Developing an AI Capability Framework for the Trilateral Security Dialogue (TSD): US, Australia, and Japan

Aiden Warren, Ph.D.
Charles T. Hunt, Ph.D.
Matthew Warren, Ph.D.
Adam Bartley, Ph.D.
Mark Bryan Manantan

July 2024

Developing an AI Capability Framework for the Trilateral Security Dialogue (TSD): US, Australia, and Japan

Aiden Warren, Ph.D.
Charles T. Hunt, Ph.D.
Matthew Warren, Ph.D.
Adam Bartley, Ph.D.
Mark Bryan Manantan

July 2024

PACIFIC F★RUM
INTERNATIONAL

About the Pacific Forum

Based in Honolulu, Pacific Forum International (Pacific Forum) is a foreign policy research institute focused on the Asia-Pacific Region. Founded in 1975, Pacific Forum collaborates with a broad network of research institutes around the Pacific Rim, drawing on Asian perspectives and disseminating project findings and recommendations to global leaders, governments, and members of the public throughout the region.

The Forum's programs encompass current and emerging political, security, economic, and cybersecurity, and critical technology policy issues, and work to help stimulate cooperative policies through rigorous research, analyses, and dialogue.

The Pacific Forum

Web: www.cyberasean.pacforum.org
Facebook: Pacific Forum
Twitter: @PacificForum
Instagram: @pacforum
Podcast: Indo-Pacific Current
Email: pacificforum@pacforum.org

Mark Bryan Manantan (Series Editor) is the Director of Cybersecurity and Critical Technologies at the Pacific Forum. At the Forum he leads the Cyber ASEAN capacity-building initiative and the US Technology and Security partnerships in the Indo-Pacific.

Contributing Authors

AIDEN WARREN, PH.D. is a Professor at the School of Global, Urban and Social Studies, RMIT University.

CHARLES T. HUNT, PH.D. is a Professor at the School of Global, Urban and Global Studies, RMIT University.

MATTHEW (MATT) WARREN, PH.D. is the Director of the RMIT University Centre for Cyber Security Research and Innovation and a Professor of Cyber Security at RMIT University.

ADAM BARTLEY, PH.D. is a post-doctoral fellow at the RMIT's Centre for Cyber Security Research and Innovation.

MARK BRYAN MANANTAN SERIES EDITOR is the Director of Cybersecurity and Critical Technologies at the Pacific Forum.

Table of Contents

Executive Summary

Artificial Intelligence (AI) has emerged as a pivotal driver of economic growth and strategic relationships among nations, a trend expected to persist in the foreseeable future. Recognizing the imperative for economic competitiveness and strategic stability, the consensus on the necessity for collaborative efforts to establish AI standards, policies, guidelines, and ethical safeguards has gained widespread acceptance among major global actors.

Against the backdrop of escalating global tensions, fragmentation of economic trade and supply chains, and intensified technological competition, countries such as Australia, Japan, and the US are converging in terms of strategy and policy to navigate the transformative domains of AI and associated next-generation technologies.

The members of the Trilateral Security Dialogue (TSD) possess significant strengths in AI. However, the challenge lies in establishing common and practical grounds to harness their comparative advantages as they face mounting challenges from the bifurcation of AI standards, and normative frameworks, investment constraints, talent shortages, and competing interests among their respective private sectors. In this context, developing interoperable AI systems present complex challenges and overcoming them requires continual risk assessment and heightened investment in capabilities to stay ahead of adversaries. As nations navigate these challenges, collaboration and innovation are paramount for maintaining strategic relevance in an increasingly AI-driven world.

As such, collaborative AI sharing and development among the TSD partners will play a pivotal role in shaping their collective strategic posture and ensuring future competitiveness. As the global reliance on AI technologies further deepens, pooling resources and expertise has become essential for harnessing the full potential of AI-driven innovations. Through strategic minilateral frameworks like the TSD, the three nations can capitalize on their complementary strengths and capabilities to collectively address emerging security challenges, bolster regional stability, and maintain technological leadership.

Collaborative AI development not only fosters interoperability and synergy in defense capabilities, but also enhances information-sharing and intelligence cooperation, ultimately strengthening the resilience of each nation's defense posture. Moreover, such collaboration promotes economic growth, fosters innovation ecosystems, and cultivates a shared commitment to ethical AI development, ensuring that Japan, Australia, and the US remain at the forefront of technological advancement and strategic influence in the years ahead.

Amid the proliferation of US-led minilateralism, the TSD serves as a key arrangement for bridging AI capability shortfalls between the United States (US), Australia, and Japan. The TSD's longstanding history of fostering strategic cooperation, coupled with its deep institutionalization, positions it uniquely to address the unique challenges presented by AI development and deployment. By convening regular high-level meetings and facilitating collaboration across defense intelligence, technology capacity building, and interoperability exercises, the TSD creates an environment conducive to sharing expertise, aligning resources, and collectively addressing emerging security imperatives in the dynamic Indo-Pacific region.

Supported by the Australian Department of Defence's Strategic Policy Grants Program, researchers from RMIT University and the Pacific Forum engaged with experts, practitioners, and professionals across Australia, Japan, and the US in a series of high-level dialogues to discuss and understand TSD member perceptions and the inclination for advancing AI cooperation. Based on the three workshops, with the aim of moving toward a collaborative TSD AI agenda, this report presents an AI capability framework and policy tool that seeks to consolidate existing policies and initiatives and establish a common approach to AI development and innovation. Building on internationally agreed principles and best practices like the AI Partnership for Defense, the Global Partnership on Artificial Intelligence, the Organisation for Economic Co-operation and Development (OECD) AI Principles, the Hiroshima AI Process, and the Political Declaration on the Responsible Use of Artificial Intelligence, the AI capability framework advances four key elements: Innovation, Ethics, Interoperability, and Security.

