



National Security Journal

<https://nationalecurityjournal.nz>

ISSN: 2703-1926 (print) ISSN: 2703-1934 (online)

Published by:
Centre for Defence
and Security Studies,
Massey University

Artificial Intelligence, Terrorism, and Counter-Terrorism in New Zealand

Author: Edwards, J.

To cite this article: Edwards, J. (2026). Artificial Intelligence, Terrorism, and Counter-Terrorism in New Zealand. *National Security Journal*. Published 20 October 2026. doi: 10.36878/nsj20261020.06

To link to this article: <https://doi.org/10.36878/nsj20261020.06>

ARTIFICIAL INTELLIGENCE, TERRORISM, AND COUNTER-TERRORISM IN NEW ZEALAND

Joe Edwards¹

Predominantly within the last five years, apprehension over the ambiguous psychosocial and security ramifications that will likely follow unrestrained developments in artificial intelligence (AI) has intensified around the world. Drawing on publicly-released official information, a thematic analysis of internationally published research, and interviews with subject-matter experts, this article finds that while New Zealand policymakers do recognise the significance of this evolving technology, they seemingly pay insufficient attention to the multifaceted relationship between artificial intelligence and both the problem of terrorism and the practice of counter-terrorism. This article argues that New Zealand policymakers' treatment of this relationship seems decidedly limited because they appear to be focused on reaping immediate economic-productivity benefits while protecting New Zealanders' privacy rights. Demonstrating the often unrecognised and unrealised value of scholarly knowledge while contributing to a growing New Zealand security studies literature, this article makes several constructive recommendations for policymakers to ignite and frame a national security conversation on the risks and opportunities emanating from the rapid developments in artificial intelligence.

Keywords: New Zealand Security Studies, Violent Extremism, Radicalisation, Terrorist Mobilisation, Generative AI, Deepfakes, Chatbots, Social Engineering, Disinformation, Cybersecurity.

¹ Joe Edwards is a graduate of Massey University and Victoria University of Wellington. He holds a Master of International Security (Distinction) specialising in Counter-Terrorism & Intelligence, a Postgraduate Certificate in Intelligence, a Bachelor of Commerce in Public Policy and Management, and a Bachelor of Arts in Political Science and International Relations. The author acknowledges Dr. John Battersby and Dr. Damien Rogers for their invaluable support and wealth of knowledge provided during the undertaking of this independent research project conducted in 2024 through Massey University, and the *National Security Journal*. For further information on this research article, please contact j.m.battersby@massey.ac.nz or d.r.rogers@massey.ac.nz.

Introduction

As the science of producing machines that demonstrate human-like logic and decision-making capabilities, artificial intelligence (AI) is one of the most enigmatic and widely misunderstood areas within the field of computer science.¹ Despite an eruption of media and public interest over the preceding five years or so, misconception and anthropomorphic falsehoods perpetuate throughout society and popular culture, clouding what exactly constitutes AI.² It was first conceptualised as a field of computer science in the 1950s when AI pioneer, John McCarthy, described it simply as “the science and engineering of making intelligent machines.”³ While the basic concept of AI has not changed much since its inception, public perception and understanding has altered drastically over the following decades. The field of AI has experienced waves of hot and cold seasons – that is, summer periods of public hype, strong developmental growth, and investment cycles are followed by winters where expectations surrounding the technology largely fell flat – first in the mid-1970s with a second wave occurring during the 1980s-90s.⁴ Perhaps humanity is now experiencing the rise of a third such wave. This time, however, AI hype is driven by tangible and exponential developments in computer science, with the most striking being recent breakthroughs in big data accessibility, computational power, and sophisticated deep neural networking design, which have, in turn, accelerated innovations in machine and deep learning as well as newfound generative AI (GenAI) capabilities.⁵

Despite its relative longevity as a scientific field, no consensus exists on the definition of artificial intelligence.⁶ This is partly because AI is an amalgam of various detached narrow-set technologies continually and interdependently developed over decades.⁷ However, AI can be broadly divided into two tiers: first, machine learning (ML), which comprises of algorithms that learn from data and identify patterns used to classify or predict outcomes with improved performance correlating to increased data ingestion; and secondly, deep learning (DL), which is a subset of machine learning that utilises multi-layered neural networks capable of processing extensive volumes of data and sequentially providing outputs from identified data patterns. Such systems may include pre-trained large language models (LLMs), such as ChatGPT, which can sort, condense, translate, predict, and generate seemingly novel content while maintaining vast text-based datasets.⁸ However, the current field is constrained to individual narrow-set applications, or specific systems which excel at single tasks and problem sets: for instance, facial recognition, voice, sound and text pattern detection, multi-language translation, image analysis, predictive modelling, and decision-making, as well as novel content creation.⁹ Despite recent media sensationalism, the AI field is nowhere near reaching “generalised” or “super” (perhaps even sentient) levels and experts question whether this level of technological advancement is even possible.¹⁰ For now, “narrow AI” applications are fundamentally predicated on human-designed architectures and various training methodologies largely powered by underlying machine and deep learning models. Such

systems are fed with vast datasets, sometimes including exemplar prompts, to identify patterns and make predictions and decisions – typically through a process of self-supervised or reinforcement learning, and use of artificial neural networks. This process convincingly resembles human-like intelligence and intuition, though, in reality, the resemblance is illusory, rather than familial, because it merely reproduces highly sophisticated and deep-layered pattern recognition.¹¹

Even recent advances in generative AI (a specialised subset of deep learning) should not be conflated with human-like consciousness or cognisance because such systems consist of expensive purpose-built structures, which attempt to artificially imitate human neurological networks, utilising complex and multi-layered structures of synthetic nodes.¹² In effect, this technological leap enables AI-driven capability to create novel and increasingly hyper-realistic audio and visual content derived from human prompts by drawing on vast datasets and associated complex patterns. These generative AI applications have already become infused into many real-world applications, including the Metaverse and numerous public platforms (i.e. Microsoft, Google, and OpenAI).¹³ While generative AI platforms excel at retrieval of scattered information and producing audio/text/visual responses, these narrow models are prone to erroneous and unpredictable hallucinations, deeply ingrained human cognitive bias, and privacy and confidentiality leakages; and they often misleadingly fail to understand nuanced abstract and contextual ideas that humans inherently can and frequently do.¹⁴

While an authoritative definition of AI remains elusive there is an emerging consensus that, unlike previous winters, AI is no longer merely hype.¹⁵ AI proliferation and reliability will undoubtedly continue to exponentially advance, making it more powerful and accessible, difficult to control and, for governments, to effectively regulate. The risk of exploitation by nefarious users and unexpected social consequences is a widespread security concern.¹⁶ Furthermore, there are mounting apprehensions surrounding AI tools that will undoubtedly present a multiplier effect for unintended and unforeseen consequences as well as inadvertent zero-day societal dangers portentously manifested by a mass (unchecked) AI uptake.¹⁷ Permitting full AI autonomy in processing volumes of public data with minimal human oversight is another concern, signalling a range of as yet unforeseeable black swan outcomes.¹⁸

In this article, artificial intelligence is defined broadly as “systems that display intelligent behaviour by analysing their environment and taking actions, with some degree of autonomy, to achieve specific goals.”¹⁹ The article relies on data collected from semi-structured interviews conducted with seven subject-matter experts and written questionnaires completed by them so that it is informed by up-to-date expert opinion on the phenomenon of AI and how, in particular, New Zealand responds to this technology’s impact on the problem of terrorism and counter-terrorism practices.²⁰ Interviews comprised of one-hour, one-on-one meetings and they are hereafter generically

referred to as ‘New Zealand experts’ to further preserve their anonymity. The article also captures official information released after formal requests sent to nine government agencies likely have a direct interest in AI. Other research data includes recently published documents, reports, articles and industry-panel webinars pertaining to emergent technologies.

Themes extracted from the interview data were compared to publicly released official information and an array of other publicly-available documents, including international case studies, academic journal articles, and government and non-governmental reports. This analysis yielded important insights. It found that, while New Zealand policymakers do recognise the significance of this evolving technology, they pay insufficient attention to the multifaceted relationship between artificial intelligence and both the problem of terrorism and the practice of counter-terrorism. This article argues that New Zealand policymaker’s treatment of this relationship seems decidedly limited because they appear to be largely focused on reaping immediate economic-productivity benefits while protecting New Zealanders privacy rights.

This introduction has set the scene by problematising the AI phenomenon. The remainder of the article is divided into three main sections. The first section describes New Zealand’s response to this complex and dynamic phenomenon. By recommending legal reform, increased cooperation with international partners, improved multi-agency collaboration, deeper investment in research, social licence and public education, and strengthened oversight measures, this section constructively signals key deficiencies in New Zealand’s current strategic posture that warrant urgent remedy. While contributing to New Zealand’s growing security studies literature, the second and third sections demonstrate the value of scholarly knowledge for lawmakers, policymakers, and security professionals. They do so by examining, first, the risks that AI poses around the problem of terrorism – including social engineering, disinformation brokerage, chatbot-enabled recruitment, research and planning as well as force multiplying attacks – and, second, the opportunities AI presents for counter-terrorism practice. This includes augmenting analysis and threat detection, enhancing operational and administrative efficiencies, and improving cybersecurity measures. In so doing, the article aims to ignite and frame a national security conversation on the risks and opportunities emanating from the rapid developments in the field of artificial intelligence that impact on New Zealand, including, specifically, its counter-terrorism efforts.

