Te Pourewa Mātaki
Inspector-General of
Intelligence and Security

Artificial intelligence frameworks and regulation

An intelligence perspective

Brendan Horsley
Inspector-General of Intelligence & Security
August 2024

CONTENTS

INTRODUCTION

1. The IGIS 2023/24 Work Programme proposed a review of the intelligence and security agencies' use, or planned use, of artificial intelligence (AI) tools. This research paper sets out domestic and international AI regulation, frameworks, and policies, and publicly available information on how AI is being used by intelligence and security agencies. A subsequent report will examine the use of AI by the New Zealand Security Intelligence Service (NZSIS) and the Government Communications Security Bureau (GCSB).

RESEARCH OBJECTIVES AND CRITERIA

2. Research for this paper sought to identify domestic and international principles or models (developed or emerging) for the use of AI, at a high level and specifically for intelligence purposes.

3. Guiding questions were:

   3.1. What are AI and machine learning and how are they used?

   3.2. What international models, principles, frameworks, or legislation have been implemented (eg multilateral or country-specific) for the responsible use of AI, particularly by government agencies?

   3.3. What international models, principles, frameworks or legislation are under development (multilateral and country-specific)?

   3.4. Internationally, how is AI being incorporated into the work of intelligence and security agencies?

   3.5. What AI policies / procedures do international intelligence and security agencies have in place?

   3.6. What challenges have been identified for oversight of AI (broadly and for intelligence purposes)?

   3.7. What issues / concerns have been raised with intelligence and security agencies using AI?

   3.8. What principles / frameworks are there for the use of AI in New Zealand, particularly by government agencies?

WHAT IS ARTIFICIAL INTELLIGENCE?

4. Artificial intelligence is a field of computer science that seeks to create computer systems that can perform tasks, or generate outputs, that normally require a human to do.[1] For example: translate a language, solve a problem, or make a decision.

5. AI is an umbrella term comprising many sub-fields. The most prominent are:[2]

Definition	Explanation	Examples
Machine learning (ML)	Machine learning uses algorithms trained to learn from existing data and improve upon that data to make decisions or predictions. The algorithms extract patterns and learn implicit rules from the data.	Image recognition software, virtual assistants (eg Siri or Alexa)
Deep learning	A sub-field of machine learning where machines are trained to learn by example and constantly refine internal features to improve the outcome. This is similar to how a human brain learns and processes information.	Task automation, chat bots, biomedical science (eg cancer detection)
Generative AI (GenAI) or artificial generative intelligence (AGI)	Generative AI uses prompts or questions to generate text or images that (sometimes) resembles human-created content. The tools match user prompts to patterns in training data and probabilistically 'fill in the blank'.	ChatGPT, Google Gemini, AI image generators (eg Midjourney or Stable Diffusion)
Large language models (LLM)	Large language models are a subset of GenAI designed to process and generate human-like text, understanding context and nuance. They can perform a variety of text analysis tasks, for example summarising large volumes of text, or gauging sentiment from data.	GPT-4, Whisper, Google's PaLM

INTERNATIONAL FRAMEWORKS

6. To date international frameworks for the development and use of AI are mostly 'soft law', such as non-binding recommendations, multilateral agreements, or guidelines. Some commonly agreed-upon guardrails are in place until individual countries enact their own legislation, or are intended to operate alongside legislation.[3]

[1] See, eg Helen Toner "What Are Generative AI, Large Language Models, and Foundation Models?" (Center for Security and Emerging Technology, 12 May 2023) cset.georgetown.edu.

[2] See, eg above n 1, and Department of Internal Affairs, National Cyber Security Centre, and Statistics New Zealand *Initial advice on Generative Artificial Intelligence in the Public Service* (July 2023).

[3] David Leslie and others *Artificial intelligence, human rights, democracy, and the rule of law: a primer* (The Council of Europe, 2021).

7. Governance approaches can be divided into four categories:[4]

	Risk-based	Rules-based	Principles-based	Outcomes-based
Definition	Identifies and mitigates risks and/or harms AI systems may cause	Detailed and specific rules, standards, or requirements for AI systems	Basic principles and/or guidelines, leaving interpretation to individual organisations	Sets measurable outcomes without specifying actions required for compliance
Benefits	• Tailored to application • Proportional to the risk • Flexible to changing risk levels	• Reduces complexity • Consistent enforcement	• Fosters innovation • Adaptable to new technology • Encourages the sharing of best practices	• Supports efficiency • Flexible to change • Fosters innovation • Cost-effective compliance
Challenges	• Risk assessments are complex • creates barriers to market entry in high-risk areas • Assessment and enforcement can be complex	• Quite rigid • Increased compliance costs • Hard to enforce	• Potential inconsistencies with interpretation of principles • Impractical compliance and enforcement • Easy to misuse or abuse	• The measurable outcomes are vague • Limited accountability • Limited control over process and transparency
Examples	European Union AI Act (2024)[5]	China's Interim Measures of the Management of Generative AI Services (2023)[6]	Canada's Voluntary Code of Conduct for AI (2023)[7]	Japan's Governance Guidelines for Implementation of AI Principles (2022)[8]

The OECD AI Principles

8. In May 2019 the Recommendation of the Council on Artificial Intelligence (AI) was adopted by the OECD.[9] This was the first inter-governmental standard on AI. It aimed to foster innovation and trust in AI, promote the responsible stewardship of trustworthy AI, and respect human rights and democratic values. The use of AI would have to incorporate privacy, digital security, and risk management.

9. The Recommendation set out five principles for member states to adhere to and incorporate in domestic guidance and/or legislation:

[4] I note that there is not always a clear distinction between categories, or there may be an overlap between them. World Economic Forum *Generative AI Governance: Shaping a collective global future* (2024); International Association of Privacy Professionals (IAPP) *Global AI Legislation Tracker* (February 2024).

[5] Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence [2024] OJ L 2024/1689.

[6] See above n 4.

[7] Innovation, Science and Economic Development Canada *Voluntary Code of Conduct on the Responsible Development and Management of Advanced Generative AI Systems* (September 2023).

[8] Ministry of Economy, Trade, and Industry *Governance Guidelines for Implementation of AI Principles Version 1.1* (28 January 2022).

[9] Organisation for Economic Cooperation and Development (OECD) Recommendation of the Council on Artificial Intelligence (2019) OECD/LEGAL/0449.

- inclusive growth and sustainable well-being;
- human-centred values and well-being;
- transparency and explainability;
- robustness, security and safety; and
- accountability.

10. In 2023, the OECD reviewed members' implementation of AI legislation or guidance. It noted that uptake depended on the wealth of the nation, other priorities facing the nation, and the progression of other regulations that overlapped AI (such as privacy, algorithm charters, and data ethics).[10] The OECD noted New Zealand, Chile, and Singapore had included AI in a trade agreement, but otherwise did not comment on New Zealand's implementation.

The Group of 20 (G20)

11. In 2019, the G20 adopted five principles for the development, regulation, and responsible stewardship of AI, reaffirmed by all G20 members in 2023:[11]

- AI must be human-centred, and respect privacy, human rights, diversity, democratic values, and safeguards to protect society;
- AI must contribute to the growth and well-being for individuals, society, and the planet;
- AI must be transparent and explainable, it must be disclosed when used, and used responsibly so people know when they are engaging with AI;
- AI must be robust, secure, and safe, and risks must be constantly assessed and addressed; and
- organisations and individuals who develop, deploy and operate AI systems are accountable for the use of the tools.

