508
[101] Soren Krach, Frieder M. Paulus, Maren Bodden, and Tilo Kircher, "The Rewarding nature of social interactions," *Frontiers in Behavioral Neuroscience* 4, no.1 (2010)
[102] Jeff Horwitz and Deepa Seetharaman, "Facebook Executives Shut Down Effots to Make the Site Less Divisive," *The Wall Street Journal,* May 26, 2020, https://www.wsj.com/articles/facebook-knows-it-encourages-division-top-executives-nixed-solutions-11590507499
[103] Amanda Lotz, "Profit, not free speech, governs media companies' decisions on controversy," *The Conversation,* August 10, 2018, https://theconversation.com/profit-not-free-speech-governs-media-companies-decisions-on-controversy-101292
[104] Kashmir Hill, "The Secretive Company That Might End Privacy as We Know it," The New York Times, January 18, 2020, https://www.nytimes.com/2020/01/18/technology/clearview-privacy-facial-recognition.html
[105] Ibid.
[106] A. Tarantola, "Why Clearview AI is a threat to us all," Engadget, February 12, 2020, https://www.engadget.com/2020-02-12-clearview-ai-police-surveillance-explained.html
[107] Ibid.
[108] Thomas Brewster, "Hackers Use Little Stickers to Trick Tesla Autopilot Into the Wrong Lane," Forbes, April 1, 2019, https://www.forbes.com/sites/thomasbrewster/2019/04/01/hackers-use-little-stickers-to-trick-tesla-autopilot-into-the-wrong-lane/?sh=6188f1a17c18

[109] Brendan F. Klare, Mark J. Burge, Joshua C. Klontz, Richard W. Vorde Bruegge, and Anil K. Jain, "Face Recogniton Performance: Role of Demographic Information," *IEEE Transactions on Information Forensics and Security* 7, no. 6 (2012), 1789-1800
[110] Patrick Grother, Mei Ngan, Kayee Hanaoka, *Face Recognition Vendor Test FRVT) Part 3: Demographic Effects*, (Washington, DC: National Institute of Standards and Technology, December 2019), https://nvlpubs.nist.gov/nistpubs/ir/2019/nist.ir.8280.pdf
[111] Ibid.
[112] Ibid.
[113] Alex Najibi, "Racial Discrimination in Face Recognition Technology," Blog, Science Policy, Special Edition: Science Policy and Social Justice, Harvard University, October 24, 2020, https://sitn.hms.harvard.edu/flash/2020/racial-discrimination-in-face-recognition-technology/
[114] Thomas J. Cowper and Michael E. Buerger, "Improving Our View of the World: Police and Augmented Reality Technology," *FBI Futures Working Group,* 2011, https://www.fbi.gov/file-repository/stats-services-publications-police-augmented-reality-technology-pdf/view
[115] Patricia Cogswell in discussion with the author, July 2021
[116] Ibid.
[117] Russell Brandom, "How facial recognition helped police identify the Capital Gazette shooter," The Verge, June 29, 2018, https://www.theverge.com/2018/6/29/17518364/facial-recognition-police-identify-capital-gazette-shooter
[118] Harwell and Timberg, "How America's surveillance"
[119] Homeland Security and Public Safety Division, Artificial Intelligence (AI) in Homeland Security
[120] Jose de Jesus Rocha-Salazar, Maria Jesuss Segovia Vargas, and Maria del Mar Camacho Minano, "Money Laundering and terrorism financing detection using neural networks and an abnormality indicator," *Expert Systems with Applications* 169, no. 114470 (169), https://www.sciencedirect.com/science/article/pii/S0957417420311209
[121] Gian Maria Campedelli, Iain Cruickshank, and Kathleen M. Carley, "A Complex networks approach to find latent clusters of terrorist groups," *Applied Network Science* 4, no. 1 (2019), 1-22
[122] Matthew Collin, "What the FinCEN leaks reveal about the ongoing war on dirty money," *Brookings Institute,* September 25, 2020, https://www.brookings.edu/blog/up-front/2020/09/25/what-the-fincen-leaks-reveal-about-the-ongoing-war-on-dirty-money/
[123] Ibid.
[124] Ibid; Richard K. Gordon, "Losing the War Against Dirty Money: Rethinking Global Standards on Preventing Money Laundering and Terrorism Financing," *Duke Journal of Comparative & international Law* 21, 2011
[125] FATF, *Opportunities and Challenges of New Technologies for AML/CFT* (Paris, France: FATF, July 2021), https://www.fatf-gafi.org/media/fatf/documents/reports/Opportunities-Challenges-of-New-Technologies-for-AML-CFT.pdf
[126] Brian Drake in discussion with the author, July 2021
[127] Ibid.
[128] Ibid.
[129] Eric Halliday and Rachel Hanna, "How the Federal Government Investigates and Prosecutes Domestic Terrorism," *Lawfare,* February 16, 2021, https://www.lawfareblog.com/how-federal-government-investigates-and-prosecutes-domestic-terrorism
[130] Jigsaw, *Countermeasures*
[131] 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/
[132] Patrcia Cogswell in discussion with the author, July 2021
[133] Shane Quinlan in discussion with the author, August 2021
[134] Ibid.
[135] Ibid.
[136] Jeff Schogol, "Why It's so difficult for the military to weed out extremists," Task & Purpose, February 19, 2021, https://taskandpurpose.com/news/military-extremists-screening/
[137] Hoffman, *Inside Terrorism,* 182-183
[138] 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/
[139] Sara Sidner and Mallory Simon, "How robot, explosives took out Dallas sniper in unprecedented way," CNN, July 16, 2016, https://www.cnn.com/2016/07/12/us/dallas-police-robot-c4-explosives

[140] Sheena Chestnut Greitens, Myunghee Lee, and Emir Yazici, "Counterterrorism and Preventive Repression," *International Security* 44, no.3 (Winter 2019/2020), 9-47
[141] Ibid.