Through collaborative efforts, including talent exchange, technology sharing, and public-private partnerships, *Innovation* focuses on advancing AI research and development. *Ethics* emphasizes the adoption of AI principles and standards to ensure responsible, reliable and human-centric AI deployment, while *Interoperability* underscores the importance of joint military exercises and knowledge sharing to enhance operational resilience. Lastly, *Security* highlights the implementation of security-by-design principles to safeguard AI-enabled technologies against cyber-enabled threats and ensure data privacy throughout the AI lifecycle. In employing

the proposed AI capability framework, the TSD members can strengthen their collective AI capabilities and address emerging security challenges in a strategic, functional and pragmatic fashion.

Introduction

The transformative impact of AI on the future of national and regional security is unquestionable. The implications for cybersecurity, nuclear deterrence, space capabilities, and information warfare are many. In accepting such assumptions, policymakers have taken an active role in shaping AI development, seeking stronger and more integrated industrial policies, developing national investment strategies for AI innovation, and building partnerships with international actors for force interoperability and defense. While the hype has permeated nearly all aspects of defense and security policymaking, the real challenge is finding the most effective and feasible vehicle to achieve such stated goals. This report unpacks the challenges and opportunities behind the proposition that the Trilateral Security Dialogue (TSD) can become an effective conduit for Australia, Japan, and the US to advance collaborative approaches to AI collaboration. It then outlines what practical steps are paramount to achieve a successful technological collaboration that aligns with their strategic and defense needs.

As AI amplifies and performs the tasks that humans currently undertake, its dual use application in military domains has sparked a race towards development and innovation in strategic areas. The national AI strategies of the US, Japan, and some European Union (EU) member countries, for instance, have noted the ability for disruptive states to impart significant asymmetrical advantages on others, causing harm, confusion, or strategic gain. For instance, at the technical level, computer vision algorithms can be employed to deface signs/data and upset systems operations while at the strategic level, AI-enabled machines can improve decision-making in complex environments, potentially transforming military affairs from the development of new hardware to logistics and tactical battlefield performance. AI can also be employed as part of election campaign tampering, mis- and dis-information campaigns, espionage and intellectual property theft, and supply chain and logistical manipulation, among other disruptions.

The Trilateral partners enjoy strong advantages in strategic technology sectors and, specifically, AI. However, as national policies across the world converge to address the requirements for driving competitive advantage in AI, compounding challenges have arisen. Australia and Japan face critical challenges in terms of investment, talent shortages, intergovernmental adoption, and security concerns across the information and cyber, maritime, air, space, and land domains.

Meanwhile, the geopolitical challenges of China's military and economic rise in the Indo-Pacific, and its diverging views towards the development of AI-enabled technologies and cyber governance, has fueled greater uncertainty on the application of AI-enabled defense capabilities in national military and naval systems. In this context, the increasing ubiquity of AI applications in the military domain offers key challenges among TSD members. These include: 1) not all applications or elements of AI design and development can be covered by one nation; 2) governments must continuously update their risk appetite due to the evolving nature of the technology; and 3) more resources must be allocated wisely to developing capabilities to get ahead of adversaries.

This report addresses the stated broader concerns on AI development and deployment, and explores the potential for collaboration among Australia, Japan, and the US within the TSD framework. Supported by the Australian Department of Defence's Strategic Policy Grants Program, researchers from RMIT University and the Pacific Forum have engaged with stakeholders across Australia, Japan, and the US in a series of expert dialogues to discuss and evaluate TSD member perceptions and inclination for boosting AI cooperation across the TSD platform. The insights in this report provide the foundation for the conception and implementation of an AI capability framework. As a useful policy tool, the AI capability framework explores how the TSD can establish a common roadmap among members to identify similar aims, policies, and best practices. It builds on the lessons learned and on insights and best practices from a multitude of experts, not just in the fields of AI, but also in governance, ethics, security, computer science, cognitive psychology, and international relations. In essence, the AI capability framework aims to guide strategic policymaking processes, offering key foundational components that can help guide the implementation of initiatives towards robust AI collaborative partnership.

The report contains four main parts. The first section charts the rise of technology minilateral groupings, highlighting how smaller arrangements composed of like-minded and agile states are better equipped to implement and foster alignment. The second section examines the changing strategic and technological environment, and outlines the need for Australia, Japan, and the US to pursue a tailored approach to burden-sharing and collaborative defense cooperation in the Indo-Pacific region. The third discusses the comparative advantages and unique value-added attributes of the TSD in pursuit of that collaborative modality. The fourth part presents the project's "AI capability framework." The final section presents the key policy recommendations for TSD members.

Rising Minilateralism: Addressing the Gap in Regional Multilateral Institutions

Minilateral groupings, like the TSD, offer their members a more efficient and agile path for mitigating challenges beyond the capabilities of traditional multilateral groupings. They are generally small, trust-based, and networked groups that share set values and aims on a particular set of issues. In the Indo-Pacific, the critically evolving threat of geopolitical tensions stemming from China's assertive challenge to the rules-based international order has caused a novel response in states to address the challenges via the expansion of high-level and technical relationships across states, governments, militaries, and advance technology and industry sectors.[1] With its prospects of expediency, minilaterals were considered more appealing to "secondary regional powers"— small-to-medium states "that are neither 'great' nor 'superpowers'" because they are more manageable, voluntary, leader-level focused, and regional instead of global.[2] Such arrangements are highly apposite to the strategic realities of secondary powers, offering greater agency in a security landscape that is increasingly uncertain and unpredictable, and where hedging strategies require more flexible arrangements beyond existing multilateral bodies.[3] Emergent minilaterals in this context include the Quad (Australia, Japan, India, US), the Trilateral Security Dialogue (Japan, Australia, US), the US-Japan-ROK Trilateral, and, more recently, the Japan-Philippines-US trilateral summit, among others.