The Group of Seven (G7)

12. In October 2023, the G7[12] agreed to a voluntary code of conduct for AI systems, building on the OECD Principles.[13] The code focuses on risk mitigation (eg pre-deployment risk assessments and mitigation), monitoring, assessing and reporting of incidents and any misuse of AI.

[10] OECD "AI Policy Observatory" oecd.ai; World Economic Forum *Generative AI Governance: Shaping a collective global future* (2024).

[11] Group of Twenty (G20) *AI Principles* (2019); Group of Twenty *G20 Ministerial Statement on Trade and Digital Economy* (2019).

[12] Canada, Germany, France, USA, Italy, Japan, the United Kingdom, and the European Union.

[13] *Hiroshima Process International Code of Conduct for Advanced AI Systems* (30 October 2023).

13. The code of conduct identified priorities for research and development, including content authentication; measures to protect data rights and mitigate societal, safety and security risks; and technical standards.

European Union (EU) AI Act

14. The European Union AI Act was passed on 13 March 2024 and entered into force on 2 April 2024.[14] Its provisions will be implemented over 24-36 months.[15] The Act sets out rules for developing and using AI in the European Union.

15. All AI systems and models must be used in an ethical, safe, and respectful manner, adhering to fundamental rights. The Act established a European AI Office, to develop tools for assessing the capabilities of general-purpose AI models, monitor the implementation of rules, identify emerging risks, investigate potential infringements, and support the enforcement of regulations on prohibited AI practices and high-risk systems.

16. Some AI tools are exempt from the EU AI Act, including those used for military and national security purposes.[16]

17. Regulation is based on the level of risk:[17]

    17.1. **Unacceptable risk:** AI tools in this category are prohibited, such as tools that manipulate human behaviour,[18] use real-time remote biometric identification (eg facial recognition) in public spaces, and those used for social scoring.[19]

    17.2. **High-risk:** These are applications that could pose significant threats to health, safety, or the fundamental rights of persons, eg those used in education, health, recruitment, law enforcement, and justice. Such tools are subject to obligations of quality, transparency, human supervision, and security. They must be evaluated before they are used, and during their life-cycle.

    17.3. **General-purpose AI:** These are tools trained on large amounts of data and capable of performing a wide range of tasks, including foundation models (eg ChatGPT) that are subject to transparency requirements.

    17.4. **Limited risk:** Limited risk tools are subject to transparency obligations aimed at informing users they are interacting with an AI system, and allowing users to exercise their choices. Examples include chatbots.

[14] Above n 5.

[15] "The Act" artificialintelligenceact.eu.

[16] Council of the EU *Artificial intelligence Act: Council and Parliament strike a deal on the first rules for AI in the world* (9 December 2023).

[17] "High-level summary of the Act" artificialintelligenceact.eu.

[18] This includes AI tools that use subliminal techniques and systems to exploit vulnerabilities in specific groups (eg age, ethnicity, or mental health status) in a manner that can affect decision-making. This can range from physical or psychological harm to individuals (eg encouraging eating disorders in youth) through to manipulating the way people vote, and undermining democracy.

[19] AI-based social scoring tools assess an individual's fraud risk when they apply for a government benefit or financial loan, based on socio-economic status and other personal information.

17.5. **Minimal risk:** These tools do not require regulation. Examples include spam-filters and AI video games.

The United Nations (UN)

18. In March 2024, the United Nations General Assembly unanimously adopted a non-binding resolution on AI.[20] The resolution calls for all member states to ensure "safe, secure, and trustworthy AI systems" are developed responsibly and in line with human rights and international law.

19. The resolution urges member states to:

19.1. stop using AI tools that do not comply with international human rights law, or pose undue risks to human rights;

19.2. cooperate and support developing nations, so they can benefit from inclusive and equitable access to new technology, digital literacy, and innovation;

19.3. utilise AI to meet the UN's Sustainable Development Goals; and

19.4. discuss AI governance and keep abreast of international best practices, including data governance.

AI LEGISLATION AND POLICY IN THE FIVE EYES NATIONS

Canada

20. In June 2022 the Artificial Intelligence and Data Act was tabled in the Canadian Parliament.[21] In early 2024 the Bill was at the committee stage. The Canadian Government aligned the Bill with the EU AI Act, taking a risk-based approach to the use of AI. The law would require high-impact AI systems to meet existing safety and human rights expectations and would prohibit irresponsible and malicious use of AI.[22]

21. Guidance and a code of practice for use of generative AI in the public sector are under development in Canada.[23] In the interim, AI regulation and guidance is found in:

21.1. The Directive on Automated Decision-Making, which requires algorithmic impact assessments to be completed before use of any automated decision-making systems, including AI, by the Canadian Government;[24]

21.2. The Voluntary Code of Conduct on the Responsible Development and Management of Advanced Generative AI Systems, under which signatories agree to principles including

[20] United Nations General Assembly *Seizing the opportunities of safe, secure and trustworthy artificial intelligence system for sustainable development* (11 March 2024) A/78/L.49.

[21] Government of Canada *The Artificial Intelligence Data Act (AIDA) - Companion document* (2023). The Act was tabled as part of Bill C-27, the Digital Charter Implementation Act, 2022.

[22] International Association of Privacy Professionals *Global AI Legislation Tracker* (February 2024).

[23] Group of Seven (G7) *G7 Hiroshima Process on Generative Artificial Intelligence* (2023).

[24] Government of Canada *Directive on Automated Decision-Making* (1 April 2019).

accountability, transparency, safety, fairness and equity, human oversight and monitoring, validity and robustness;[25] and

21.3. The Principles for the development, provision, and use of Generative AI systems, which sets out the privacy obligations for those using AI tools in Canada.[26] These include: legal consent for personal information to be used by AI; personal information can only be used for the appropriate purpose; necessity and proportionality must be established; it must be clear that AI uses personal information; and the risks of this must be identified. There must be accountability for compliance, limited collection of personal information, and individuals must be allowed access to their data.

Australia

22. Australia does not have any specific law or regulation for AI, which is governed by existing laws such as the Privacy Act 1988, the Data Availability and Transparency Act 2022, and the Consumer Data Right.[27]

United Kingdom

23. The United Kingdom does not have an AI law.[28] It was taking a "context-based, proportionate, and adaptable" approach to the regulation of AI, relying on existing legislation to act as guardrails, such as the Data Protection Act 2018, the Equality Act 2010, and the National Security and Investment Act 2021.[29] In mid-2024 the newly elected UK Government indicated an intention to develop legislation to place "requirements" on those developing the most powerful AI models.[30]

United States of America

24. In 2020 the Office of the Director of National Intelligence (ODNI) produced guidelines for the US intelligence community on how to procure, design, build, use, protect, consume, and manage AI in an ethical manner.[31] The framework provides a number of questions decision-makers must consider when acquiring or using an AI system. These cover:

24.1. documenting the purpose, parameters, limitations, and design of AI systems, and how this is balanced against acceptable risk;

24.2. the legal and policy considerations of using an AI tool, eg civil liberties, privacy, data collection and governance;

[25] Innovation, Science and Economic Development Canada *Voluntary Code of Conduct on the Responsible Development and Management of Advanced Generative AI Systems* (September 2023).