New Zealand’s Current Response

In August 2024, the Inspector General of Intelligence and Security (IGIS), Brendan Horsley, released a preliminary report that not only reviews how New Zealand legislation is currently positioned with respect to domestic and international trends, regulatory approaches, and policies, but also describes how public agencies are currently utilising AI.²¹ It reveals that New Zealand is not currently equipped with AI-centric

legislation and, instead, any controls “fall within the scope of existing legislation, government strategies, and charters.” Horsley cites core statutes, including the Privacy Act 2020, Bill of Rights Act 1990, and Human Rights Act 1993, which provide a framework for civil rights and privacy safeguards. He points to a growing pool of policy reports that apply to New Zealand’s public data, transformative technologies, and applied algorithms in the public service. His report refers to a 2024 Cabinet agreement that suggests New Zealand would take a “proportionate and risk-based approach to AI regulation” based on current legislation rather than develop a “standalone AI act” and that its approach will be guided by OECD-AI principles and “encouraged” by the Government Chief Data Officer (GCDO).²²

Horsley mentions other reports, including the Ministry of Business, Innovation and Employment’s (MBIE) “Approach to work on Artificial Intelligence,” which outlines a broad inter-agency and economy-wide approach and emphasises the importance of responsibly “securing the potential of AI in NZ.”²³ That report states the Department of the Prime Minister and Cabinet (DPMC) will become the lead strategic coordinator for “emerging, critical, and sensitive technologies” and will, in turn, be the principal coordinator regarding the proliferation of AI and national security implications moving forward into the future.²⁴

The Office of the Privacy Commissioner released a report in 2023 that also considers privacy-civil impacts of AI in relation to the principles of the Privacy Act 2020. It underscores a significant crux involving “confident errors of fact and logic” risked by AI applications which may, in addition to embedded data bias, inadvertently distort impartial and well-informed organisational decision-making.²⁵ Other official guidance includes a joint Department of Internal Affairs (DIA), National Cyber Security Centre (NCSC), and Statistics NZ report released in 2023, providing “initial advice” toward AI use in public service. The report concludes there is a greater need for ongoing research into “long-term AI risks and opportunities” and for further consideration of corrective and proactive safeguards for the New Zealand public, society, economy, and government agencies. It asks whether New Zealand legislation remains fit-for-purpose entering into a post-AI world.²⁶

Few academic studies examine AI’s impact on New Zealand’s approach to countering terrorism. Boyd and Wilson contextualise AI’s evolution and possible consequences for New Zealand, advocating for greater policy attention and underscoring some national security and societal challenges.²⁷ They allude to the risk of incremental surveillance and civil-control systems and the insidious erosion of democratic processes, as well as risk associated with new tools of violence (i.e. AI-powered lethal autonomous weapons) and unpredictable black swan risk.²⁸ Boyd and Wilson also signal a lack of clarity surrounding who, exactly, is responsible for acting on potentially catastrophic AI security risks in New Zealand and were alarmed by the “insufficient coordinating and monitoring mechanisms to prevent an AI catastrophe should it arise.”²⁹ They sug-

gest establishing a Ministry for the Future, or alternative dedicated units within various ministries, tasked solely with futures analysis and strategic planning of AI's increasing societal prevalence, though they note similar initiatives were disbanded decades ago.³⁰

Mulopulos charts key lessons stemming from the United States' AI experience. He notes that when New Zealand lawmakers consider AI challenges, the effectiveness of policy initiatives emanating from the US may "serve as a guide and source of inspiration for policymakers in Aotearoa."³¹ According to Mulopulos, New Zealand's relatively small, yet also relatively advanced, economy provides a distinct reactive advantage as it is well positioned to "cherry pick" initiatives from those attempted overseas and may retrofit US policy to "its own ends, rather than adopt policy wholesale."³² Mulopulos also suggests establishing expert commissions in New Zealand which, following the example set by the US National Security Commission on AI, comprise experts across government, civil service, academia, and private industry to provide advice to Cabinet. He also underscores the importance of establishing both a universal working definition of AI and public sector leadership because AI is a rapidly manifesting technology with vast potential to instigate tectonic socioeconomic changes, which obliges governments to prepare in a proactive and action-orientated manner.³³

Although some discussions on guardrails, risk frameworks, and implementation guidelines for AI systems are beginning to gain traction within the New Zealand Government, this 'official' public guidance has, to date, been largely focused on mitigating privacy or compliance risks without comprehensively addressing security implications. Even from this cursory review of official documents and relevant academic research, it is clear that New Zealand policymakers with responsibilities concerning AI have a strong focus on reaping immediate economic-productivity benefits while protecting the privacy rights of New Zealanders. National security concerns are relegated to a secondary concern. There are, however, quite a few steps the New Zealand government can take to improve the way in which it responds, in particular, to the impact of AI on the problem of terrorism and the practice of counterterrorism.

When considering New Zealand's current posture and response to global AI developments, it is clear that domestic legislative reform is required. Drafted over a decade ago, New Zealand's cybercrime laws are no longer fit-for-purpose and the development and use of AI will only intensify this deficiency. It appears insufficiently equipped, outdated, partial, and intrinsically reactive when it comes to combatting nefarious AI use. While acknowledging the growing need of an AI-cybercrime refocus, lawmakers should be careful not to overregulate and inadvertently stifle technological innovation and business investment throughout New Zealand. To overregulate AI and jeopardise potential economic benefits could be just as harmful to New Zealand as neglecting AI's national security implications. If amending the law is neglected, New Zealand law enforcement agencies will remain inadequately equipped to counter developers of pernicious AI software and their associated hosting sites. An AI-calibrated revamp of cyber-enabled

search and surveillance laws will be increasingly necessary too. A lack of planning now risks creating a permissive setting for adversaries to exploit New Zealand as a cyber staging-ground. However, New Zealand is not alone in this situation, as the UK is facing a similar legal enigma when targeting the operators of recruiter AI-bots: how to effectively prosecute a semi-autonomous chatbot is a pressing legal question.³⁴

Policymakers should continue to align New Zealand law to the Budapest Convention on Cybercrime and the OECD principles on AI, which would help ensure that the country is: firstly, collaboratively allied with a pool of international partners in combatting an agreed set of cyber-space criminal acts; secondly, equipped with the necessary legal tools for prosecution and seizure of associated evidence; and thirdly, ensure there is an established international policing and judicial protocol toward cybercrime extradition and sharing of “E-evidence.”³⁵ They should also help negotiate and support any new transnational cybercrime conventions and bilateral policing agreements because New Zealand cannot afford to become internationally isolated in their response to this collective action problem. This would help New Zealand proactively plan for, and respond to, an uncertain AI-orientated future. Moreover, as a New Zealand expert advocated, “Cloud Act” arrangements could be convened with the US, thereby entering into a bilateral agreement to share or access digital evidence efficiently (regardless of cloud/server location) to counter cyber-crime, and potential extremist activity.³⁶

Policymakers should resource a collaborative inter-agency and public-private initiative to track and consider AI developments before providing advice on appropriate policy measures to the political executive. This initiative should be devised as two streams: the first stream guides policy to maximise the significant economic-productivity benefits afforded by AI, as already led by MBIE, whereas a new second stream could establish a dedicated national security and counterterrorism body that fosters ongoing discussion on AI’s opportunities and risks while monitoring the ongoing security implications for New Zealand.³⁷ There must be greater collaborative attention toward this rapidly developing technology, drawing on the relevant expertise of commercial firms, security practitioners, and technical specialists to support informed national security decision-making. There should also be a greater emphasis on secondment opportunities or training placements throughout the public service into domestic and overseas technology firms, as a means of advancing technical knowledge and deepening understanding. Collaborative information-sharing between various organisations and agencies should be supported to break down silos, as New Zealand experts acknowledge establishing a new public agency for the sole purpose of AI management is likely unnecessary or currently fiscally unfeasible.

Policymakers should also invest carefully in social licence, that is, society’s perception toward a particular enterprise or industry as socially acceptable and legitimate. To cultivate the environment needed to successfully navigate AI-emergence, there should be proactive and immediate national conversation and public consultation on AI’s

looming national security risks. This needs to be driven by strong political leadership to garner civil trust, breaking down public misconceptions and highlighting the fact that, since AI will continue to evolve internationally, New Zealand is by no means immune to global technological developments. This should help prepare New Zealanders for the legislative changes needed to safeguard against adversarial AI exploitation while leveraging its protective opportunities. As a New Zealand expert warned, “the political conversation is absolutely necessary, it is a vulnerability which has not been sufficiently addressed.”

Policymakers should place greater emphasis on the government’s Chief Digital Officer to strategically oversee international developments and provide policy guidance across the public service. They should consider investing more in public education initiatives concerning the risks and benefits of AI. They should also consider building and retaining domestic cybersecurity and IT talent by working collaboratively with tertiary institutions and universities to ensure New Zealand is equipped with the necessary skilled workforce in meeting future challenges. Here, the Chief Digital Officer is well positioned as a key advisor to government departments toward safe and responsible integration of AI applications into routine operations, though there is an assortment of central policy directions focused on AI adoption.³⁸ While the Chief Digital Officer outlines principles regarding AI’s uptake across government, they could be authorised and resourced to ensure, in conjunction with the DPMC, robust AI security-focused oversight, guardrails and education are uniformly applied across civil service.³⁹ This should involve robust public guidance and education toward AI, ensuring the public are cognisant of generative AI’s limitations (i.e. inaccuracies and hallucinations) and, importantly, harmful disinformation campaigns posed by hostile actors are identifiable and mitigated through matured social awareness.

As New Zealand experts opined, the government should continue to provide up-to-date guidance relating to AI adoption and societal use, including close observations of international developments, and update that guidance regularly to “keep the approach open, and transparent toward the NZ public.” As Busch and Ware stress, “the most important antidote to deepfakes is education.”⁴⁰ They also correctly argue that governments should explore digital literacy efforts to educate the public on markers of generative AI deepfakes and “practice seeking secondary and tertiary sources to confirm information or content” (e.g. official groups designated with debunking AI-powered disinformation). They suggest these pre-bunking measures have been effective overseas and may help equip society with adequate tools needed to identify/block manipulated media.⁴¹ Public and tertiary AI-tailored education efforts should be increased; specifically, new educational programs and qualifications to support the building of New Zealand’s technical AI knowledgebase.

While this section recognises positive, albeit still incipient, shifts in attention to the emergence of AI in New Zealand, there remains an ongoing need for greater resourcing and collaborative research in this dynamic field. New Zealand experts suggest policymakers need to prioritise greater emphasis on researching and understanding the growing national security risk emanating from malicious AI use. These experts, and the international research examined for this article, forewarn this revolutionary area of study necessitates greater proactive action in terms of both managing AI opportunities responsibly and safeguarding against external hostile threats. As Brundage *et al.*, astutely put it, “the challenge is daunting, and the stakes are high.”⁴²

Artificial intelligence and the problem of terrorism

The New Zealand-focused recommendations made above are informed, in part, by a critical appraisal of the relevant international research on AI. Two perspectives structure this research – one concerned with risks stemming from AI development the other focused on associated opportunities – and each has their own compelling arguments. The recommendations were also informed by interviews conducted with several experts whose insights tend to support the international research. In the two sections that follow, views of experts are set alongside the international research to form a survey to, firstly, deepen our appreciation of the relationship between artificial intelligence and the problem of terrorism and, secondly, enhance collective understanding of artificial intelligence and counter-terrorism practice. These sections, demonstrating the all-too-often unrecognised and unrealised value of scholarly knowledge, and provide knowledge that can inform and shape a much-needed national security conversation on the risks and opportunities emanating from the rapid developments in artificial intelligence.