A significant part of this framework building can be attributed to what is conceived as the erosion of the US "hub-and-spoke" alliance network. This mindset has not necessarily included the plausibility that the alliance network would end, but rather that guarantees of US assurance could no longer be accepted as a given. Such calculations have been weighed against the unpredictability of US defense and security commitments under former President Donald Trump's "America First" Policy that demanded greater burden sharing among US allies. These concerns have impacted the trust matrixes in US extended deterrence and capabilities. Consequently, Trump's disdain for US alliances has left lingering regional perceptions of the US' staying power, and the normative and economic influence of American institutional and popular pull, in doubt.[4] Another explanation offers that current minilateral development represents a new form of "virtual" alliance-making by Washington. Looking to establish "intra-spoke" security frameworks that move beyond the hub-and-spoke system, the US is redefining its regional strategic relationships to reinforce the order and values of international relations.[5]

Further explanation for the move toward minilateral groupings as strategic carriers of national, predominantly security agendas is the weakening of regional multilateral structures and institutions as mediums for conflict resolution. In Southeast Asia the traditional multilateral security bodies—the ASEAN Regional Forum (ARF) and the East Asia Summit (EAS)—have neither been sufficiently addressing shared concerns nor contributing to regional capacity building. As William Tow has argued, the ARF and EAS's institutional passiveness has not: fully served member needs; contributed to broader regional stability through the mediation of disagreement and conflict; or addressed the rise of threats presented in AI, quantum computing, malicious cyber activities, and other technology advancements with miliary-end use capacities.[6] Others, however, point to the role of minilaterals as potential containment-enforcing bodies designed to manage Chinese expansionism. One concern is that such groups will in the long term perpetuate a club of "like-minded" nations that may further exacerbate regional divisions amid fraying geopolitical conditions.[7]

The Evolving Strategic and Technological Environment

The evolution of strategic minilateral agreements has notably progressed from predominantly diplomatic arrangements, like the Quadrilateral Security Dialogue (Quad), that offer a quasi-deterrence formation on strategic issues (Chinese maritime and land claims), to technological capacity building partnerships that seek to incorporate technical responses to regional challenges, like AUKUS and Chip 4.

The Quad emerged as an ad hoc grouping to coordinate humanitarian assistance and disaster relief efforts in the wake of the 2004 Indian Ocean earthquake and tsunami. The transformation from nascent strategic forum between four democracies to a formal grouping gained momentum in the following years due to shared concerns about China's rising influence and assertiveness in the Indo-Pacific region.[8] Since 2019, these concerns have expanded to include new aims and an explicit focus on issues pertaining to maritime security, cybersecurity, infrastructure development, and economic cooperation.

By 2019, four emerging trends had contributed to a new focus on minilateral formations specifically targeting technological exchange, collaboration, and progress.[9] The first was that new habits of cooperation had become more entrenched between the Quad partners, and that such specified "minilateral" groups had become noticeably more useful than established multilateral frameworks. The second was the shift in Australia and India towards a more defensive posture vis-à-vis China after aggressive Chinese actions (for India, the 2020 Galwan clash; for Australia, the 2020 14-point demand and subsequent broad-based economic sanctions), shredding former assumptions about economic and security balancing strategies. The third characteristic reflected China's strategic regional growth and assertiveness, and the corresponding failure of both extent multilateral institutions and contemporary alliance agreements to respond to this behavior.[10] While such realizations occurred particularly among Japan and Australia, it undoubtedly touched all members of the Quad. Finally, and more recently since 2023, the emergence of AI as an undeniable and strategic force in national economic and security development prompted the states to address observable shortfalls in aggregate capabilities, namely around skills development, research and design, infrastructure, and value and supply chain security.

The Quad's Critical and Emerging Technology (CET) Working Group was the next phase of institutionalization. Established in March 2021, the CET Working Group coordinates efforts in pursuit of an "open, accessible, and secure technology ecosystem."[11] However, while the Quad has formed multiple working groups, issued a series of joint statements on technology principles, standards, development, governance, and on shared democratic values, action has generally been considered "fairly modest" with "discussions between government, the private sector, and civil society on these issues[…] still at an embryonic stage."[12] Others have charged that while the Quad has hosted a range of discussion and fora, there has been much less in terms of deliverables, meaning that "substantial collective progress is missing."[13]

More pointedly, India's "caveated" participation has led to perceptions that the grouping will predominantly remain a non-traditional security organization with loose aims of deterring Chinese revisionism. For instance, critics note that New Delhi has continued to "procure Russian defense equipment that poses barriers to interoperability with Quad partners" while also addressing its foreign policy goals separately, and "at odds," with other Quad members. This became an important point of contention following Russia's invasion of Ukraine, India's diplomatic response to it, and on questions of regional security provision.[14] At the 2024 Raisina Dialogue, Indian Minister of External Affairs S. Jaishankar advocated for strengthening ties with Russia, stating that "…it makes sense to give Russia multiple options" rather than only China. Prior to these comments, India rejected Japan's request to transport humanitarian aid supplies for Ukrainian refugees via India, further indicating that strategic minilateral collaboration in the Quad would be highly selective.[15] This followed on from previous Indian rejections, in 2019, of "Japan's Osaka Track, a framework to promote data transfers in favor of expanding data localization."[16]

India's reluctance to join closer strategic aims has caused some to temper expectations on the Quad to solely addressing issues such as climate change and health pandemics. The creation of AUKUS a week before the first face-to-face meeting among Quad leaders in September 2021 has suggested that more technologically focused structures have taken on new significance and have been located outside of the Quad, and specifically in AUKUS.[17] More significantly, the fluctuations of Quad alignment on strategic and technological issues from a more active to passive dynamics has highlighted the challenges of addressing new and novel agendas through more robust yet adaptive minilateral arrangements.