[26] The Canadian Office of the Privacy Commissioner *Principles for the development, provision, and use of Generative AI systems* (7 December 2023). "Use" includes collection, disclosure, deletion, or training an AI tool on personal information.

[27] Above n 22, at 4. The Consumer Data Right is enacted in the Competition and Consumer Act 2010.

[28] Above n 22, at 13.

[29] UK Government *AI Regulation White Paper* (March 2023).

[30] The King's Speech 2024, Prime Minister's Office, 10 Downing St and His Majesty King Charles III (17 July 2024).

[31] Office of the Director of National Intelligence *Artificial Intelligence ethics framework for the Intelligence Community* (June 2020).

24.3. when and how a human will be involved with the AI tool, including accountability of AI outputs and decisions, and how bias is mitigated to ensure objectivity of the tool, and that there is transparency, explainability and interpretability; and

24.4. periodic review of the tools.

25. On 30 October 2023 President Biden signed an Executive Order on AI.[32] The Order has eight guiding principles and priorities:[33]

25.1. Promote safe and secure AI by establishing robust evaluations of AI systems and mitigating risks before systems are deployed.

25.2. Promote responsible innovation, competition, and collaboration in AI.

25.3. AI development is built on the views of workers, labour unions, educators, and employers to support responsible uses of AI that improves workers' lives and positively augments human work.

25.4. Advance equity and civil rights by holding those developing and deploying AI accountable to standards that protect against harmful discrimination and abuse through technical evaluations, oversight, and regulation.

25.5. Enact safeguards against fraud, unintended bias, infringements on privacy and intellectual property rights, and other harms from AI in critical fields, eg healthcare and financial services.

25.6. Preserve privacy and civil liberties and ensure that the collection, use, and retention of data is lawful, secure, and promotes privacy to prevent AI-powered systems from exploiting or exposing personal data, while combating the broader legal and societal risks of improper collection and use of personal data.

25.7. Manage risks and increase internal government capacity to support responsible use of AI.

25.8. Develop a framework to manage the risks, unlock potential, and promote common approaches to shared challenges of AI.

26. The Order requires the National Security Council to develop AI guidelines for the US military, intelligence and national security community, to ensure they use AI safely, ethically, and effectively. It will also direct actions to counter the use of AI by adversaries.[34] Some parts of the US national security system (eg the FBI and the Department of Homeland Security) are exempt from the Executive Order, however.[35]

[32] The White House *Fact Sheet: President Biden issues Executive Order on safe, secure, and trustworthy artificial intelligence* (30 October 2023).

[33] *Executive Order [14110] on the safe, secure, and trustworthy development and use of artificial intelligence* (30 October 2023).

[34] At 4.8 (a)-(b).

[35] See Faiza Patel and Patrick C. Toomey "National security carve-outs undermine AI regulations" (21 December 2023) justsecurity.org.

27. In November 2023, two US House Representatives introduced a Bill (known as the "Five AIs Act") proposing to establish a Five Eyes "Strategic Artificial Intelligence Working Group".[36] The Bill was developed and announced without consultation with America's Five Eyes partners. It proposes that the US Secretary of Defense and Director of National Intelligence set up a working group to "develop and coordinate" an AI initiative within the Five Eyes network and members of the Five Eyes Intelligence Oversight and Review Council (FIORC).[37] The Bill's future is unclear.

New Zealand

28. New Zealand has no AI-specific legislation. Instead, AI falls within the scope of existing legislation, government strategies, and charters.

29. Existing legislation includes:

    29.1. The Privacy Act 2020, which sets out the Information Privacy Principles on how personal information can be collected, shared, and used.[38] The storage of personal information must have appropriate safeguards, and individuals can access and correct their information. Information must be accurate, complete, relevant, up to date, and not misleading.[39] Appropriate retention processes must be in place.[40] The Act applies to any personal information processed by AI.

    29.2. The New Zealand Bill of Rights Act 1990, which affirms rights potentially affected by government development or use of AI tools, including rights to freedom from discrimination and security against unreasonable search and seizure.

    29.3. The Human Rights Act 1993, which prohibits discrimination on the basis of age, sex, marital status, religious or ethical belief, colour, race, ethnic or national origins, disability, political opinion, employment status, family status and sexual orientation.

30. The New Zealand Government has developed a number of guidance documents, strategies and charters that apply to data, digital transformation, and algorithms:

    30.1. The Strategy for a Digital Public Service (2020)[41] and the Government Data Strategy and Roadmap 2021[42] call for the public sector to operate digitally to enhance efficiency and deliver outcomes meeting New Zealand's needs.

    30.2. The Algorithm Charter for Aotearoa New Zealand (2020) commits signatory government agencies to using algorithms in a consistent, transparent, and accountable manner and

[36] House of Representatives HR 6425 – 118th Congress (2023-24): To direct the Secretary of Defense to establish a working group to develop and coordinate an artificial intelligence initiative among the Five Eyes countries, and for other purposes; Senate S 4306 – 118th Congress (2023-24). See also Sydney Freedberg "AI for Five Eyes? New Bill pushes AI collaboration with UK, Australia, Canada, New Zealand" *Breaking Defense* (online ed, 22 November 2023).

[37] My office is a member of FIORC.

[38] Privacy Act 2020, s 22 Principles. I note that 2, 3, and 4(b) do not apply to personal information collected by intelligence and security agencies.

[39] Principles 5 – 8.

[40] Principles 9, 10 and 11(1)(g).

[41] *Rautaki mō tētahi Rāngai Kāwanatanga Matihiko - Strategy for a Digital Public Service* (Department of Internal Affairs, March 2020).

[42] New Zealand Government *Government Data Strategy and Roadmap 2021* (September 2021).

provides guidance on the assessment of any tool used to ensure it is ethical and legal.[43] Supporting the charter is the Algorithmic Impact Assessment toolkit.[44] The intelligence agencies are not signatories to the Algorithm Charter and I will explore their reasons for this in my review of their use of AI.

31. In July 2023, the New Zealand Government published interim guidance and advice on the use of generative AI tools in the public service.[45] It is intended to help agencies make informed decisions about using GenAI and balance the benefits and risks. The guidance covers security and privacy of information, eg not using data classified as SENSITIVE or above, and not putting personal information into public-facing tools.

32. In September 2023, the Privacy Commissioner published guidance on AI and the Information Privacy Principles.[46] It suggests a privacy impact assessment before a tool is used and that use of AI tools should consider Te Ao Māori perspectives on privacy.[47] The guidance identifies questions to ask when assessing a tool before use.