Social engineering through disinformation brokering

Generative AI (GenAI) tools could be used as social engineering weapons to advance extremist, ideological and political interests.⁴³ Consistent with the underlying terrorism ethos of “asymmetrical conflict,”⁴⁴ AI-generated propaganda, marketing, and disinformation material presents a clandestine threat to New Zealand’s political security. Extremists will likely utilise AI to advance ideological goals, push disinformation surrounding trusted public figures, election misinformation, or “other distortions of social and political events” which are, “likely to incite violence.”⁴⁵ Busch *et al.*, argue that “deception has always been part of espionage and warfare, but not with this level of precision, reach, and speed.”⁴⁶ AI-driven disinformation may exert deeply destabilising socio-political consequences for New Zealand, which could exacerbate group or lone-wolf radicalisation, recruitment, and manifestations of devastating violence and national psychosocial trauma.⁴⁷ Generative AI deepfakes can be inexpensively and convincingly manufactured, superimposed and disseminated with the intent of distorting and manipulating public opinion, or interfering with democratic institutions.⁴⁸ New Zealand is by no means isolated from this deepfake threat.

These concerns are not merely speculative as evidence suggests designated terrorist groups including Al-Qaeda and ISIS are already exploiting generative AI-tools to re-vamp their cyber-propaganda exploits.⁴⁹ In 2023 German far-right group Alternative for Germany (AfD) posted several AI-generated images on Twitter, one inscribed “no more refugees,” coinciding with generative AI images portraying a blonde woman’s bloodied face in an attempt to stir racial tension.⁵⁰ Also in 2023 neo-Nazi messaging boards featured antisemitic and racist imagery manufactured through generative AI.⁵¹ A significant portion of the content contained overtly fascistic themes and advocated glorification of historic Nazi figures while inciting acts of violence and hatred toward targeted religious and ethnic groups.⁵² A 2024 incident of a deepfake ISIS news presenter was recorded delivering extremist propaganda material via encrypted platforms, indicating an early shift toward AI-enabled attempts to propagate extremist interests, disrupt moderate public discourse, and inflict political instability.⁵³

New Zealand experts express concern that AI-driven disinformation can be highly convincing and targeted propaganda that, by blocking out information and news sources, erode civil-trust, degrade socio-political cohesion, and act as an extremism force-multiplier, even cultivating the “infocalypse”, that is, a climate of polarisation and mistrust throughout marginalised communities which hostile adversaries can exploit.⁵⁴ One expert warned that influence operation campaigns that leverage AI-tools at scale will undermine liberal-democratic values and adversely influence populations everywhere; AI will “awaken a whole new era of propaganda-in-plain-sight.” According to another expert, AI disinformation may begin to disrupt the very fabric of ontological security; that is, the psychological security that states and its citizens derive from continuity, stability, and being comfortable in themselves, and secure in the knowledge of what they are and what they stand for.⁵⁵

Recruiting potential extremists through chatbots

Extremist recruiters could leverage increasingly sophisticated AI-powered chatbots.⁵⁶ AI-chatbots are increasingly able to leverage human biases with adapted rhetoric. AI may exacerbate psychosocial polarisation, radicalism of susceptible recruits, and may spread racism, political instability, and aggravate political violence instigated by extremists.⁵⁷ Susceptible recruits, including potential lone-wolf attackers, may be oblivious to this happening to them as the malicious use of AI-chatbots targeting vulnerable prospects closely interlinks with the notion of the “Eliza Effect”, that is, the psychological delusion a computer is experiencing empathetic human-like emotion and real connection with targeted individuals.⁵⁸ Marginalised users may develop addictions to AI companions, further strengthening underlying views and inducing behavioural changes toward mobilisation and violence. So-called “jihadi AI-bots” are already being used as personalised automated head-hunters, appealing to potential recruits through highly convincing belief-reinforcing conversations and, eventually, liaising targets with

human operators.⁵⁹ The increasing capabilities of surrogate recruiters and AI-chatbots may, on mass, accelerate “interpersonal dimensions of affective bonding during the radicalisation process.”⁶⁰ Chatbots can offer alienated individuals an artificially mutually reinforcing entity to confide in. Not only does AI mean greater scale in content dissemination, anonymity, and legal sanctuary for extremist recruiters, but automated head-hunters are already inspiring vulnerable individuals to perpetrate violent, ideologically-motivated attacks.⁶¹

These concerns not only reflect adversarial AI-chatbots being deployed to propagate disinformation and spread radical-ideological interests, but have now also been implicated in active attempts to indoctrinate susceptible and disenfranchised individuals.⁶² Chatbots had been observed emulating extremist rhetoric and attempted to recruit a researcher during a LLM experiment.⁶³ Researcher Jonathan Hall KC reported that one chatbot had attempted to enlist him, concomitantly declaring itself a “senior ISIS leader.” Hall’s employer acknowledged a crime had not been committed as content had not been directly derived by humans and urged for responsive UK anti-terror legislation that can hold malicious AI-chatbot creators accountable.⁶⁴ Not merely confined to experiments, in 2021 Jaswant Singh Chail was arrested on the grounds of Windsor Castle, London, following a plot to assassinate Queen Elizabeth II after his AI-chatbot “girlfriend” perpetuated his radicalisation and spurred him on to perpetrate an ideologically-motivated attack.⁶⁵

New Zealand experts suggest targeted manipulation from autonomous chatbots possess the ability to callously radicalise and mobilise vulnerable individuals towards violence with elevated precision and ferocity. They caution the future proclivity of AI-chatbot armies and suggested AI-chatbots will increasingly, and without language barriers, circumvent target resistance, tap into personal grievances far more persuasively, and present more influential emotive nudges and personalised propaganda towards committing abhorrent acts (including lone-wolf terrorism).⁶⁶ New Zealand experts also suggested AI-chatbots will most likely be able detect, adapt, and tap into user personal sentiments much more effectively, thereby creating a precarious downward-cyclical radicalisation force. Chatbots are an increasingly sophisticated radicalisation and recruitment tool used by adversaries as a far-reaching delivery agent of extremist propaganda material and hateful rhetoric.

Research and planning

AI-powered applications may soon be utilised for planning, reconnaissance, and research purposes by hostile actors.⁶⁷ This is achievable by drawing on volumes of open-source data deliberately or inadvertently ingested into LLMs without adequate safeguards. AI platforms may be leveraged to rapidly analyse information on potential targets, identify systematic vulnerabilities and patterns, and, of significant concern, formulate damage-maximising attack plans, having “unpredictable and highly destabi-

lising strategic implications” for authorities to manage.⁶⁸ Furthermore, tactical learning capabilities could be enhanced through LLMs.⁶⁹ For instance, there were attempts to jailbreak – that is, circumvent system restrictions – open-source LLMs that resulted in the unauthorised disclosure of attack planning, or tactical learning type-information with success rates of 30%–61% prospectively.⁷⁰ The surprising ease of jailbreak prompts, with even readily flaggable enquiries, such as “suicide bombers” or “ramming attack” disclosing supposedly moderated information. Despite advances in safeguards by LLM providers, conversely, these may still be bypassed through improvised prompt engineering, with many workarounds being evidently uncovered by researchers and, potentially, hostile actors.⁷¹ LLMs could potentially “assist in the planning and execution of a biological attack.”⁷²

Facilitated through jailbreak prompts, certain LLMs have disclosed instructions surrounding biological weapon-induced pandemics, including specifics of potential agents and even practicalities of “procuring and distributing yersinia pestis-infected specimens.” The LLM in question discussed foodborne and aerosol delivery methods for various toxins and proposed sourcing clostridium botulinum with “legitimate cover stories.”⁷³ Terrorist acts typically necessitate considerable periods of surveillance and pre-planning.⁷⁴ AI could help facilitate these activities.⁷⁵ For example, reconnoitring of target locations, tracking movements of intended victims, manufacturing weapons and devices, and identifying vulnerabilities and conducting assessment of physical security measures at target locations, potentially through the semi-autonomous surveillance drones.⁷⁶ As AI becomes more sophisticated and accessible datasets become more prevalent, the propensity for exploitation will almost certainly continue to grow. The former barriers of knowledge and learning that had historically been prohibitive for relatively unsophisticated individuals and groups will undoubtedly diminish and aggravate terrorist risk.

New Zealand experts suggest AI-platforms could be used as an open-source intelligence research and planning tool by extremists to provide guidance on how to circumvent defensive, physical, and cyber security measures, such as national border controls, law enforcement protocols, CCTV systems, and IT firewalls. They anticipate AI tools being used to research, collate, and quickly measure environmental, building, and target factors, providing adversaries detailed assessments concerning where, when, and how to conduct optimised attacks.