AUKUS

The AUKUS trilateral partnership differs considerably from the Quad by offering clear benchmarks for strategic growth and outcomes through the facilitation of nuclear-powered submarine (SSN) technological exchange. This agreement has broken several barriers on technological exchange, including, as Fraser and Soliman highlight, a 66-year moratorium on sharing nuclear propulsion technologies.[18] The new focus on technological sharing also highlighted the "need to reevaluate, restructure and invigorate alliance relationships through the combination of information-sharing and capability-building."[19] While there remains ongoing questions surrounding the final outcome, and indeed guarantees, of the SSN program, much hype has continued to build around Pillar II on emerging technologies.

Pillar II and the comparatively fast-paced development of trilateral activities across cyber, AI, quantum, and undersea capabilities have demonstrated that states like Australia, the United Kingdom, and the US are looking for more tangible outcomes and congruence across strategic aims, with a focus on interoperability and particularly around technological advances. For instance, the initial four workstreams (cyber, AI, quantum, and undersea) were quickly expanded to eight in 2022 to include electronic warfare, hypersonic and counter-hypersonic capabilities, innovation, and information sharing. Further expansions were announced in December 2023 to develop deep space radar capabilities to identify emerging threats in space.[20]

In the context of AI specifically, AUKUS members have accelerated trial and testing of systems in contested environments, with the first AI autonomy trial taking place in April 2023, testing collaborative swarming systems to detect and track military targets in real time. A second trial, TORVICE, testing autonomous vehicle behavior under adverse conditions, quickly followed. Meanwhile, in December 2023, the three nations announced that AI algorithms had been employed on multiple systems to "enhance force protection, precision targeting, and intelligence, surveillance, and reconnaissance," known as Resilient and Autonomous Artificial Intelligence Technologies (RAAIT).[21]

These actions and the corresponding diplomatic program undertaken by AUKUS partners has represented a new and even urgent approach to strategic minilateral development. Compared to the Quad, as Tomohiko Satake has written, a significant factor in AUKUS development, and particularly for Australia, is that it has become a "national endeavor," more so than any other arrangement in recent times.[22] Part of this momentum is facilitated by the perception that defense innovation systems have suffered a relative decline compared to states such as China and Russia,

and that the need to develop resilience against new asymmetric capabilities advanced by AI and other technological advancements must include multiple actors working in unison. For the US, the diffusion of systems, roles, technologies, and research programs among a membership base of trusted and likeminded partners will ultimately reduce aggregate technological and power deficits, and contribute to a broader strategy of deterrence that restrains Chinese technological authoritarianism.[23]

With the rising trend of minilateral arrangements, the possibility of adding new members have also come to the fore. The movement towards engaging Tokyo in technological partnership platforms has received significant attention, with recent comments by Australia's Prime Minister Anthony Albanese suggesting that Japan, among other nations such as New Zealand and Canada, could be invited to join the AUKUS Pillar II platform.[24] Such calls have followed further recommendations in the UK House of Commons Foreign Affairs Committee that Japan and possibly South Korea be included in Pillar II.[25] These suggestions have been swiftly walked back, with potential membership currently off the table. But as the commander of US naval submarine forces Vice Admiral Robert Gaucher has expressed, Japan is already in many ways a Pillar II partner "with or without AUKUS."[26] In other words, the foundations for technological collaboration in AI already exist in other minilateral arrangements and are under active consideration for expansion to broader strategic technological groupings. As Michael Auslin has further argued, the addition of Japan to AUKUS "would represent the natural evolution of the group."

However, the question of a potential "JAUKUS" (Japan, Australia, United Kingdom, US quadrilateral technological group) expansion has met with several challenges, illustrating the move toward evolving minilateral organizations, and not necessarily the consolidation of advanced technological cooperation in one grouping. The first consideration must be defining what aspects of Pillar II can or should be included in any expansion involving Japanese partnership. One issue that has consistently emerged in discussions of AUKUS and even Five Eyes enlargement, for instance, has been Tokyo's slow movement toward the adoption of counter-espionage laws. Japan's Unfair Competition Prevention Act, for instance, offers only limited measures and penalties for espionage, with necessary articulation lacking across business and education institution security, and principals on foreign states and researchers.[27] Another is Japan's security clearance system on issues such as sensitive economic information which has provided challenges for the private sector, and particularly international multinationals, who view intellectual property theft as a burgeoning problem.[28] Meanwhile, given, as Tsuruoka Michito notes, that "many of the Pillar 2 projects are not entirely new but have origins in "The Technical Cooperation Program (TTCP)," a technology cooperation framework among the Five Eyes countries,"[29] what barriers must first

be overcome, and what assurances, signatures, or compliances need to first be obtained by other partners before Japan's inclusion? Should AUKUS partners first work together to deliver the capabilities they've promised, some have further asked, before even considering expansion? These challenges so far don't have adequate answers.

AUKUS proffers a workable interface for collaboration across multiple advanced technology levels (or work streams) because of the high degree of trustworthiness and technical and legal compatibility needed to implement cooperation. In other words, deep cultural, political, linguistic, and historical complementarities have enabled AUKUS to exist, albeit with still many ongoing challenges. As John Blaxland has written on the subject, there is still a reluctance to go beyond the three core members of AUKUS. The delicacy of the configuration, despite great intimacy among the partners, has led to little appetite for expansion and what that might mean. AUKUS remains for the time being a "fragile endeavour, in part because all three members are rambunctious democracies that are going to have multiple elections in the lifetime of the project."[30]

For its part, Japan has expressed its support for AUKUS, however, amid the fanfare of its potential inclusion in Pillar II, some Japanese scholars and experts have noted that it will be more strategic for Japan to invest in existing minilateral arrangements through which it is an original founding member and that serve its national interests. Dovetailing on these comments, Australia's Ambassador to Japan Justin Hayhurst, in addressing questions on Australia's membership in technology minilaterals to which Japan is excluded, framed the answer as the need to advance new platforms amid existing frameworks: "Australia and Japan are really deepening our security cooperation and collaboration, including in defense and intelligence. We don't need to be the same, be in the same groups in the same ways all of the time. Those partnerships [AUKUS and Five Eyes] reflect very specific purposes and histories and systems."[31]

With the growing appetite to advance minilateral groupings, either in the context of the Quad or AUKUS, the most notable challenge is identifying what is strategically and operationally feasible among the key members to achieve any concrete breakthroughs. The reality is that, policymakers are grappling with the urgency of addressing the myriad of challenges associated with AI as a dual-use technology amid limited resources and shifting domestic priorities.