33. In November 2023, the National Cyber Security Centre joined agencies from 17 other countries endorsing the Guidelines for Secure AI System Development,[48] led by the UK National Cyber Security Centre.[49]

34. In June 2024, Cabinet agreed New Zealand would take a propionate and risked-based approach to AI regulation, using current legislation rather than developing a standalone AI Act.[50] New Zealand's approach to AI would be guided by the OECD AI Principles. The Government Chief Data Officer would encourage the public sector to adopt and use AI technologies responsibly, while managing the risks. The Department of Prime Minister and Cabinet would coordinate work across government on national security risks and opportunities associated with new technology, including AI.[51]

35. Independently of government, Te Mana Raraunga, the Maori Data Sovereignty Network, has developed the Māori Data Sovereignty Principles, values-based directives relevant to AI tools that process information relating to Maori people, language, culture, resources, or environments.[52]

[43] New Zealand Government *Algorithm charter for Aotearoa New Zealand* (July 2020).

[44] Available on data.govt.nz

[45] Department of Internal Affairs, National Cyber Security Centre, and Statistics New Zealand *Initial advice of the use of generative artificial intelligence in the public service* (July 2023).

[46] Office of the Privacy Commissioner *Artificial intelligence and the Information Privacy Principles* (September 2023).

[47] This may include the use of overseas AI systems that do not work accurately for Māori, the collection of Māori data and information that may be considered taonga, and the exclusion from consultation on processes and decisions regarding the use of AI that may impact upon Māori.

[48] UK National Cyber Security Centre *Guidelines for secure AI system development* (27 November 2023).

[49] Government Communications Security Bureau/NCSC "Joint Guidance: Guidelines for Secure AI System Development" (technical advisory, 28 November 2023).

[50] Cabinet Economic Policy Committee Minute of Decision "Approach and Work on Artificial Intelligence" ECO-24-MIN-0119 (26 June 2024); Cabinet paper "Approach to Work on Artificial Intelligence" (June 2024).

[51] Cabinet paper "Approach to Work on Artificial Intelligence" (June 2024) at [23].

[52] Te Mana Raraunga (Māori Data Sovereignty Network) *Principles of Māori Data Sovereignty Brief #1* (October 2018).

ARTIFICIAL INTELLIGENCE IN THE SECURITY AND INTELLIGENCE SECTOR

36. Some uses of AI by intelligence and security agencies have been disclosed by agencies or oversight bodies, or reported in academic or news media publications.

37. The Australian Inspector-General of Intelligence and Security, for example, has reported that the Australian intelligence community is “cautiously integrating” AI technologies into its analysis, primarily to analyse large, unstructured, and complex data sets.[53] The Australian Signals Directorate (ASD) has published an ethical framework on how it will integrate and use AI.[54]

38. The United States National Security Agency (NSA) has established an artificial intelligence security centre, which it describes as a “key part” of its cybersecurity mission.[55] In 2020 the CIA launched CIA Labs, where CIA staff collaborate with academia and industry on, amongst other things, artificial intelligence to support CIA activities.[56]

39. Artificial intelligence can enhance, or support, a number of intelligence and analysis disciplines and activities, but intelligence and security agencies are research and adopt AI to counter use of it by adversaries. Cyber criminals, for example, are reported to be utilising AI to conduct phishing or social engineering attacks.[57] State and non-state actors are reported to be using AI technologies to create and spread mis- and dis-information, hate speech, and conduct foreign interference (eg in elections).[58] Terrorist groups are reportedly using AI to recruit members, spread terrorist propaganda, and conduct attacks using autonomous weapons powered by AI.[59]

Counter-terrorism

40. MI5 uses AI to counter terrorism and violent extremism, for example:[60]

- AI assists analysts to assess whether someone watching violent and extremist videos poses a risk for radicalisation and committing a terrorist act;

- AI is used to detect violence (eg beheadings and torture) in videos so staff do not have to view the material; and

[53] Office of the Inspector-General of Intelligence and Security (Australia) *Preliminary Inquiry – Use of Artificial Intelligence by Intelligence Agencies* (Canberra, 29 May 2024).

[54] Australian Signals Directorate *Ethical Artificial Intelligence in the Australian Signals Directorate* (January 2023).

[55] “Artificial Intelligence Security Centre” on nsa.gov (accessed 5 August 2024).

[56] For more information, see “CIA Labs” on cia.gov.

[57] “FBI warns of Increasing Threat of Cyber Criminals Utilising Artificial Intelligence” (8 May 2024) on fbi.gov; Europol Innovation Lab *ChatGPT – The Impact of Large Language Models on Law Enforcement, a Tech Watch Flash Report* (Luxemburg, 2023) at 7-9.

[58] United Nations *Our Common Agenda Policy Brief 8: Information Integrity on Digital Platforms* (June 2023) at 19; The European External Action Service *2nd EEAS Report on Foreign Information Manipulation and Interference Threats: A Framework for Networked Defence* (January 2024) at 11.

[59] The United Nations Interregional Crime and Justice Institute (UNICRI) and the Cyber Security and New Technologies Unit of the United Nations Counter-Terrorism Centre (UNCCT) *Algorithms and Terrorism: The Malicious Use of Artificial Intelligence for Terrorist Purposes* (2021); Tech Against terrorism *Early terrorist experimentation with generative artificial intelligence services* (November 2023); Clarisa Nelu “Exploitation of Generative AI by Terrorist Groups” (10 June 2024) on icct.nl.

[60] Ken McCallum, Director General MI5 “Maths and the MI5: The calculations that keep the country safe” (speech, 30 June 2023) on mi5.gov.uk.

- machine learning models are used to translate and transcribe intercepted communications, such as phone calls.

41. The FBI reportedly uses AI to enable its assessment of threats, tips, and leads received from various sources, including social media, and electronic leads.[61] This includes using natural language processing models to review synopses of calls to tip lines, to check if anything important has been missed by the call taker. The AI is reportedly trained using the expertise of experienced call takers.

Cybersecurity

42. GCHQ, ASD, and the Canadian Security Establishment (CSE) have acknowledged using AI to combat malicious cybersecurity threats, with applications including:[62]

- identification of malicious software by analysing patterns of activity on networks and devices at scale, characterising criminal or hostile behaviour, updating their understanding of the threat and correlating it with reported activity elsewhere;

- defending against cyber-attacks by looking continually for patterns of complex behaviour, learning from the patterns to identify malicious activity such as website requests or suspicious emails (eg phishing campaigns); and

- characterising, analysing, and tracing malicious software to its origin, for attribution and take-down purposes.

Combating transnational organised crime

43. According to GCHQ, AI can be used to combat transnational organised crime activity by analysing large-scale and disparate data.[63] This involves mapping criminal networks and identifying those involved in crime; conducting analysis of financial data; and analysing data to provide geographical information on illicit activity.

Combating disinformation

44. Some countries are reported to use AI-enabled tools to develop and spread misinformation at scale and speed, including deepfake video and audio material.[64] To combat this, AI can be used to fact-check data against verified and trusted sources; detect deepfake videos or content; and identify and block social media bots used to spread disinformation.[65]

[61] Madison Alder "FBI's AI work includes 'Shark Tank'-style idea exploration, tip line use case" (5 June 2024) on fedscoop.com; Faiza Patel and Patrick Toomey "Bringing Transparency to National Security Uses of Artificial Intelligence" (4 April 2024) on justsecurity.org.

[62] GCHQ *Pioneering a New National Security: The Ethics of Artificial Intelligence* (24 February 2021); ASD, above n 54, at 6; Canadian Security Establishment *Annual Report 2023-24* (2024) at 49.