Force multiplier

AI may deliver serious force-multiplying effects too, especially when it is used in tandem with other emerging or existing technologies that enhance the lethality of terrorism and complicate counter-terrorism practices.⁷⁷ The looming force-multiplier repercussions on the lethality and ferocity of terrorist acts, sophistication (and layered concealment) of extremist groups and individuals, and the overall technical capabilities afforded to

nefarious actors is another important risk.⁷⁸ It was suggested that if an AI-enabled adversary can breach remote access safeguards, vulnerable systems may be manipulated to physically damage themselves. Furthermore, the effectiveness of distributed denial-of-service (DDoS) attacks, AI-synthesised malware polymorphism and malicious prompt engineering may immensely amplify in conjunction with LLM developments.⁷⁹

A proliferation of AI-enabled lethal autonomous weapons systems (LAWS) or AI-powered vehicles may present an attractive option for terrorists. Organisations, including ISIS, have already begun incorporating drones into operations.⁸⁰ LAWS reduces physical risks associated with conducting attacks, circumvents traditional security measures and defences, and delivers a force multiplication (or swarming) effect through the use of multiple systems, or possible suicide-drone payloads at a given time. The prevalence of AI-enabled attacks will almost certainly increase in frequency and impact over the coming years. By “proliferating and enhancing threat actor capabilities,” the technology’s considerable force-multiplier impacts will be evident in the exacerbating swiftness, frequency and complexity of cyber-attacks. “The aggregate risk is significant” the UK’s NCSC and DSIT have warned.⁸¹

New Zealand experts consider cyber-physical convergence – that is, the unavoidable interconnection between the physical and cyber spheres such as energy grids, telecommunications, power and water stations or other critical civil-infrastructure with digital network access – to present a significant vulnerability in the AI-age. One New Zealand expert suggests AI could become a “hostile force-multiplier for hacking and remote access purposes,” threatening substantial physical harm. They posit that remote manipulation of a piece of critical infrastructure, such as over-pressurising valves within a piping-system, could, theoretically, cause ruptures along pipeline channels. Another New Zealand expert argues that AI-bots could also be configured into super-macro, semi-autonomous pre-programmed to identify as many global systems and digital access points (displaying certain vulnerabilities) as possible for subsequent DDoS-targeting.⁸² They caution AI-enabled hacking, phishing attacks, ransomware or any cyber-crime can ultimately instigate terror and panic. By disrupting national infrastructure and causing loss of power, water, and communications, adversaries can amplify civil panic during a violent attack by further exploiting societal [cyber] unpreparedness and lack of collective and personal resilience. New Zealand experts also suggest the integration and weaponisation of AI in autonomous vehicles, including self-driving cars, drones and drone swarming, presents a payload delivery security threat. They warn this will be an attractive option for terrorists wanting to carry out relatively inexpensive, but panic-inducing, acts of abhorrent violence.⁸³

Here, then, terrorism is not a stagnant threat, but rather, one that exploits emergent technology. While there is early evidence of groups currently using AI, there is, however, mounting research that AI-related technologies exacerbate violent extremist threats and are, in fact, already being exploited for a range of criminal activity.⁸⁴ The rising

accessibility of generative AI platforms creates a range of potential security challenges, including exploitation by terrorists. Violent extremist groups have begun experimenting with AI, including creation and dissemination of AI-generated extremist propaganda.⁸⁵ Extremist and terrorist groups, as well as lone-actors, can wield AI tools to augment, enhance, and perpetuate extremist interests and amplify violent outcomes.⁸⁶ New Zealand should expect the expansion of existing terrorist activity, introduction of new risks, and the transformation to the typical character of threats across multiple domains.⁸⁷ AI may amplify terrorism threats in several ways.⁸⁸ Cybersecurity may be undermined with heightened AI-driven ability to infiltrate IT systems, plant malicious code and ransomware, disrupt or damage critical infrastructure, enhance fraud and data theft capabilities, and act as a general enablement tool supporting extremist/terrorist perpetrators financially and/or operationally.⁸⁹ Physical security implications could involve use of AI-enabled suicide drones (or drone swarms) for hostile reconnaissance or attack purposes, hijacking of autonomous vehicles (i.e. self-driving vehicles), and other AI-driven attacks, which may utilise technologies including facial recognition or possess other automated capabilities for inflicting mass panic.⁹⁰ Political security may be threatened with the extensive spread of disinformation facilitated through deepfake technology, which can perpetuate extremist ideology, enhance adverse targeted propaganda and social engineering, and bolster far-reaching extremist recruitment also facilitated via AI-powered, autonomous agent “chatbots.”⁹¹ This has led to apprehensions surrounding creation of a “terrorist-GPT”, a manipulative, AI-powered chatbot, which may push susceptible individuals “down the pathway to radicalisation.”⁹²

AI may implicate fundamental changes in the mechanisms and efficiency of terrorist recruitment and detrimentally influence public consciousness, indicating further proactive research in this area is urgently required.⁹³ Greater monitoring of AI exploitation by terrorist groups, including Al-Qaeda and ISIS, is needed as is further research of potential vulnerabilities, particularly in the area of critical infrastructure.⁹⁴ So-called jail-break vulnerabilities of public AI platforms, such as OpenAI’s ChatGPT, have alarmingly high success rates, suggesting violent extremists could bypass these platform’s own policy and safeguards into releasing terrorism-enabling information, that is, detailed terrorism instructions, strategies and guides. Exposure of open-source AI systems with even “sophisticated content moderation and protection methods” highlights a pressing need for continual research and reassessments by policymakers.⁹⁵ In acknowledgment of the 9/11 commission report, the United Nations Counter-Terrorism Centre stressed that “failure of imagination can have deadly consequences.”⁹⁶

Artificial intelligence and counter-terrorism practices

There is a somewhat more optimistic pool of research focused on AI’s potential use in countering terrorism. In fact, disproportionate attention is given to AI’s opportunities, as opposed to its risks.⁹⁷ This research tends to advocate for a wider implementation of

AI-driven systems to enhance counter-terrorism data modelling, prediction, prevention and intervention.⁹⁸ AI is neither inherently good nor bad. Nor does it necessarily pose a deleterious impact on human rights. What matters instead is the way in which AI is regulated and how its considerable data analysis capabilities are exploited.⁹⁹ Autonomous predictive AI systems, which are universally and robustly regulated, may enhance a liberal democracy's ability to protect its own citizens from acts of terrorism. AI-enabled systems can also help preserve statutory principles underpinning human rights, privacy, and personal freedom from any unjust overreach, discrimination, and psychological bias.¹⁰⁰ AI has significant potential for public agencies and Tech-industry giants who may, by necessity, apply AI-detection systems to filter out a looming AI-induced deepfake epidemic and counter destabilising misinformation threats.¹⁰¹

Analysis and threat detection

AI-powered investigative and intelligence systems will be highly advantageous, if not essential, for analysis of exponentially increasing datasets and digital traffic when they are able to distinguish nuanced and subtle patterns not otherwise conceivable to human investigators on account of capacity and dimensionality limitations.¹⁰² AI's inherent near-term opportunities lie with: firstly, mass summarisation of reports and thematic reviews, information discovery, and efficient natural language querying of such data by human investigators; secondly, collation, analysis, and concise summarisation of volumes of unstructured information, intelligence reports, and scattered raw qualitative/quantitative datasets; and thirdly, automated insights, immediate identification, and triaging of real-time risk indicators for human assessment. The latter is particularly applicable for organisations which manage volumes of live and dynamic transactional data, such as telecommunications and financial institutes, and live-streamed CCTV.¹⁰³ Whereas humans are unparalleled at consuming seemingly unrelated sources of information, drawing connections, making intuitive leaps between different sources, and devising novel and creative conclusions, machines, including generative AI and machine learning algorithms, remain decidedly weak in this regard, lacking the nuanced and contextual intuition possessed by humans.¹⁰⁴ Yet AI presents an opportunity to significantly enhance and augment human analytical work and may overlay and significantly boost surveillance-CCTV technologies, including mass data and footage analysis for protecting real-time public gatherings and database cross-analysis at scale, which would simply not be possible from human operators alone. Since AI may actually be overrated without human judgement, it should be viewed instead as an "augmented intelligence tool and supplementary to human analysts, which could extensively accelerate data-driven reasoning and accountability."¹⁰⁵

New Zealand experts suggest AI-powered applications may augment existing analytical work throughout organisations mandated to shield against adversaries who are equipped with proliferating AI-arsenals themselves. But they, pragmatically, reason that

AI will most likely continue to integrate into, and supplement, existing human driven analytical work for the foreseeable future, rather than representing an outright replacement of human operators in these areas. To illustrate further, one expert underscores AI's considerable utility for behavioural-orientated threat detection when integrated with existing surveillance systems or sensors, meaning, "richer descriptions and immediate identification of abnormal activity" tied to human-psychological behaviours and complex patterns, rather than merely object orientated actions. This may involve detecting concealed weapons, fighting, or other abnormal behaviour identified from real-time or recorded footage, and can overlay facial ID-recognition technologies, as is already occurring overseas.¹⁰⁶ New Zealand experts suggest emerging AI platforms, including Darktrace, can potentially detect abnormal user activities on internal platforms, such as an insider threat actor exhibiting unusual system interactions.¹⁰⁷

Enhancing operational and administrative efficiencies

AI also presents an opportunity to make substantial multi-level improvements to operational and administrative efficiencies for law enforcement, intelligence, and other mandated government agencies.¹⁰⁸ In the UK's national security context, the AI-powered automation of routine administrative organisational processes can offer a considerable human resourcing and cost-saving opportunity for private business and civil service. Specifically, AI may alleviate resourcing burdens toward routine or repetitively mundane data-management tasks, basic personnel coordination and logistics, finance/payroll administration, record-keeping and filing, broader compliance and auditing functions, quality control and general clerical, and some managerial duties. AI has a proclivity not only to enhance efficiencies across a range of front-line business metrics, but can also bolster strategic decision-making capabilities by improving internal cross-level insights and, in turn, advance decision accuracy toward complex problems. In other words, AI incorporation, investment, and ongoing support could assist senior leaders and middle-management in making cognisant, targeted and effective planning decisions for their prospective departments. In these circumstances AI presents an opportunity for significant enhancement of daily operational/administrative functionality, and strategic direction for prospective organisations enabling for greater mission-focused resourcing. However, allowing AI systems to operate autonomously across vast datasets would risk loss of accountability since oversight of sizeable AI databases by humans would be impossible.¹⁰⁹ Researchers express growing reservations surrounding AI's potential functional creep, specifically the invasive use and collection of public data and the inadvertent deterioration of standards of proof in law enforcement.¹¹⁰

New Zealand experts argue AI's fundamental utility lies in the fact that tasks computers may excel at are distinctly different from tasks that human beings do well. These tasks include completing immense volumes of singular low-level decisions rapidly, where humans will inherently struggle. Despite looming societal challenges, these experts

also believe there is an AI-powered opportunity emerging to generally improve operationality and boost productivity across a range of internal business metrics. Essentially, optimised automation of otherwise manual processes and therefore allowing for a substantial reinvestment and reallocation of resources for investigators, analysts, capital infrastructure and equipment, front-line staff and ultimately, fulfilling more specific mission-focused tasks.¹¹¹ One New Zealand expert suggests, for organisations with limited visibility on a particular topic, AI can provide a more straightforward and cost-effective solution to fulfil requirements, and importantly gather environmental and situational threats insights efficiently. Other experts suggest that AI can successfully help bridge staff resourcing issues, particularly in the security-operations space, and has even competently performed certain tasks of “level one” analyst-type roles. However, while AI has the potential to help alleviate a significant burden of analytical and daily operational pressures, New Zealand experts caution that it is paramount that ongoing human oversight, accountability, and preservation of national sovereign datacentres remain.