Revitalizing the Trilateral Security Dialogue (TSD)

This policy report asserts that the TSD remains to be the most adaptive and feasible mechanism that affords Australia, Japan, and the US the latitude to explore new areas of collaboration in AI. While the Quad, AUKUS, and more recently Chip 4 points to an increasing appetite towards technological collaboration, the domestic and national interests of each country still underwrite the political bandwidth, and resources that they can commit to the minilateral project. To this end, the TSD has fewer friction points that can impede practical collaboration given the high strategic complementarity that exists among the three member countries as evinced by the grouping's low-key but consistent engagements.

It should be noted at this point that the purview of locating strategic AI collaboration within the TSD is more than about merely right-sizing collaborative efforts or maximizing high-level alignment coupling across strategic innovative areas. Addressing AI capability gaps, and the needs of future AI workforce and machine learning data training requirements, can be managed, networked, and scaled at orders of magnitude through collaborative efforts and based on collective interests. Evidently, Japan is a leading innovation state in AI that can contribute strategically and mutually towards these aims. Its ability to act as a force multiplier to joint efforts already undertaken by the US and Australia in AUKUS, albeit for the time being in the domain of AI, should be the core consideration.

The TSD is a two-decade old strategic minilateral arrangement that has established, over time, a deepening institutionalzation to include defense intelligence, capacity building in advanced technology areas, disaster management and training, as well as military exercises aimed at high-end warfighting interoperability. Across the 22 years since its founding, efforts to regularly meet, discuss, strategize, and share information, policy experience, regulatory adjustments, political changes, and security considerations has led to an "unbroken" high-level consistency of interaction transcending broader conceptions of "like-mindedness" evident in, for instance, the Quad.[32] This has led some to suggest that the TSD already functions as the "inner core of

coordination among the US, Japan, and Australia in the face of strategic competition in the Indo-Pacific."[33]

Indeed, existing bilateral strategic partnerships between all three members have grown over time to work as "building blocks" for deepening TSD interaction.[34] These include not simply military exercises and bilateral ministerial meetings, but also examples like the Japan-Australia Agreement on the Transfer of Defense Equipment and Technology, and enhanced acquisition and cross-servicing agreements, as well as Reciprocal Access Agreements, across all three countries. During the historic US-Japan summit held in April 2024, Biden and Kishida signaled more collaboration among TSD members toward the creation of a joint air defense network to deepen interoperability and defense planning.[35] With the enforcement of the Japan-Australia Reciprocal Access Agreement in 2023, the building blocks for creating an integrated air and missile defense system are present. It also offers another milieu to expand cooperation among the three countries altogether, including critical and emerging technologies like AI.

It is relevant to note that the US views Australia and Japan as not merely responsible treaty allies, but also as its "most likeminded" and capable partners in the region.[36] For Tokyo, leaders have emphatically stated that while the US remains the primary strategic partner, Australia has become a close second. In Canberra, such sentiments serve to illustrate the momentum that has occurred with the TSD platform, despite often irregular meetings, and the synergy that has emerged over time in documents like Australia's 2023 Defence Strategic Review, Japan's 2022 National Security Strategy and the US' Indo-Pacific Strategy.[37] Meanwhile, established habits of cooperation, dialogue, and shared narratives of security, explicit in their respective security and defense strategies, have expanded trust networks, molded behaviors to forms of appropriate cultural, political, and regulatory conduct, and broadened cross-border exchanges and linkages—more so than the Quad or AUKUS framework can currently facilitate. Undoubtedly, these habits have led Japanese analysts to claim, as Thomas Wilkins writes, that the TSD "creates a safe space through which the country can be acclimatized in the company of familiar allies and partners, including as it self-adjusts to the implications of its own new defense policy settings, which may appear less menacing when operationalized in the company of others."[38] One recent study surveying experts and leaders in Australia, Japan, and the US recorded the sentiment that TSD member defense policies had generally moved from the era of mutual "interoperability" to the era of mutual "interchangeability,"[39] thus heralding a new era of strategic collaboration.

In the context of sensitive technology sharing across the AI spectrum, including areas such as machine learning, algorithm and hardware accelerators, natural language processing, data

analytics, and integrated circuit design, this "intimacy" facilitates a readymade platform for AI collaboration. Additionally, it distinguishes the TSD from other minilaterals that combine more routine instruments of policy action, or include partners with less developed institutional processes, experience, and even interaction with one or more members of the TSD. This distinction is important, as Koga notes, in determining the success of strategic minilaterals, which must "develop an optimal division of labor among themselves" in instrumentalizing outcomes.[40] In contrast to other frameworks where the form and functions of collaboration are not rooted in any deep consensus, the omission of these distinct arrangements is likely to curtail its capacity to undertake hard-security initiatives. Compared to the Quad, the formal alliance partnerships of Japan and Australia with the US allow them greater latitude to further explore sensitive areas of policy cooperation in a trilateral setting. In such formalized agreements, it becomes practically feasible for the TSD to effectively coordinate and realign resources, and implement adjustments, to accomplish the member's strategic goals due to established roles and expectations.

Introducing the TSD AI Capability Framework

This section demonstrates how the TSD members can implement a collaborative AI agenda that boosts capabilities and security. The AI capability framework presented here aims to explore how Canberra, Tokyo, and Washington D.C. can further expand the current portfolio of their collaboration in the tech domain. While recent statements from TSD meetings point to a substantive interest in technological collaboration, this project seeks to probe such intent deeper. The conception of the AI capability framework aims to support defense and security policymakers by consolidating the patchwork of existing policies, initiatives, and guidelines to explore credible pathways toward a common approach for AI development, ethics, security, and ultimately, interoperability across all three countries.