[63] GCHQ, above n 62.

[64] GCHQ, above n 62. See also Organisation for Economic Cooperation and Development *Initial policy considerations for generative artificial intelligence: OECD artificial intelligence papers* (September 2023) at 15.

[65] GCHQ, above n 62.

Data processing and augmented intelligence analysis

45. In the United States, the Department of Homeland Security has acknowledged that AI has the ability to process large amounts of data (eg bulk datasets) at speed, enabling analysts to derive insights from the data more quickly than is possible from manual review.[66] The CIA reportedly uses AI to read, review, monitor and analyse news media articles for geopolitical trends, and emerging crises in real-time.[67]

46. The NSA has noted that AI can be used to filter, flag, and triage data of interest.[68] As an example, AI could identify connections between disparate data and identify previously unknown connections between targets.[69]

47. AI could also be used to conduct sentiment analysis,[70] keyword matching, object detection from satellite images and video feeds,[71] and for identifying patterns or anomalies in large sets of data.[72] Such automation is often referred to as a "human-machine analysis team" and can reduce the administrative burden on analysts.[73]

48. The NSA is reported to use AI in signals intelligence to identify who is speaking and for speech-to-text transcription and translation.[74] Internally, NSA reportedly uses AI to enable compliance with legislation and policy controls and enable its audit unit.[75]

49. In late 2022, CSE deployed an automated translation tool which was developed in-house using machine learning.[76] CSE analysts can use the tool to translate content from more than 100 languages.

Risks

50. Risks identified with the use of AI can be split into the general and the security-specific.

[66] Department of Homeland Security *Privacy Impact Assessment for the DHS Data Analytics Tools* (13 June 2023) at 5.

[67] Patrick Turner "Spies like AI: The future of artificial intelligence for the US intelligence community" (27 January 2020) on defenseone.com; Frank Konkel "The CIA is taking a 'crawl, walk, run' approach to GenAI" (28 March 2024) on nextgov.com.

[68] Above n 52, at 7.

[69] GCHQ, above n 62. See also Alexander Babuta and others *Artificial Intelligence and UK National Security: Policy Considerations* (RUSI Occasional Paper, April 2020).

[70] Sentiment analysis is the automated analysis of digital text and other data to assess what emotion or attitude is being expressed (eg favourable, neutral, or negative). It is used widely for commercial purposes, eg to assess attitudes to companies or products from analysis of social media. The detection and analysis of emotion is still in development, and are not yet generally regarded as definitive or wholly reliable.

[71] R A Marcum "Rapid Broad Area Search and Detection of Chinese Surface-To-Air Missile Sites Using Deep Convolutional Neural Networks" (2017) Journal of Applied Remote Sensing 11(4*)*; Alexander Blanchard and Mariarosaria Taddeo "The Ethics of Artificial Intelligence for Intelligence Analysis: A Review of the Key Challenges with Recommendations" (April 2023) Digital Society 2(12).

[72] Blanchard and Taddeo, above n 71.

[73] Babuta, above n 69.

[74] Ronja Kneip "Another layer of opacity: How spies use AI and why we should talk about it" (20 December 2019) on aboutintel.eu; see also Turner, above n 67.

[75] Turner, above n 67.

[76] Canadian Security Establishment, above n 62, at 49.

General risks

50.1. **Transparency**: AI tools may not produce explainable results, due to intellectual property rights or the privacy of data used to train the tool.[77] It might not be apparent when action or analysis by an intelligence agency has been informed by an AI system.

50.2. **Bias:** AI tools sometimes provide biased and discriminatory outputs or decisions. Bias can be introduced in the development of an AI tool (eg from the developers' unconscious biases) and/or from the data the tool is trained on (eg poor quality, inaccurate, incomplete, or biased data).[78] If this occurs AI outputs can perpetuate social biases by discriminating based on gender,[79] ethnicity, and race.[80] If decision-makers rely on AI outputs without adequate recognition of assessment of this risk, the rights of affected people may be breached and biases further entrenched.[81]

50.3. **Liability/accountability**: Where people are adversely affected by state actions guided or determined by AI decisions or outputs and have a claim to redress, liability or accountability may have to be settled in circumstances where no human can be identified as directly responsible.[82] This might fall on proximate decision-makers who in fact had no influence on the operation of the algorithm – or it might be difficult to settle accountability at all.

50.4. **Privacy**: AI models are trained on large amounts of data, which may include personal information.[83] Big data analysis and the use of bulk datasets for machine learning typically requires repurposing data, contrary to the usual principle that data collection by the state should be for a named and specific purpose.[84] Data exploitation may be based on 'consent' for sharing that is pro forma rather than substantive (eg where users of applications have "accepted" lengthy terms of use with no realistic likelihood they have read and understood them).

[77] Jenna Burrell "How the Machine 'Thinks': Understanding Opacity in the Machine Learning Algorithms" (January 2016) Big Data and Society 3(1).

[78] DCAF - Geneva Centre for Security Governance *SSR Backgrounder Series: Intelligence oversight in the age of digitalization* (29 March 2024) at 6; Sandra Watcher and others "Bias Preservation in Machine Learning: The Legality of Fairness Metrics Under EU Non-Discrimination Law" (15 January 2021) 123 West Virginia Law Review 3; Céline Castets-Renard "Human Rights and Algorithmic Impact Assessment for Predictive Policing" in Hans W Micklitz and others (eds) *Constitutional Challenges in the Algorithmic Society* (Cambridge University Press, Cambridge, 2021) at 93-110; Leslie, above n 3, at 17.

[79] Roberto Iriondo "Amazon scraps secret AI recruiting engine that showed biases against women" (Carnegie Melon University Machine Learning Department, 11 October 2018) ml.cmu.edu.

[80] KM Vogel and others "The impact of AI on intelligence analysis: Tackling issues of collaboration, algorithmic transparency, accountability, and management" (2021) Intelligence and National Security 36(6); Blanchard and Taddeo, above n 71; Babuta, above n 69; David Leslie *Understanding artificial intelligence ethics and safety: A guide for the responsible design and implementation of AI systems in the public sector* (The Alan Turing Institute, 2019).

[81] Catarina Santos Botelho "The end of deception? – Counteracting algorithmic discrimination in the digital age" (2023) forthcoming in Giovanni De Gregorio and others (eds) *The Oxford Handbook on Digital Constitutionalism* (Oxford University Press, Oxford, 2024); Castets-Renard, above n 78.

[82] Jim Dempsey and Susan Landau "Challenging the machine: Contestability in government AI systems" (11 March 2024) on lawfaremedia.org; Kars Alfrink and others "Contestable AI by Design: Towards a Framework" (2023) Minds and Machines 33 at 613-639; and Leslie, above n 80.

[83] Office of the High Commissioner for Human Rights, United Nations *The right to privacy in the digital age: report (2021)* A/HRC/48/31.

[84] Alicia Solow-Niedeman "Information Privacy and the Interference Economy" (2022) 117 Northwestern University Law Review 357.