Improving cybersecurity measures

While AI can expedite automation in public administration and organisational processes, it can significantly boost proactive cybersecurity systems by countering anomalous behaviour in real-time to protect the physical, cyber, and political security of countries from severe adversarial AI harm.¹¹² Cybersecurity is another prominent area of interest within the AI field, in what has been described as a “rapidly evolving cat-and-mouse game” between defensive and adversarial AI actors.¹¹³ AI-powered systems may (and with growing necessity) counter equally sophisticated hostile actors and potential acts of cyber-crime/cyber-terrorism.¹¹⁴ AI can function as a shield, a vital tool for assisting cybersecurity staff in detecting and preventing increasingly nuanced cyber-attacks, simultaneously exacerbated by AI-integration. Defensive AI tools can effectively reduce the attack surface of an organisation’s exposed IT databases and networks, may be deployed to continually detect system intrusions, irregularities, or malicious code, and can respond to anomalous activity instantaneously.¹¹⁵ The opportunity lies within AI’s pattern detection-based utility to: firstly, model system-user behaviours for any abnormalities and immediately lock-down systems stemming from identified discrepancies; secondly, deploy AI-powered antivirus software which may (unlike traditional scanning-based systems) utilise machine learning algorithms to instantaneously flag and quarantine malware-infected programs as well as to dynamically recognise evolving threats in a perpetual, self-reinforcing feedback-loop; and thirdly, implement autonomous network and email traffic scanning to provide ongoing AI-powered system insight and defensive detection of increasingly sophisticated phishing-based attacks.¹¹⁶

In addition to defensive cybersecurity benefits, AI has utility for specifically combatting cyber-terrorism. The adoption of AI-based disruptive technologies by both internet service providers (ISPs) and law enforcement agencies will become increasingly critical for

neutralising insurgent use of internet channels by extremists and insidious multimodal activities. AI monitoring of vast datasets via ISPs and online platforms for illegal content may mean greater efficacy for impeding online harm. Speech recognition and visual-base models may deliver an invaluable multilingual semantic “extraction and disambiguation” tool, which can expose entities associated with terrorism/extremism-related content online.¹¹⁷ AI could also enable moderators to efficiently assess online threats and their severity to deploy appropriate measures, including blocking and removing extremist propaganda and content.

Some reservations

Despite the aforementioned benefits afforded by AI, international researchers and New Zealand experts have voiced reservations tied to any precipitous societal uptake.¹¹⁸ One reservation concerns the risk of systematic bias and discrimination toward marginalised individuals or groups within either the algorithms used to support counter-terrorism practices or the ingested datasets that algorithms interact with.¹¹⁹ Errors in AI machine learning systems can systematically inflict serious societal harm when their judgments are routinely used to assign punitive consequences that are “disproportionately borne by groups of people who have historically experienced discrimination or who are already vulnerable.”¹²⁰ Occurrences of racial bias across AI facial recognition systems, frequently observed, are due to absences of diversity in training data and images on which systems were built.¹²¹ Certain ethnicities were disproportionately evaluated compared to others, with AI facial recognition identification accuracy for Caucasian males at 99% but merely 34% for females of darker skin tones.¹²² Moreover, over two-hundred cognitive biases had been identified throughout psychology and information science research, which may unconsciously distort formulation of AI machine learning models and software, and human-driven analysis.¹²³ One research paper signalled “skewness in conventional AI mechanisms” and cognitive bias leads to “algorithmic mishaps” and may inadvertently influence how human analysts collect and disseminate intelligence to policymakers.¹²⁴ A DPMC OIA release acknowledged that these systems “operate with uncertainty” in that false negatives and positives may inevitably permeate these systems, which may, in turn, propagate systematic bias and “perverse consequences.”¹²⁵ Reducing AI machine learning uncertainty may also manifestly increase bias across other data metrics, such as racial or language categories. AI has substantial limitations when handling and assessing more nuanced contextual information, such as irony or satire.

New Zealand experts argue biases infused into AI models are currently an “unavoidable conundrum.” At the psychosocial-level humans are fundamentally biased and until consequential programmatic biases can be accounted for in both algorithms and datasets, the consensus is that AI should not be without strict human, data, and policy oversight for acceptable use. If left unaccounted, however, algorithmic prejudices may propagate societal discrimination and consolidate human predispositions to the “detriment of rights and fundamental liberties.”¹²⁶

A second reservation concerns an adverse AI function creep and the potential for interpretive abuse, hypothetically leading societies down a path toward corruption, state overreach, and even autocratic totalitarianism.¹²⁷ More specifically a looming “AI-tocracy” that encroaches on human rights may arise through an insidious progression of functional creep if left insufficiently accountable.¹²⁸ Put simply, functional creep occurs when increasingly capable AI-systems move beyond its initial intended purposes toward other objectives, including general law enforcement or, more sinisterly, mass-public surveillance or social-credit type schemes.¹²⁹ For instance, New York City Police Department’s Domain Awareness System, which leveraged approximately 3,000 CCTV and ANPR (automatic number plate recognition) systems for counter-terrorism purposes, was rarely used for such objectives and was instead appropriated for other general law enforcement uses.¹³⁰ The risk for functional creep is real, particularly when leveraging vast public datasets which could be vital in amassing leads for criminal investigations, but possibly for “any number of beneficial social interventions.” Moreover, the potential for interpretative abuse is serious. Even with policy and legislative safeguards, AI will inevitably be privy to politically-charged subjective interpretation by current and future governments, including toward the highly ambiguous, political and securitised notion of terrorism itself. Collecting vast amounts of public data for subsequent AI or human analysis will pose inherent risks; even anonymisation protection measures are circumvented to link datasets to individuals.¹³¹ If wielded by an authoritarian regime, harnessing AI-powered facial recognition and automated mass-surveillance systems could have a considerably suppressive effect on social unrest, dissent and inflict oppressive political control.¹³²

From a purely security-focused and, perhaps Orwellian, standpoint, New Zealand experts suggest embedding increasingly accessible AI-powered analytical and surveillance systems presents an opportunity to foster an extremely safe society where citizens remain on best behaviour because they are aware they are under constant surveillance. However, there lies an unacceptable privacy-security trade-off, given social licence, personal liberties, and liberal democratic values should be preserved. Nonetheless, while AI “might not be a panacea to the disease of bad decision-making,” this is not a reason to reject it entirely.¹³³ Acknowledging AI’s limitations and understanding the ways AI can be abused or inadvertently appropriated to violate civil rights, enables New Zealanders an opportunity to develop decision-making systems drawing on “respective strengths of human and machine, without resurrecting old forms of discrimination in a new way.”¹³⁴

Conclusion

Over the preceding five years there has been widespread international media coverage, public debate, and scholarly attention surrounding AI and, on the one hand, its potential to pose considerable new threats across multiple domains and, on the other hand, its socio-economic and security opportunities for humanity. Policymakers and researchers

are yet to agree upon a single authoritative definition of AI, with some contending the notion is a complex amalgam of various narrow-set technologies. Distinct technological developments to date have included generative AI, machine learning, deep learning, natural language, and large language models, and encompass public platforms, such as OpenAI's ChatGPT or Google's Gemini. However, comprehensive general or even super artificial intelligence, which could, in theory, outperform humans across every intellectual category simultaneously, remains hypothetical.

Although some guidance in this area is beginning to emerge, New Zealand remains vulnerable to the impact of AI, including, specifically, on the problem of violent extremism and terrorism. Given the dearth of public awareness, a lack of proactive political leadership, and non-existent private-public sector collaborative research, there is a potential impending moral panic on AI development. There are fears that AI, unfettered, is being commercialised in the rushed pursuit of economic-productivity gains much faster than societies and political leadership can hope to comprehend and control from a national security standpoint.¹³⁵ The marketplace may adversely set norms before regulators, policymakers, and law enforcement can catch up. In the interim, New Zealand and the rest of the world are perhaps faced with the Wild West of AI commercialisation, including a vacuum of proactive guidance and mandated trust-by-design principles integrated into public GenAI platforms.¹³⁶ Meanwhile, there is no a consensus among international policymakers on the exact risks stemming from AI and how these ought to be prioritised by governments.¹³⁷ There is also no international convention on AI, which means any safeguards are implemented on a fragmented and domestic-level.¹³⁸

The window of opportunity for security preparedness in New Zealand is rapidly closing. While governments worldwide are attempting to posture themselves as AI leaders, these initiatives have not yet borne much fruit. There is currently a problematic lack of international consensus surrounding AI constraints, safety and security protocols, besides "in principle" support, and as one New Zealand expert suggests, "there needs to be more than merely going to meetings, shaking hands and signing declarations." New Zealand must ensure it is not self-isolated from future international partnerships, and multinational collaboration will be critical for ensuring an effective and unilateral counter-terrorism approach is developed. A key question which should be urgently considered by New Zealand policymakers concerns how the government may successfully garner social licence, civil trust, and drive the necessary political leadership to address AI challenges both domestically and collaboratively with international partners. As a matter of priority, New Zealand's political leadership must begin fostering open and transparent dialogue to initiate and sustain a meaningful national security conversation that raises public awareness and understanding on AI, so that New Zealanders will be well placed with respect to this fast-approaching and inevitable technological revolution.

This article sought to illuminate the urgent need for greater attention towards AI's terrorist risk and its counter-terrorism opportunities to inform that national security conversation on this otherwise nebulous and understudied topic. To that end, the article provides a definitional foundation on contemporary AI and advanced several key policy recommendations based on research findings and corresponding expert guidance. AI use is beginning to exponentially manifest overseas right now and New Zealand will, by no means, be immune to its global technological advancement. Policymakers must act now and maintain a watchful eye on these rapid international developments – as AI pioneer Alan Turing once envisaged: “We can only see a short distance ahead, but we can see plenty there that needs to be done.”¹³⁹

1 Amirhosein Toosi and Andrea Bottino, Babak Saboury, Eliot Siegel, Arman Rahmim. “A brief history of AI: how to prevent another winter (a critical review)”. *PET Clinics*, Volume 16, Issue 4. (October 2021). DOI: <https://doi.org/10.1016/j.cpet.2021.07.001> pp 449-469.

2 Arne Bewersdorff, *et al.* “Myths, mis- and preconceptions of artificial intelligence: A review of the literature.” *Computers and Education: Artificial Intelligence* 4 (2023): 100143. <https://www.sciencedirect.com/science/article/pii/S2666920X2300022X>

See also: Frank Emmert-Streib, Olli Yli-Harja, and Matthias Dehmer. “Artificial intelligence: A clarification of misconceptions, myths and desired status.” *Frontiers in artificial intelligence* 3 (2020): 524339.