Building on internationally agreed principles and best practices like the AI Partnership for Defense, Global Partnership on Artificial Intelligence, the Organization for Economic Co-operation and Development (OECD) AI Principles, Hiroshima AI Process, and the Political Declaration on the Responsible Use of Artificial Intelligence, this report advances the following elements comprising the proposed AI capability framework: *Innovation*, *Ethics*, *Interoperability*, and *Security*. These four elements were presented during the three country consultations among the respective AI multi-stakeholder communities (in Washington D.C., Melbourne, and Tokyo).

Innovation. This component refers to the capacity of the three states to advance research and development AI-enabled technologies. It explores the exchange of talent, technology, and data among universities and research institutions as well as domestic start-up communities, and multinational companies through public-private partnerships, to cultivate international collaboration.

Ethics. Although the race is on to develop the next AI breakthrough, the US, Japan, and Australia continue to advocate for a human-centric approach to AI development. Thus, the adoption of ethical, responsible, and reliable AI, grounded on the principles, standards, and norms from the above-mentioned frameworks remains paramount throughout the AI lifecycle to avoid the unintended consequences of AI, like bias and misrepresentation. This will ensure further that the collection of data that is used for training AI models is anonymized and does not compromise any personal information that may impinge on data privacy and security, or fundamental human rights.

Interoperability. As the application of AI-enabled technologies in defense and security gains further ground, the three countries must start fielding such technologies in joint military exercises to test conceptual assumptions. Even though AI-enabled technologies have yet to reach the anticipated "game-changing" status that will dramatically shift military operations, sharing best practices and lessons learned in promising areas like logistics or cybersecurity is essential to improve resilience and take stock of lessons learned that may contribute to the development of operational doctrines.

Security. As the US, Japan, and Australia seek pathways to share AI-ready data to train machine learning models, they must implement principles and concepts like security or safety-by-design applied throughout an AI application's lifecycle.[41] This will ensure that training models are secured against adversarial AI, model subversion, and data poisoning. Observing security or safety-by-design from data collection to the design process will enhance the robustness of training models. Because most AI-enabled technologies are developed and acquired from third-party contractors, establishing a baseline certification and accreditation can help ensure the maintenance of high-quality control standards and due diligence.[42]

Drawing from the insights, feedback, and input from the project's three iterative workshops held in Washington D.C., Melbourne, and Tokyo throughout 2023, the following section unpacks the *Innovation*, *Ethics*, *Interoperability*, and *Security* of the AI capability framework. Rather than just a mere conceptual exercise, the project team sought experts' recommendations to operationalize the foundational components of the AI capability framework within the TSD context. These include:

Trilateral Security Dialogue AI Capability Framework

INNOVATION

ETHICS

INTEROPERABILITY

SECURITY
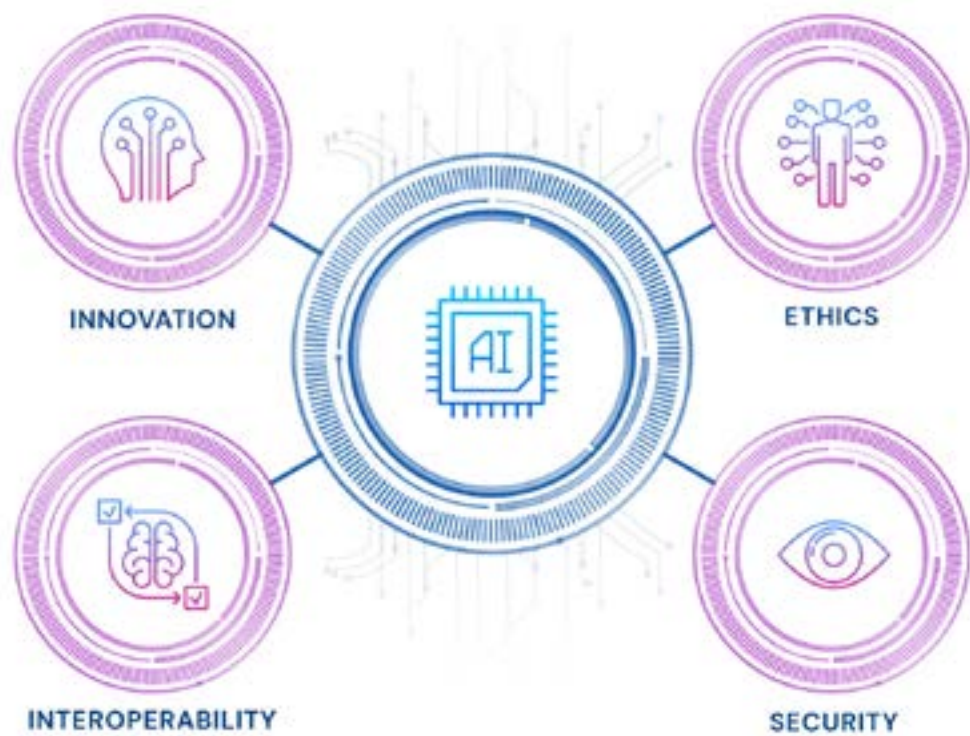
Innovation

TSD members' domestic AI research is currently being achieved in isolation, and with separate ambitions in mind. Collaboration is hampered by different bureaucratic and cultural innovation environments resulting in a fragmented and incohesive lack of vision, which in turn minimizes the value of research. As such, TSD members should address operational barriers to create an ecosystem of AI innovation and research. This should seek a unified approach to research and the combination of resources.