50.5. **Unreliability of results:** AI models can be unreliable and lack transparency as to how a result was achieved (often called the 'black box' effect). AI can create false or misleading information, and omit critical information. As AI is based on deep-learning models it can behave in unpredictable ways, and produce outcomes the tools' developers do not understand.[85]

50.6. **Human autonomy**: The ability to influence human choices, decisions, and perceptions by manipulation of information is a predominant concern with AI. When done for commercial purposes this has been dubbed "surveillance capitalism".[86] When done for purposes of political interference or social disruption it is of potentially even greater concern.[87]

Intelligence and security risks

51. Initiatives to guide or regulate AI there are often carve-outs for national security purposes, reducing applicability to intelligence and security agencies.[88] As noted earlier the European Union AI Act and US Executive Order on AI are examples.[89]

52. Risks identified with intelligence and security agencies using AI tools in their work include:

52.1. **Necessity and proportionality** of any collection of information guided by AI may be hard to determine. There is a risk of data being collected to improve the function of the AI rather than more directly to meet an intelligence need.[90] If predictive AI is used to distinguish suspicious from normal behaviour it is likely to require broad collection, including information generated by people of no interest to intelligence agencies.[91]

52.2. **Social licence**: Specific uses of AI for national security activities are likely to be classified, making it hard for intelligence and security agencies (and oversight) to reassure the public about how AI is used, and that the rights and privacy of citizens are upheld.[92]

52.3. **Proprietary information contained in AI tools:** Intelligence and security agencies are not immune to the 'black box effect' as, for example, commercial tools may contain proprietary information on how AI is used which will not be disclosed to the agencies.[93]

52.4. **Suitability of tools:** Some AI tools acceptable for everyday purposes may miss nuance important in intelligence and security – eg AI translation tools may miss subtleties or

[85] Pegah Maham and Sabrina Küspert *Governing General Purpose AI: A Comprehensive Map of Unreliability, Misuse and Systemic Risks* (Stiftung Neue Verantwortung, July 2023) at 3, 14, 18, 21; Melissa Heikkilä "Nobody knows how AI works" (MIT Technology Review, 5 March 2024) on technologyreview.com.

[86] Shoshana Zuboff "Big other: Surveillance Capitalism and the Prospects of an Information Civilization" (2015) Journal of Information Technology 30(1).

[87] Above n 3, at 16.

[88] Above n 35.

[89] The European Union exempted national security purposes form the AI Act as "national security remains the sole responsibility of Member States in accordance with Article 4(2) TEU [Treaty on European Union] and by the specific operational needs of national security activities and specific rules" of Member States (Artificial Intelligence Act, s 24).

[90] Blanchard and Taddeo, above n 71.

[91] Kathleen McKendrick *Artificial intelligence Prediction and Counterterrorism* (Chatham House, August 2019).

[92] Alexander Babuta "A New Generation of Intelligence: National Security and Surveillance in the Age of AI" (19 February 2019) on rusi.org.

[93] Bram Vaassen "AI, Opacity, and Personal Autonomy" (2022) Philosophy and Technology 35(88).

inflections in language (such as sarcasm or humour) that a human translator could identify and factor in to the assessment of meaning.[94]

OVERSIGHT OF INTELLIGENCE AND SECURITY AGENCY USE OF AI

Oversight reviews

53. In May 2024 the Australian Inspector-General of Intelligence and Security released an unclassified summary of its review into the use of AI by the Australian intelligence community.[95] The review found that the maturity and use of AI varied across the six agencies it oversees. The policies and strategies employed by the agencies were driven by ethical, compliance, and legal considerations, including human rights. Some agencies had implemented AI governance boards. Use of AI in the agencies reviewed was primarily focused on "enhancing operational efficiency through enabling or enhancing human analysts" and AI was "not being used for autonomous action or decision making."[96] The Australian intelligence and security agencies planned to use AI more deeply, and had strategic frameworks and ethical guidelines in place.

54. While not identifying any legality, propriety, or human rights concerns, the Australian IGIS made three recommendations to the agencies:

- continue to refine and adapt governance frameworks, policies and approval processes for the use of AI and the data that enables it;
- continue to enhance the transparency and auditability of AI systems; and
- continue to engage with Australian Government policy makers to ensure legislation and other frameworks governing the use of AI, remain applicable to, or do not preclude, the activities of the Australian intelligence community.

Challenges for oversight

55. A working group of the Five Eyes Intelligence Oversight and Review Council (FIORC) identified challenges for overseeing the use of AI by intelligence and security agencies including:[97]

55.1. **Acquiring expertise in AI**: Oversight staff need to upskill in AI, and there is currently limited training for the oversight of AI.[98] Recruiting skilled people is difficult as most countries have a skill shortage in AI (particularly in the public sector).[99] There is

[94] Célia Tavares and others "Artificial intelligence: A blessing or a threat to language service providers in Portugal" (20 October 2023) Informatics 10(4); Gabriel Nicholas and Aliya Bhatia "Lost in translation: Large language models in non-English context analysis" (Centre for Democracy and Technology, May 2023) at 25-26.

[95] Above n 53.

[96] Above n 53, at 6.

[97] FIORC AI Working Group *AI Paper* (September 2023).

[98] DCAF - Geneva Centre for Security Governance, above n 78, at 5.

[99] World Economic Forum *Jobs of Tomorrow: Large Language Models and Jobs* (2023); the National Security Commission on Artificial Intelligence *2021 Final Report* (2021).

competition with the private sector and academia, and those with the relevant skills expect high wages, flexible work arrangements, or cannot gain a security clearance.[100]

55.2. **Identifying applicable standards**: Among the Five Eyes countries there are varying levels of legislative control and guidance (ie 'soft' controls). In the absence of hard law and where there are exceptions for intelligence and security agencies from soft controls, the appropriate standards for assessing agency activity are unclear and contestable.

55.3. **Achieving transparency and accountability**: Independent oversight of AI tools can only occur if the oversight body is aware of the tools in use. If an intelligence or security agency takes action, or makes decisions, based on AI outputs, it may be difficult for the agency to explain how the AI tool came to that determination and justify its actions. Currently, there is no requirement to identify intelligence analysis that is derived from AI. This could potentially lead to issues if the intelligence is flawed, and an agency takes action or uses the intelligence. If a New Zealand intelligence and security agency used an AI tool supplied by a Five Eyes partner agency, there might be limited understanding within the New Zealand agency of the tool and how it works (ie the 'black box effect').

[100] Christopher Moran and others "The US Intelligence Community, Global Security, and AI: From Secret Intelligence to Smart Spying" (2023) Journal of Global Security Studies 8(2).

BIBLIOGRAPHY

Legislation and other legal instruments

New Zealand

Human Rights Act 1993.

New Zealand Bill of Rights Act 1990.

Privacy Act 2020.

Australia

Data Availability and Transparency Act 2022.

Privacy Act 1988.

Treasury Laws Amendment (Consumer Data Right) Act 2019.

Canada

Bill C-27 Digital Charter Implementation Act 2022.

United Kingdom

Data Protection Act 2018.

Equality Act 2010.

National Security and Investment Act 2021.

United States

House of Representatives HR 6425 – 118th Congress (2023-24): To direct the Secretary of Defense to establish a working group to develop and coordinate an artificial intelligence initiative among the Five Eyes countries, and for other purposes.