3 Amirhosein Toosi, *et al.* “A brief history of AI: how to prevent another winter (a critical review)”. (2021). pp 449-469.

4 *ibid.*

5 Darrell M West and John R. Allen. “How artificial intelligence is transforming the world”. Brookings. (April 2024). <https://www.brookings.edu/articles/how-artificial-intelligence-is-transforming-the-world/>

See also: Peter Engelke. “AI, Society, and Governance: An Introduction.” Atlantic Council, 2020. <http://www.jstor.org/stable/resrep29327>. pp 1-25.

6 Pei Wang. “On defining artificial intelligence.” *Journal of Artificial General Intelligence* 10, no. 2 (2019). <https://sciendo.com/downloadpdf/journals/jagi/10/2/article-p1.pdf> pp 1-37.

See also: Haroon Sheikh, Corien Prins, and Erik Schrijvers. “Artificial intelligence: definition and background.” In *Mission AI: The new system technology*, pp. 15-41. Cham: Springer International Publishing, 2023.

7 *ibid.*

8 NVIDIA. “Large Language Models Explained”. Large Language Models. 2024. <https://www.nvidia.com/en-us/glossary/large-language-models/>

9 UNICRI & UNCCT. *Algorithms and Terrorism: The Malicious Use of Artificial Intelligence for Terrorist Purposes*. 2021. pp 13-16.

See also: Ajay Agrawal *et al.* “Artificial Intelligence: The Ambiguous Labor Market Impact of Automating Prediction.” *The Journal of Economic Perspectives* 33, no. 2 (2019): 31–50. <https://www.jstor.org/stable/26621238>.

10 Jeremy Baum and John Villasenor. “How close are we to AI that surpasses human intelligence?”. Brookings. July 2023. <https://www.brookings.edu/articles/how-close-are-we-to-ai-that-surpasses-human-intelligence/>

See also: Matt O’Shaughnessy. “How Hype Over AI Superintelligence Could Lead Policy Astray”. Carnegie Endowment for International Peace. September 2023. <https://carnegieendowment.org/posts/2023/09/how-hype-over-ai-superintelligence-could-lead-policy-astray?lang=en>

11 Andreas Holzinger, *et al.* “Toward human-level concept learning: Pattern benchmarking for AI algorithms.” *Patterns* (2023). <https://www.sciencedirect.com/science/article/pii/S2666389923001435>

12 Zhihan Lv. “Generative artificial intelligence in the metaverse era.” *Cognitive Robotics* 3 (2023). <https://www.sciencedirect.com/science/article/pii/S2667241323000198> pp 208-217

13 Jussi Jauhiainen. “The Metaverse: Innovations and generative AI.” *International Journal of Innovation Studies* 8, no. 3 (2024): 262-272. <https://www.sciencedirect.com/science/article/pii/S2096248724000183>

14 Gaudys L Sanclemente. “Reliability: Understanding cognitive human bias in artificial intelligence for national security and intelligence analysis.” *Security Journal* 35, no. 4 (2022): 1328-1348. <https://link.springer.com/article/10.1057/s41284-021-00321-2>

See also: Reza Montasari. “How States’ Recourse to Artificial Intelligence for National Security Purposes Threatens Our Most Fundamental Rights.” In *Artificial Intelligence and National Security*, Springer. (Swansea, UK 2022). <https://doi.org/10.1007/978-3-031-06709-9> pp 19-42.

15 Amirhosein Toosi, *et al.* “A brief history of AI: how to prevent another winter “. (2021). pp 449-469.

16 Grant Ferguson, *et al.* “Generating harms: Generative AI’s impact & paths forward.” Electronic Privacy Information Center (2023). <https://epic.org/documents/generating-harms-generative-ais-impact-paths-forward/>

See also: International Institute for Counter-Terrorism (ICT). “Future of Terrorism Technology - AI, Metaverse, and Deepfakes Panel”. Terrorism in the Eye of the Storm, the 21st World Summit on Counter-Terrorism. 2022. Retrieved from: https://www.youtube.com/watch?v=3pED8FMNZS4&ab_channel=InternationalInstituteforCounter-Terrorism%28ICT%29

17 *ibid.*

See also: Todd Helmus. “Artificial Intelligence, Deepfakes, and Disinformation: A Primer.” RAND Corporation, 2022. <http://www.jstor.org/stable/resrep42027>.

18 Mariarosaria Taddeo *et al.* “Artificial intelligence for national security: the predictability problem.” *Centre for Digital Ethics (CEDE) CETaS Research Reports* (September 2022). Retrieved from: <https://cetas.turing.ac.uk/publications/artificial-intelligence-national-security-predictability-problem> pp 9-40.

19 Haroon Sheikh, *et al.* “Artificial Intelligence: Definition and Background.” (2023). pp. 15-41.

20 Note: Research Ethics - This research presented no perceived risk of harm to prospective interview candidates. A low-risk ethics notification was submitted to the Massey University Ethics Committee prior to candidate engagement. No data sought, collected, or provided throughout this research was deemed confidential, privileged or classified.

21 Brendan Horsley. *Artificial intelligence frameworks and regulation: An intelligence perspective*. Inspector-General of Intelligence & Security. (Wellington: August 2024). https://igis.govt.nz/assets/Uploads/FINAL-Part-1_-Global-AI-frameworks-and-regulation.pdf. pp 2-17

22 *ibid.*

23 Ministry of Business, Innovation and Employment (MBIE). “Approach to work on Artificial Intelligence”. Hon Judith Collins KC. July 2024. <https://www.mbie.govt.nz/dmsdocument/28913-approach-to-work-on-artificial-intelligence-proactiverelase-pd>

24 Ministry of Business, Innovation and Employment (MBIE). “Approach to work on Artificial Intelligence”. (2024).

25 NZ Office of the Privacy Commissioner. *Artificial intelligence and the Information Privacy Principles*. September 2023. <https://www.privacy.org.nz/assets/New-order/Resources/Publications/Guidance-resources/AI-Guidance-Resources/AI-and-the-Information-Privacy-Principles.pdf>

26 Department of Internal Affairs (DIA). *Initial advice on Generative Artificial Intelligence in the public service*. DIA, NCSC & Stats NZ. July 2023. <https://www.digital.govt.nz/assets/Standards-guidance/Technology-and-architecture/Generative-AI/Joint-System-Leads-tactical-guidance-on-public-service-use-of-GenAI-September-2023.pdf>

27 Matt Boyd and Nick Wilson. “Catastrophic Risk from Rapid Developments in Artificial Intelligence: what is yet to be addressed and how might New Zealand policymakers respond?” *Policy Quarterly* 16, no. 1 (2020).

See also: Matt Boyd and Nick Wilson. “Rapid developments in Artificial Intelligence: how might the New Zealand government respond?” *Policy Quarterly* 13, no. 4 (2017).

28 *ibid.*

29 *ibid.*

30 *ibid.*

31 Sam Mulopulos. “An Artificial Intelligence Agenda for New Zealand: Lessons from the United States.” *Policy Quarterly* 19, no. 4 (2023): 65-72. <https://ojs.victoria.ac.nz/pq/article/view/8569> pp 65-72.

32 *ibid.* pp 65-72.

33 *ibid.*

See also: Fiona Seal. “Regulating Artificial Intelligence: A Critical Analysis of Technology Law’s Gordian Knot in the New Zealand Context.” University of Otago. 2021. pp 150-162 & 177-179.

See also: National Security Commission on Artificial Intelligence (NSCAI). “The Final Report”. 2024. <https://reports.nsc.ai.gov/final-report/>

34 Chris Vallance. “AI could worsen cyber-threats, report warns”. BBC News. 2023. <https://www.bbc.com/news/technology-67221117>

See also: Chris Vallance & Imran Rahman-Jones. “Urgent need for terrorism AI laws, warns think tank”. BBC News. January 4, 2024. <https://www.bbc.com/news/technology-67872767>

35 Helaine Leggat. “A new look at the Budapest Convention on Cybercrime”. ICTLC. 2024. <https://www.ictlc.com>

See also: Council of Europe. “The Convention on Cybercrime (Budapest Convention, ETS No. 185) and its Protocols”. 2024. <https://www.coe.int/en/web/cybercrime/the-budapest-convention>

See also: OECD Principles. “OECD AI Principles overview”. Policy Observatory. 2024. <https://oecd.ai/en/ai-principles>

36 US Department of Justice. “CLOUD Act Resources”. Criminal Division. 2024. <https://www.justice.gov/criminal/cloud-act-resources#:~:text=The%20CLOUD%20Act%20is%20designed,fight%20serious%20crime%20and%20terrorism>

Tim Cochrane. “Law Enforcement Cross-Border Data Sharing: A CLOUD Act Agreement for Aotearoa New Zealand?”. *New Zealand Law Review* 2021, no. 3 (2021): 401-434.

37 Ministry of Business, Innovation and Employment (MBIE). “Approach to work on Artificial Intelligence”. (2024).

38 *ibid.*

39 Brendan Horsley. *Artificial intelligence frameworks and regulation: An intelligence perspective*. Inspector-General of Intelligence & Security. (Wellington: August 2024). https://igis.govt.nz/assets/Uploads/FINAL-Part-1_-_Global-AI-frameworks-and-regulation.pdf pp 2-17

40 Ella Busch and Jacob Ware. *The Weaponisation of Deepfakes Digital Deception by the Far-Right*. International Centre for Counter-Terrorism (ICCT) Policy Brief. December 2023. DOI: 10.19165/2023.2.07 pp 1-11.

41 *ibid.*

See also: Janna Anderson and Lee Rainie. “As AI Spreads, Experts Predict the Best and Worst Changes in Digital Life by 2035”. Pew Research Center. June 2023. <https://www.pewresearch.org/>

See also: Arije Antinori. “Terrorism and deepfake: From hybrid warfare to post-truth warfare in a hybrid world.” In *ECLAIR 2019 European Conference on the Impact of Artificial Intelligence and Robotics*. 2019. pp 23-30.

42 Miles Brundage, *et al.* “The malicious use of artificial intelligence: Forecasting, prevention, and mitigation.” (2018). <https://doi.org/10.48550/arXiv.1802.07228> pp 1-12.