AI Research Security Dialogue

Establishing a dialogue to address these barriers would need to cover issues of intellectual-property theft, academic infiltration, norms, and regulations on academic publishing, and what this means for the broader concerns of economic security. Such a dialogue would develop

further methods for government departments to foster a robust and strategic AI research corridor among TSD partners. Transnational corridors of this sort, much as they have been previously associated with strategic economic integration pathways, will force policymakers to shift thinking about geography and time in the technology and development space from years to decades—a necessary adjustment given the often-short-term focus of governments due to short election cycles and shifting budget priorities.

If the plan must include a connective formula for institutionalizing collaboration on AI within the TSD in close alignment with perceived risks and benefits of AI technologies to human and societal values, a research and security dialogue will be a strategic necessity.[43] Like the AUKUS framework for working groups, any such dialogue must adopt a systems-focused format with room for agile thinking and adaptation. One avenue for thinking about working groups or dialogue frameworks is by first addressing the AI gaps at the national level. For instance, the US outperforms both Japan and Australia in machine learning (ML) by a significant margin, suggesting that US strengths in this area will be best for leading an ML working group on strategic AI aims.

For broader understanding, interaction, and time saving, this security dialogue should include key members of security commissions and committees in all three countries. Dialogue starting points include:[44]

HIGH IMPACT RESEARCH TRANSLATION

AI TALENT SHORTAGE CHALLENGES

CAPACITY BUILDING AND SHARING POTENTIAL ACROSS:

- Cyber technologies
- Machine Learning
- Natural Language Processing
- AI algorithms and hardware accelerators
- Advanced data analytics
- Adversarial AI
- Quantum computing

TRI-NATIONAL CLOUD RESEARCH DATABASE

CYBER AND DATA SECURITY

Tri-national cloud AI research repository

Experts from all three states agree that an agile data program will be necessary to incorporate new and novel approaches to AI development, particularly as shared development is required to enhance collaboration. Data categorization challenges exist due to inherent biases in national, ethnic, linguistic, processual, and legal norms of behavior in each nation. Accordingly, trialing and testing AI platforms on multiple data sources can reduce inaccuracies and build resilience in algorithms.

Building capability toward a tri-national cloud data bank and research repository can begin with a data resource index. This would encourage the exchange of legally appropriate/acceptable data sets between nations. There are several benefits to such sharing, including the ability to test and evaluate models across different systems, cultures, and national indicators. Building cross-cultural data training into AI systems would also build agility and learning within algorithms and clarify applications for interoperable use. Once established, the index can be transitioned into a shared interface through CloudBank, and as Imbrie and others note, this will help to incubate "an international network of research universities collaborating on these technologies." Similar suggestions call for providing cloud computing credits to researchers without access to large and diverse data sources. This would require governments to approach agreements with a focus on safe harbor laws and data consortia agreements that allow for controlled experimentation, while preserving privacy.

Limited data sharing agreements between Australian and the US, such as the Agreement on Access to Electronic Data for the Purpose of Countering Serious Crime, have recently come into force, highlighting that a nascent process for such exchange is already in motion.[45] This arrangement has been supported through existing and updated regulatory institutions, such as the US' Clarifying Lawful Overseas Use of Data (CLOUD) Act and Australia's Telecommunications (Interception and Access) Act 1979. It should be noted that while data sharing arrangements for countering serious crimes requires the special protection of data, which may for instance include child sexual abuse, other purely AI research-based data applications will require much less sensitive data sets. Certainly, all three states share data and information in some capacity, but until now, those relationships have lacked systemization. Additionally, while all three TSD partners have expressed the importance of securing data and information from malicious cyber actors, Japan continues to face assurance concerns following damaging cyber attacks on the Department of Defense in 2020. Agreeing to a framework for cyber and data security will be required to move forward on data exchange.

Advanced Technology Programs

The success of platforms like DARPA (Defense Advance Research Projects Agency) in the US demonstrates that advanced technology programs have a crucial role to play in strategic AI capability development. Acquisition, Technology & Logistics Agency (ATLA), a similar organisation in Japan, and Advanced Strategic Capabilities Accelerator (ASCA) in Australia, have created the grounds to advance AI research, data exchange, development, and experimentation. Combining talent from each nation is pivotal to maximise AI research potential, and this is possible through leveraging each state's advanced technology programs, each with their own set of unique contacts and talent. This will unify TSD research, at least in select areas, and synthesise a greater overall AI research capacity, enabling a comprehensive research vision.

Aspects of the US's "pathfinder" program model can also be adopted by Australia and Japan. A pathfinder program is designed to be ambitious in its capabilities, requiring significant investment, but delivering equally valuable strategic significance. The most notable pathfinder program is Project Maven, which is intended to automate object detection in intelligence imagery and video, serving US intelligence capabilities.[46] Recently, Australia has also seen success through the adoption of similar pathfinder programs such as the Ghost Bat program, which created an uncrewed aircraft through partnerships from defense, government, and industry. Significant investment in such programs will lend capacity and expertise to broader strategic programs and build leverage through knowledge communities that can be transported to new projects, or to lead integration efforts in new domains, as identified from the proposed AI Research Security Dialogue.

Regulatory Sandboxes

Regulatory sandboxes will help move AI projects and collaboration forward by generating understanding of legal, compliance, ethical, and linguistic hurdles that each government and military must consider. While such sandboxes currently do exist, these have been geared toward domestic security and economic considerations, such as with technology reforms for customs and borders processes.[47] To some extent, the AUKUS Pillars I and II arrangements have added to key government and industry competencies in building out expertise and knowledge for transborder legal, process, and standards adjustment on advanced technologies.[48] For the US and Australia, these competencies are progressing, particularly across amendments to the US

International Traffic in Arms Regulations, which now include guidelines for nontraditional defense contractors in the commercial world to work on AUKUS problems. However, these will need to be extended to include Japan to fully leverage the prospect for AI collaboration.