Senate S 4306 – 118th Congress (2023-24). A bill to direct the Secretary of Defense to establish a working group to develop and coordinate an artificial intelligence initiative among the Five Eyes countries, and for other purposes (the “Five AIs Act 2024”).

Organisation for Economic Cooperation and Development

Recommendation of the Council on Artificial Intelligence (2019) OECD/LEGAL/0449.

European Union

Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence [2024] OJ L 2024/1689.

United Nations

Seizing the opportunities of safe, secure and trustworthy artificial intelligence systems for sustainable development: draft resolution A/78/L.49 (11 March 2024).

Government materials

New Zealand

Cabinet Economic Policy Committee Minute of Decision "Approach and Work on Artificial Intelligence" ECO-24-MIN-0119 (26 June 2024) on mbie.govt.nz.

Cabinet paper "Approach to Work on Artificial Intelligence" (June 2024) on mbie.govt.nz.

Department of Internal Affairs *Rautaki mō tētahi Rāngai Kāwanatanga Matihiko - Strategy for a Digital Public Service* (March 2020) digital.govt.nz

Department of Internal Affairs, National Cyber Security Centre, and Statistics New Zealand *Initial advice on Generative Artificial Intelligence in the Public Service* (July 2023).

Department of Internal Affairs, National Cyber Security Centre, and Statistics New Zealand *Generative Artificial Intelligence system leaders' guidance for use of gen-AI across the New Zealand public service* (Summary, September 2023).

Government Communications Security Bureau/NCSC "Joint Guidance: Guidelines for Secure AI System Development" (technical advisory, 28 November 2023).

New Zealand Government *Government Data Strategy and Roadmap 2021* (September 2021).

New Zealand Government *Algorithm charter for Aotearoa New Zealand* (July 2020).

Inspector-General of Intelligence and Security *IGIS 2023/24 Work Programme* (June 2023).

Office of the Privacy Commissioner *Artificial intelligence and the Information Privacy Principles* (September 2023).

Statistics New Zealand *Algorithm Assessment Report* (October 2018).

Statistics New Zealand "Algorithmic Impact Assessment toolkit" on data.govt.nz

Australia

Australia's Chief Scientist *Rapid response information report: Generative AI: Language models and multimodal foundation models* (24 March 2023).

Australian Signals Directorate *Ethical Artificial Intelligence in the Australian Signals Directorate* (January 2023).

Office of the Inspector-General of Intelligence and Security *Preliminary Inquiry – Use of Artificial Intelligence by Intelligence Agencies* (29 May 2024).

Department of Industry, Science and Resources *Safe and responsible AI in Australia* (June 2023).

Canada

Canadian Security Establishment *Annual Report 2023-24* (2024).

Government of Canada *The Artificial Intelligence Data Act (AIDA) - Companion document* (2023).

Government of Canada *Directive on Automated Decision-Making* (1 April 2019).

Innovation, Science and Economic Development Canada *Voluntary Code of Conduct on the Responsible Development and Management of Advanced Generative AI Systems* (September 2023).

The Canadian Office of the Privacy Commissioner *Principles for the development, provision, and use of Generative AI systems* (7 December 2023).

Japan

Ministry of Economy, Trade, and Industry *Governance Guidelines for Implementation of AI Principles Version 1.1* (28 January 2022).

United Kingdom

Department for Science, Innovation and Technology *A pro-innovation approach to AI regulation* (29 March 2023).

Department for Science, Innovation and Technology *Safety and security risk of generative artificial intelligence to 2025 (Annex B)* (25 October 2023).

UK National Cyber Security Centre *Guidelines for secure AI system development* (27 November 2023).

GCHQ *Pioneering a New National Security: The Ethics of Artificial Intelligence* (24 February 2021).

United States of America

Department of Defense *Data, Analytics, and Artificial Intelligence Adoption Strategy Accelerating Decision Advantage* (November 2023).

Department of Homeland Security *Privacy Impact Assessment for the DHS Data Analytics Tools* (13 June 2023).

National Security Agency *The Next Wave Machine Learning 1* (2018) 22 1.

National Security Agency *The Next Wave Machine Learning 2* (2019) 22 2.

National Security Commission on Artificial Intelligence *NSCAI Final Report* (2021).

Office of the Director of National Intelligence *Artificial Intelligence ethics framework for the Intelligence Community* (June 2020).

Executive Order [14110] on the safe, secure, and trustworthy development and use of artificial intelligence (30 October 2023).

The White House *Fact Sheet: President Biden issues Executive Order on safe, secure, and trustworthy artificial intelligence* (30 October 2023).

Group of Seven (G7)

Hiroshima Process International Code of Conduct for Advanced AI Systems (30 October 2023).

Group of Twenty (G20)

AI Principles (2019).

G20 Ministerial Statement on Trade and Digital Economy (2019).

Papers and journal articles

Kars Alfrink and others "Contestable AI by design: Towards a Framework" (2023) Minds and Machines 33, 613-639.

Alexander Blanchard and Mariarosaria Taddeo "The Ethics of Artificial Intelligence for Intelligence Analysis: A Review of the Key Challenges with Recommendations" (April 2023) Digital Society 2(12).

Jenna Burrell "How the Machine 'Thinks': Understanding Opacity in the Machine Learning Algorithms" (January 2016) Big Data and Society 3(1).

Joe Devanny and others "Generative AI and Intelligence Assessment" (30 November 2023) The RUSI Journal 168(7) 16-25.

R A Marcum "Rapid Broad Area Search and Detection of Chinese Surface-To-Air Missile Sites Using Deep Convolutional Neural Networks" (2017) Journal of Applied Remote Sensing 11(4).

Christopher Moran and others "The US Intelligence Community, Global Security, and AI: From Secret Intelligence to Smart Spying" (2023) Journal of Global Security Studies 8(2).

Alicia Solow-Niedeman "Information Privacy and the Interference Economy" (2022) 117 Northwestern University Law Review 357.

Karaitiana Taiuru "Treaty Of Waitangi/Te Tiriti and Māori Ethics Guidelines for: AI, Algorithms, Data and IOT" (2020) retrieved from Taiuru.maori.nz.

Célia Tavares and others "Artificial Intelligence: A Blessing or a Threat to Language Service Providers in Portugal" (20 October 2023) Informatics 10(4).

Bram Vaassen "AI, Opacity, and Personal Autonomy" (2022) Philosophy and Technology 35(88).

K M Vogel and others "The Impact of AI on Intelligence Analysis: Tackling Issues of Collaboration, Algorithmic Transparency, Accountability, and Management" (2021) Intelligence and National Security 36(6).

Sandra Watcher and others "Bias Preservation in Machine Learning: The Legality of Fairness Metrics Under EU Non-Discrimination Law" (15 January 2021) 123 West Virginia Law Review 3.

Shoshana Zuboff "Big other: Surveillance Capitalism and the Prospects of an Information Civilization" (2015) Journal of Information Technology 30(1).

Chapters in books

Catarina Santos Botelho "The End of Deception? – Counteracting Algorithmic Discrimination in the Digital Age" (2023) forthcoming in Giovanni De Gregorio and others (eds) *The Oxford Handbook on Digital Constitutionalism* (Oxford University Press, Oxford, 2024).