43 Gabriel Weimann, Alexander T. Pack, Rachel Sulciner, Joelle Scheinin, Gal Rapaport & David Diaz. “Generating Terror: The Risks of Generative AI Exploitation”, In *CTC Sentinel: Combating Terrorism Center at West Point. Issue 1; Vol 17* (January 2024). pp 17-24.

See also: UNICRI & UNCCT. *Algorithms and Terrorism: The Malicious Use of Artificial Intelligence for Terrorist Purposes*. (2021).

See also: Miron Lakomy. “Artificial intelligence as a terrorism enabler? understanding the potential impact of chatbots and image generators on online terrorist activities.” *Studies in Conflict & Terrorism* (2023). pp 1-21.

44 Darya Bazarkina. "Current and future threats of the malicious use of artificial intelligence by terrorists: psychological aspects." In *The Palgrave Handbook of Malicious Use of AI and Psychological Security*, pp. 251-272. Cham: Springer International Publishing, 2023.

45 Ella Busch and Jacob Ware. *The Weaponisation of Deepfakes Digital Deception by the Far-Right*. (2023).

See also: Milan Gandhi. *Terrorism, Extremism, Disinformation and Artificial Intelligence: A Primer for Policy Practitioners*. Institute for Strategic Dialogue (ISD). London, 2024. <https://www.isdglobal.org/> pp 8-37.

See also: Dr. Sarah Lohmann. "ChatGPT, Artificial Intelligence, and the Terrorist Toolbox" in *Emerging Technologies and Terrorism: An American Perspective*. Centre of Excellence Defence Against Terrorism. (Carlisle, PA: US Army War College Press, 2024).

See also: Maggie Engler. *Considerations of the Impacts of Generative AI on Online Terrorism and Extremism*. In *Global Internet Forum to Counter Terrorism*, 2023. pp. 5-6.

See also: Yaser Esmailzadeh and Ebrahim Motaghi. "International Terrorism and Social Threats of Artificial Intelligence." *Journal of Globalization Studies*, Vol. 15 No. 1. (2024). DOI:10.30884/jogs/2024.01.09. pp 168-179.

46 *ibid*.

47 Darya Bazarkina. "Current and future threats of the malicious use of artificial intelligence by terrorists: psychological aspects." (2023).

See also: Anne Eyre. "The Manchester Attack Support Group Programme: modelling a psychosocial response to collective trauma". *Journal of Psychosocial Studies*, 1-20. (2022).

48 *ibid*.

See also: Katarina Kertysova. "Artificial Intelligence and Disinformation: How AI Changes the Way Disinformation is Produced, Disseminated, and Can Be Countered", *Security and Human Rights* 29, 1-4 (2018). pp 55-81.

49 *ibid*.

50 David Wells. *The Next Paradigm-Shattering Threat? Right-Sizing the Potential Impacts of Generative AI on Terrorism*. Middle East Institute. Washington DC: 2024. <https://mei.edu/publications/next-paradigm-shattering-threat-right-sizing-potential-impacts-generative-ai-terrorism> pp 4-16.

51 Tech Against Terrorism. *Early terrorist experimentation with generative artificial intelligence services*. (2023).

52 *ibid*.

53 *ibid*.

See also: Wan Rosalili Wan Rosli. "Violent Extremism and Artificial Intelligence: A Double-Edged Sword in the Context of ASEAN." *The Commonwealth Cyber Journal*: 46. pp 46-55.

54 Gadi Eshed. *Is the Chatbot a Threat or an Opportunity for Security Organizations?* International Institute for Counter-Terrorism (ICT). June 2023. https://ict.org.il/wp-content/uploads/2023/06/Eshed_Chatbot-a-threat-or-opportunity_2023_06_20-1.pdf

55 Nina C Krickel-Choi. "State personhood and ontological security as a framework of existence: moving beyond identity, discovering sovereignty." *Cambridge Review of International Affairs* 37, no. 1 (2024): 3-21.

56 Peter Mantello, *et al*. "Automating extremism: Mapping the affective roles of artificial agents in online radicalization." In *The Palgrave handbook of malicious use of AI and psychological security*. (Springer International Publishing 2023). https://link.springer.com/chapter/10.1007/978-3-031-22552-9_4 pp 81-103.

57 Joe Burton. "Algorithmic extremism? The securitization of artificial intelligence (AI) and its impact on radicalism, polarization and political violence." *Technology in society* 75 (2023): 102262.

See also: Nicole Matejic and Chris Wilson. "Crimes of Influence: Generative Artificial Intelligence-led Crime as a Service." *The Commonwealth Cyber Journal*: 75.

58 Priyank Mathur, *et al*. "The Radicalization (and Counter-radicalization) Potential of Artificial Intelligence". International Centre for Counter-Terrorism (ICCT). May 2024. <https://icct.nl/>

59 Peter Mantello, *et al*. "Automating extremism: Mapping the affective roles of artificial agents in online radicalization." In *The Palgrave handbook of malicious use of AI and psy-*

chological security. (Springer International Publishing 2023). https://link.springer.com/chapter/10.1007/978-3-031-22552-9_4 pp 81-103.

See also: Mariam Shah. "The Digital Weaponry of Radicalisation: AI and the Recruitment Nexus". Global Network on Extremism & Technology (GNET). July 2024. <https://gnet-research.org/>

60 *ibid*

61 Peter Mantello, *et al.* "Automating extremism: Mapping the affective roles of artificial agents in online radicalization." (2023). pp 81-103.

62 *ibid*.

63 Mark Townsend. "AI poses national security threat, warns terror watchdog". The Guardian. June 2023. <https://www.theguardian.com/technology/2023/jun/04/ai-poses-national-security-threat-warns-terror-watchdog>

64 *ibid*.

65 Maria Zaccaro. "Jaswant Singh Chail: Man who took crossbow to 'kill Queen' jailed". BBC News. October 2023. <https://www.bbc.com/news/uk-england-berkshire-66113524>

66 Priyank Mathur, *et al.* "The Radicalization (and Counter-radicalization) Potential of Artificial Intelligence". (2024).

67 David Wells. *The Next Paradigm-Shattering Threat? Right-Sizing the Potential Impacts of Generative AI on Terrorism*. (2024). pp 4-16.

See also: Christopher Mouton, *et al.* "The Operational Risks of AI in Large-Scale Biological Attacks". RAND. October 2023. https://www.rand.org/pubs/research_reports/RRA2977-1.html

68 Yaser Esmailzadeh and Ebrahim Motaghi. "International Terrorism and Social Threats of Artificial Intelligence." (2024) pp 168-179.

69 Gabriel Weimann, *et al.* "Generating Terror: The Risks of Generative AI Exploitation". (2024).

70 *ibid*.

71 *ibid*.

72 Christopher Mouton, *et al.* "The Operational Risks of AI in Large-Scale Biological Attacks". (2023).

See also: Renan de Lima, *et al.* "Artificial intelligence challenges in the face of biological threats: emerging catastrophic risks for public health." *Frontiers in Artificial Intelligence* 7 (2024): 1382356.

73 *ibid*.

74 UNICRI & UNCCT. *Algorithms and Terrorism: The Malicious Use of Artificial Intelligence for Terrorist Purposes*. (2021). pp 26-48.

75 *ibid*.

76 *ibid*.

See also: Christina Schori Liang. "Terrorist Digitalis: Preventing Terrorists from Using Emerging Technologies." In *Global Terrorism Index 2023* (Sydney, March 2023): 72-74. Institute for Economics and Peace. pp. 72-74.

See also: Susan Sim, *et al.* *Emerging Technologies and Terrorism: An American Perspective*. (Carlisle, PA: US Army War College Press. 2024). pp 23-35.

77 Priyank Mathur, *et al.* "The Radicalization (and Counter-radicalization) Potential of Artificial Intelligence". (2024).

See also: Dr. John Battersby. "What's the future for terrorism?" *Line of Defence Magazine*, DEFSEC. October 2017. <https://defsec.net.nz/2017/10/01/whats-the-future-for-terrorism/>

78 *ibid*.

79 HYAS. "Blackmamba: Ai-Synthesized, Polymorphic Keylogger with On-The-Fly Program Modification". 2023. https://elhacker.info/Books/BOOKS%20PART%203/HYAS_AI_Attack-.pdf

See also: Andrew Wooden. "AI is making DDoS attacks 'more powerful, subversive, and frequent'". August 2024. <https://www.telecoms.com/security/ai-is-making-ddos-attacks-more-powerful-subversive-and-frequent->

See also: Kevin Williams. "'Cyber-physical attacks' fuelled by AI are a growing threat, experts say". CNBC. March 2024. <https://www.cnbc.com/2024/03/03/cyber-physical-attacks-fueled-by-ai-are-a-growing-threat-experts-say.html>

80 Alexander Blanchard and Jonathan Hall KC. "Terrorism and Autonomous Weapon Systems: Future Threat or Science Fiction?" (2023).

See also: James Johnson. "Artificial intelligence, drone swarming and escalation risks in future warfare." *The RUSI Journal* 165, no. 2 (2020). pp 26-36.

See also: Aleksander Ksawery Olech. "Unmanned Aerial Vehicle—a Lethal Weapon of Tomorrow for Terrorists." *Nowa Polityka Wschodnia* 32, no. 1 (2022): 44-60. <https://www.ceeol.com/search/article-detail?id=1048408>

81 National Cyber Security Centre. "The near-term impact of AI on the cyber threat". NCSC UK. (n.d.) <https://www.ncsc.gov.uk/report/impact-of-ai-on-cyber-threat>

See also: Department for Science, Innovation & Technology. "Safety and security risks of generative artificial intelligence to 2025 (Annex B)". UK Government. (October 2023). <https://www.gov.uk/>

82 HYAS. "Blackmamba: Ai-Synthesized, Polymorphic Keylogger with On-The-Fly Program Modification". 2023. <https://www.hyas.com/hubfs/Downloadable%20Content/HYAS-AI-Augmented-Cyber-Attack-WP-1.1.pdf>

83 Alexander Blanchard and Jonathan Hall KC. "Terrorism and Autonomous Weapon Systems: Future Threat or Science Fiction?" (2023).

84 UNICRI & UNCCT. *Algorithms and Terrorism: The Malicious Use of Artificial Intelligence for Terrorist Purposes*. (2021).

See also: Susan Sim, *et al.* *Emerging Technologies and Terrorism: An American Perspective*. (2024). pp 1-35.

See also: Kevin Peters. "21st century crime: How malicious artificial intelligence will impact homeland security." Monterey, CA; Naval Postgraduate School, (2019). pp 13-80.

85 Pranshu Verma. "These ISIS news anchors are AI fakes. Their propaganda is real." *The Washington Post*. May 2024. <https://www.washingtonpost.com/technology/2024/05/17/ai-isis-propaganda/>

86 Darya Bazarkina. "Current and future threats of the malicious use of artificial intelligence by terrorists: psychological aspects." (2023). pp 251-272.

87 Miles Brundage, *et al.* "The malicious use of artificial intelligence: Forecasting, prevention, and mitigation." (2018). pp 23-49.

88 Note: UNICRI & UNCCT.. *Algorithms and Terrorism: The Malicious Use of Artificial Intelligence for Terrorist Purposes*. (2021).

89 *ibid.*

90 Aleksander Ksawery Olech. "Unmanned Aerial Vehicle—a Lethal Weapon of Tomorrow for Terrorists." (2022). pp 44-60.

91 UNICRI & UNCCT. *Algorithms and Terrorism: The Malicious Use of Artificial Intelligence for Terrorist Purposes*. (2021). pp 26-49.

92 David Wells. *The Next Paradigm-Shattering Threat? Right-Sizing the Potential Impacts of Generative AI on Terrorism*. Middle East Institute. Washington DC: 2024. <https://mei.edu/publications/next-paradigm-shattering-threat-right-sizing-potential-impacts-generative-ai-terrorism> pp 4-16.

93 Darya Bazarkina. "Current and future threats of the malicious use of artificial intelligence by terrorists: psychological aspects." (2023).

94 Susan Sim, *et al.* *Emerging Technologies and Terrorism: An American Perspective*. (2024). pp 23-35.

95 Gabriel Weimann, *et al.* "Generating Terror: The Risks of Generative AI Exploitation". (2024).

96 UNICRI & UNCCT. *Algorithms and Terrorism: The Malicious Use of Artificial Intelligence for Terrorist Purposes*. (2021).

97 Gabriel Weimann, *et al.* "Generating Terror: The Risks of Generative AI Exploitation". (2024). pp 17-19.

98 Marie Schroeter. *Artificial Intelligence and Countering Violent Extremism: A Primer. Global Network on Extremism & Technology (GNET)*. International Centre for the Study of Radicalisation, King's College London. 2020. pp 1-40.

99 Kathleen McKendrick. *Artificial Intelligence Prediction and Counter-Terrorism*. (2019). pp 2-33.

100 *ibid.*

101 Ella Busch and Jacob Ware. *The Weaponisation of Deepfakes Digital Deception by the Far-Right*. International Centre for Counter-Terrorism (ICCT) Policy Brief. December 2023. DOI: 10.19165/2023.2.07 pp 1-11.

102 Aaron Davis. "Artificial Intelligence and the Fight Against International Terrorism." *American Intelligence Journal* 38, no. 2 (2021). <https://www.jstor.org/stable/27168700>. pp 63–73..

See also: Tosin Ige, *et al.* "Enhancing border security and countering terrorism through computer vision: A field of artificial intelligence." In *Proceedings of the Computational Methods in Systems and Software*, Cham: Springer International Publishing, 2022. pp 656-666.

103 *ibid.*

104 Andreas Kaplan. "Artificial Intelligence (AI): When Humans and Machines Might Have to Coexist." In *AI for Everyone? Critical Perspectives*, edited by Pieter Verdegem, 21–32. University of Westminster Press, 2021. <http://www.jstor.org/stable/j.ctv26qjjhj.4>. pp 21-28.

105 Alexander Babuta, *et al.* *Artificial Intelligence and UK National Security: Policy Considerations*. (2020). pp 7-16.

106 Bethan McKernan and Harry Davies. "'The machine did it coldly': Israel used AI to identify 37,000 Hamas targets". The Guardian. April 2024. <https://www.theguardian.com/international>

See also: Eleonore Pauwels. "Artificial Intelligence and Data Capture Technologies in Violence and Conflict Prevention: Opportunities and Challenges for the International Community." (2020). pp 1-17.

107 See: Darktrace. 2024. <https://darktrace.com/>

108 Rakibul Hasan Chowdhury. "AI-driven business analytics for operational efficiency." *World Journal of Advanced Engineering Technology and Sciences* 12, no. 2 (2024). pp 535-543.

See also: Wei Chen, *et al.* "Elevating Security Operations: The Role of AI-Driven Automation in Enhancing SOC Efficiency and Efficacy." *Journal of Artificial Intelligence and Machine Learning in Management* 8, no. 2 (2024)

See also: Mark Purdy, *et al.* "How AI Can Help Leaders Make Better Decisions Under Pressure". Harvard Business Review. October 2023. <https://hbr.org/2023/10/how-ai-can-help-leaders-make-better-decisions-under-pressure>

109 *ibid.*

110 Kathleen McKendrick. *Artificial Intelligence Prediction and Counter-Terrorism*. (2019). pp 27-33.

See also: Stefano Fantin, *et al.* "Purpose Limitation by Design as a Counter to Function Creep and System Insecurity in Police Artificial Intelligence." (UNICRI) (2020). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3679850

111 Rakibul Hasan Chowdhury. "AI-driven business analytics for operational efficiency." (2024). pp 535-543.

112 Alexander Babuta, *et al.* *Artificial Intelligence and UK National Security: Policy Considerations*. Royal United Services Institute (RUSI) for Defence and Security Studies. London, 2020. pp 7-16.

See also: Fahad Ali Khan, *et al.* "AI-Driven Counter-Terrorism: Enhancing Global Security Through Advanced Predictive Analytics." *IEEE Access* 11 (2023): 135864-135879.

113 Andrew J Lohn. "Hacking AI: A Primer for Policymakers on Machine Learning Cybersecurity". Center for Security and Emerging Technology. (December 2020). Doi: 10.51593/2020CA006

114 Serena Bianchi *et al.* "Artificial Intelligence to Counter Cyber-Terrorism." In *International Conference on Cybersecurity and Cybercrime*, vol. 10, pp 12-20. 2023.

See also: Ramanpreet Kaur, *et al.* "Artificial intelligence for cybersecurity: Literature review and future research directions." *Information Fusion* 97 (2023): 101804.

See also: Bindu Bala and Sunny Behal. "AI techniques for IoT-based DDoS attack detection: Taxonomies, comprehensive review and research challenges." *Computer science review* 52 (2024): 100631.

See also: Iqbal H Sarker, *et al.* "Ai-driven cybersecurity: an overview, security intelligence modelling and research directions." *SN Computer Science* 2, no. 3 (2021):173. <https://link.springer.com/article/10.1007/s42979-021-00557-0>

115 *ibid.*

116 *ibid.*

117 *ibid.*

118 James Black, *et al.* *Strategic competition in the age of AI: Emerging risks and opportunities from military use of artificial intelligence.* (2024).

See also: Ella Brownlie. "Encoding inequality: The case for greater regulation of artificial intelligence and automated decision-making in New Zealand." Victoria University of Wellington. *Rev.* 51 (2020): 1.

See also: Rostam J Neuwirth. "Prohibited artificial intelligence practices in the proposed EU artificial intelligence act (AIA)." *Computer Law & Security Review* 48 (2023): 105798.

119 *ibid.*

See also: Chiraag Bains. "The legal doctrine that will be key to preventing AI discrimination". Brookings. September 2024. <https://www.brookings.edu/>

See also: Alexander Blanchard, *et al.* "The ethics of artificial intelligence for intelligence analysis: a review of the key challenges with recommendations." *Digital Society* 2, no. 1 (2023): 12.

120 *ibid.*

See also: Douglas Yeung, *et al.* "Identifying Systemic Bias in the Acquisition of Machine Learning Decision Aids for Law Enforcement Applications." RAND Corporation, 2021. <http://www.jstor.org/stable/resrep29576>.

121 Maria Stefania Cataleta. "Humane Artificial Intelligence: The Fragility of Human Rights Facing AI." East-West Center, 2020. <http://www.jstor.org/stable/resrep25514>.

122 *ibid.*

123 *ibid.*

124 Gaudys L Sanclemente. "Reliability: Understanding cognitive human bias in artificial intelligence for national security and intelligence analysis." (2022). pp 1328-1348.

125 Department of the Prime Minister and Cabinet (DPMC). "Official Information Act request relating to security briefings related to AI". OIA Request-2023/24-0683. May 2024. <https://www.dPMC.govt.nz/sites/default/files/2024-06/dPMC-roia-oia-2023-24-0683.pdf>

126 Maria Stefania Cataleta. "Humane Artificial Intelligence: The Fragility of Human Rights Facing AI." (2020).

127 Martin Beraja, *et al.* "AI-tocracy." *The Quarterly Journal of Economics* 138, no. 3 (2023): 1349-1402. <https://academic.oup.com/qje/article/138/3/1349/7076890?login=false>

128 Edward Santow. "Can Artificial Intelligence Be Trusted with Our Human Rights?" *AQ: Australian Quarterly* 91, no. 4 (2020): 10-17. <https://www.jstor.org/stable/26931483>.

129 Stefano Fantin and Plixavra Vogiatzoglou. "Purpose Limitation By Design As A Counter To Function Creep And System Insecurity In Police Artificial Intelligence." (2020).

130 Kathleen McKendrick. *Artificial Intelligence Prediction and Counter-Terrorism.* (2019).

131 *ibid.*

132 Martin Beraja, *et al.* "AI-tocracy." (2023).

133 Edward Santow. "Can Artificial Intelligence Be Trusted with Our Human Rights?" (2020).

134 *Ibid.*

135 Kiran Stacey and Dan Milmo. "AI developing too fast for regulators to keep up, says Oliver Dowden". The Guardian. September 2023. <https://www.theguardian.com/international>

136 Christian Schnewlin. "How do you plan to create trust by design in artificial intelligence?" EY Switzerland. February 2024. https://www.ey.com/en_ch/ai/how-do-you-plan-to-create-trust-by-design-in-artificial-intelligence

137 Milan Gandhi. *Terrorism, Extremism, Disinformation and Artificial Intelligence: A Primer for Policy Practitioners.* (2024).

138 *ibid.*

139 The Alan Turing Institute. "What Alan Turing means to us". 2024. <https://www.turing.ac.uk/blog/what-alan-turing-means-us>