Céline Castets-Renard "Human Rights and Algorithmic Impact Assessment for Predictive Policing" in Hans W Micklitz and others (eds) *Constitutional Challenges in the Algorithmic Society* (Cambridge University Press, Cambridge, 2021).

Reports

Artificial Intelligence Forum New Zealand *Artificial Intelligence: Shaping a Future New Zealand* (2018).

Artificial Intelligence Researchers Association *White Paper: Aotearoa New Zealand Artificial Intelligence - A Strategic Approach* (November 2021).

Alexander Babuta and others *Artificial Intelligence and UK National Security: Policy Considerations* (RUSI Occasional Paper, April 2020).

Anthony Burke *Robust artificial intelligence for active cyber defence* (Alan Turing Institute, March 2020).

DCAF - Geneva Centre for Security Governance *SSR Backgrounder Series: Intelligence oversight in the age of digitalization* (29 March 2024).

The European External Action Service *2nd EEAS Report on Foreign Information Manipulation and Interference Threats: A Framework for Networked Defence* (January 2024).

Europol Innovation Lab*: ChatGPT - The Impact of Large Language Models on Law Enforcement, a Tech Watch Flash Report* (Luxembourg, 2023).

International Association of Privacy Professionals (IAPP) *Global AI Legislation Tracker* (February 2024).

T Kukutai and others *Maori Data Governance Model* (Te Kāhui Raraunga, 2023).

David Leslie *Understanding artificial intelligence ethics and safety: A guide for the responsible design and implementation of AI systems in the public sector* (The Alan Turing Institute, 2019).

David Leslie and others *Artificial intelligence, human rights, democracy, and the rule of law: a primer* (The Council of Europe, 2021).

Pegah Maham and Sabrina Küspert *Governing General Purpose AI: A Comprehensive Map of Unreliability, Misuse and Systemic Risks* (Stiftung Neue Verantwortung, July 2023).

Kathleen McKendrick *Artificial Intelligence: Prediction and Counter-Terrorism* (Chatham House, 2019).

New Zealand Law Foundation and Otago University *Government Use of Artificial Intelligence in New Zealand* (2019).

Gabriel Nicholas and Aliya Bhatia *Lost in Translation: Large Language Models in Non-English Context Analysis* (Centre for Democracy and Technology, May 2023).

Organisation for Economic Cooperation and Development *Initial policy considerations for generative artificial intelligence: OECD artificial intelligence papers* (September 2023).

Corin Stone *The Integration of Artificial Intelligence in the Intelligence Community: Necessary Steps to Scale Efforts and Speed Progress* (Joint PIJIP/TLS Research Paper Series 73, 2021).

Mariarosaria Taddeo and others *Artificial Intelligence for National Security: The Predictability Problem* (Centre for Digital Ethics, September 2022).

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

Te Kotahi Research Institute *Māori Perspectives on Trust and ADM* (2020).

Te Mana Raraunga Māori Data Sovereignty Network *Principles of Māori Data Sovereignty Brief #1* (October 2018).

Te Mana Raraunga Māori Data Sovereignty Network *Submission on the Draft Algorithm Charter* (2020).

Office of the High Commissioner for Human Rights, United Nations *The right to privacy in the digital age: report (2021)* A/HRC/48/31.

The United Nations *Our Common Agenda Policy Brief 8: Information Integrity on Digital Platforms* (June 2023).

The United Nations Interregional Crime and Justice Institute (UNICRI) and the Cyber Security and New Technologies Unit of the United Nations Counter-Terrorism Centre (UNCCT) *Algorithms and Terrorism: The Malicious Use of Artificial Intelligence for Terrorist Purposes* (2021).

World Economic Forum *Generative AI Governance: Shaping a collective global future* (2024).

World Economic Forum *Jobs of Tomorrow: Large Language Models and Jobs* (2023).

World Economic Forum *Reimagining Regulation for the Age of AI: New Zealand pilot Project* (June 2020).

News media

Andrew Burt “Ethical frameworks for AI aren’t enough” (9 November 2020) *Harvard Business Review*.

Sydney Freedberg “AI for Five Eyes? New Bill pushes AI collaboration with UK, Australia, Canada, New Zealand” *Breaking Defense* (online ed, 22 November 2023).

Joe McKendrick “AI adoption skyrocketed over the last 18 months” (27 September 2021) *Harvard Business Review*.

Alexandra Sims “NZ Police are using AI to catch criminals – but the law urgently needs to catch up too” (14 October 2021) *Radio New Zealand*.

Other internet resources

“The Act” on artificialintelligenceact.eu.

“AI Policy Observatory” on oecd.ai.

Madison Alder “FBI’s AI work includes ‘Shark Tank’-style idea exploration, tip line use case” (5 June 2024) on fedscoop.com.

"Artificial Intelligence Security Centre" on nsa.gov.

Alexander Babuta "A New Generation of Intelligence: National Security and Surveillance in the Age of AI" (19 February 2019) on rusi.org.

"CIA Labs" on cia.gov.

Council of the EU "Artificial intelligence Act: Council and Parliament strike a deal on the first rules for AI in the world" (press release, 9 December 2023).

Jim Dempsey and Susan Landau "Challenging the machine: Contestability in government AI systems" (11 March 2024) on lawfaremedia.org.

"FBI warns of Increasing Threat of Cyber Criminals Utilising Artificial Intelligence" (8 May 2024) on fbi.gov

Melissa Heikkilä "Nobody knows how AI works" (MIT Technology Review, Massachusetts Institute of Technology, 5 March 2024) on technologyreview.com.

"High-level summary of the Act" on artificialintelligenceact.eu.

Intelligence and National Security Alliance National Symposium "Gaining the AI advantage: Strategic Approach to Big Data" (recording, 2022) on youtube.com.

Roberto Iriondo "Amazon scraps secret AI recruiting engine that showed biases against women" (11 October 2018) Carnegie Mellon University Machine Learning Department on ml.cmu.edu.

The King's Speech 2024, Prime Minister's Office, 10 Downing St and His Majesty King Charles III (17 July 2024)

Ronja Kneip "Another layer of opacity: How spies use AI and why we should talk about it" (20 December 2019) on aboutintel.eu.

Frank Konkel "The CIA is taking a 'crawl, walk, run' approach to GenAI" (28 March 2024) on nextgov.com.

Ken McCallum "Maths and the MI5: The calculations that keep the country safe" (speech, 30 June 2023) on mi5.gov.uk.

Daragh Murray and Pete Fussey "GCHQ's ethical approach to AI: An initial human rights-based response" (5 March 2021) on aboutintel.eu.

Clarisa Nelu "Exploitation of Generative AI by Terrorist Groups" (10 June 2024) on icct.nl.

Faiza Patel and Patrick C. Toomey "National security carve-outs undermine AI regulations" (21 December 2023) on justsecurity.org.

Faiza Patel and Patrick C. Toomey "Bringing transparency to national security uses of artificial intelligence" (4 April 2024) on justsecurity.org.

Helen Toner "What Are Generative AI, Large Language Models, and Foundation Models?" (Center for Security and Emerging Technology, 2023) on cset.georgetown.edu.

Patrick Turner "Spies like AI: The future of artificial intelligence for the US intelligence community" (27 January 2020) on defenseone.com.