Regulatory sandbox trials in an AI collaborative context will require input from university research institutes, public service, defense, and industry, and may need to run on a case-by-case basis. The trials can begin with addressing multiparty, transnational research projects and challenges around intellectual property and export regulations. These trials may require phased iterations to build capabilities among the team to meet objectives. For instance, phase one could begin with minimal exemptions or amendments to legislation; phase two can consider trials that optimize existing operating environments; and phase three can explore more complex legislative or regulatory waivers or modifications, larger investments, or extra resourcing or costs.

These sandboxes can exist across technical and legal domains and, as they are designed for experimentation, can build nuance and expertise in cross national information and technology systems. This will help in the early critical phases of experimentation to understand whether applications will have merit. Regulatory sandboxes offer the TSD the best opportunity to determine feasibility on new projects, whilst understanding and correcting regulatory and other substantive barriers to these projects' success.[49] Cross-trilateral collaboration operating across multiple regulatory zones presents challenges to AI research and regulatory sandboxes present real-world testing of new ideas as a method of understanding the challenges of operating within a multi regulatory body environment.

Ethics

While TSD members share common values, the emergence of AI as a comparatively new and widely impacting phenomenon for policymakers has meant that the common implementation of ethical regulation is disjointed. A principles framework provides a set of shared guidelines, values, and standards that participating countries agree to uphold. These include: a common understanding of AI and cybersecurity principles, definitions, and best practices; the development of norms of behavior in AI, promoting responsible state behavior and deterring malicious activities; trust building; the facilitation of cooperation in response to cyber incidents, sharing threat intelligence, and implementing joint cybersecurity measures; and reducing legal uncertainties. As part of a broader capabilities' framework, the findings highlight further development of ethical guidelines.

In a collaborative framework, clear ethical standards will be central to legal exchanges of potentially exploitative applications and data. The US has made significant advancements in this respect. US Department of Defense programs, like the Urban Reconnaissance through Supervised Autonomy program, for instance, have made ethical challenges of human information and interaction central to design and operation. Across departments and agencies, ethical codes have been written into strategies for AI employment, with DARPA taking a lead in implementation at the design phase. With that being said, the number of incidents concerning the misuse of AI is rising. According to the Stanford AI Index, the number of AI incidents and controversies has increased 26 times since 2012, with notable instances including a "deepfake video of Ukrainian President Volodymyr Zelenskyy surrendering and US prisons using call monitoring technology on their inmates."[50]

Other examples include biases in natural language processing, leading to false information, and even as "fairer" language models are being developed, these have been found to contain biases. As these challenges indicate, AI systems can be difficult to understand and interpret, particularly if machine learning has contributed to algorithm development. This is the explainability problem of AI. AI machines and algorithms have become the workhorses and increasingly the main innovators in new AI development. While TSD states have clear legal distinctions on commercial AI use and intellectual property, there is a concern that legal jurisdictions and doctrine will be unable to keep up with and/or explain the algorithm and determine fault or wrongdoing. What is clear is that the legal implications of AI capability platforms remain understudied and under-regulated. Some issues, like the domains of "conflict technologies," such as AI softbots—software-based systems with great task variance and autonomy—remain in an extended and persistent state of ambiguity. This is due to machine learning capabilities and the untethered nature of such systems; that is, their disconnect from an explainable physical location and therefore jurisdiction.

Interoperability

The significance of interoperability between Australia, Japan, and the US in the domains of AI lies in the need for seamless collaboration amidst rapid advancements in AI technology. Interoperability is defined as the ability for organizationally and culturally differentiated units or systems to operate effectively to produce an efficient and congruent outcome of purpose. Forces or systems adopt interoperability to bring force multiplier effects and innovation to challenges seen as beyond the capabilities of individualized or isolated units. For AI, interoperability

across national governments or militaries calls attention to AI-enabled outputs that may exist in one nation for employment across separate national defense systems for force integration and deterrence. Currently, there is great interest and discussion on how states like Australia may contribute to AI interoperability across force partnership agreements like AUKUS, the Quad, and the TSD.

In Australia, discussions in this space have primarily revolved around the new aspirations for national defense security in the 2020 Defence Strategic Update (DSU) and the 2023 Defence Strategic Review (DSR). While the DSU significantly shifted national attention toward new and emerging technology areas, including cyber, as domains for special attention and new spending, the DSR reinforced this attention with a focus on AUKUS pillars I and II. This evolution in defense strategy documents also occurred within the Royal Australian Navy and more broadly across the Department in Defence. The Plan Mercator Strategy 2036, for instance, for the first time highlighted the importance of AI-enabled platforms as a more detailed feature of force planning. The 2020 Robotics, Autonomous Systems and Artificial Intelligence (RAS-AI) 2040 strategy, and its Army counterpart, RAS v2.0, provide more detail and strategic guidance across broader Defence in AI. Both documents have contributed to groundbreaking exercises using automated and AI systems, such as Autonomous Warrior 2022.

More military exercises, more variety

Identifying and rectifying gaps or disconnects in autonomous systems is complex and potentially impossible without access to sensitive training data, leading to unexpected behavior that can increase risks to humans and damage trust in their effective use. The challenge is particularly problematic in decision support systems or AI-enabled commands trained using data from one military's tactics and procedures, as they may not reflect the methods of other forces, potentially leading to unexpected actions that increase risks and damage trust. In this context, there is no substitute for military and teaming exercises. The difference between training on the ground and planning in the room is still vast. On this point, field commanders are still hesitant to move beyond what is comfortable and risk acceptable. Part of this will depend on building a culture that accepts greater risk taking. In moving forward, experts indicated that a starting point to strengthen interoperability include:

The US and Australia will have an initial head start over Japan in AI-simulated and/or-enabled exercises under existing arrangements through AUKUS. This, however, should not discourage further exercises in a TSD context, but rather reinforce extent training platforms, diversify military teams in AI-enabled battlefield training, and develop further capabilities across a range of pathfinding games and exercises designed to maximise human-machine teaming. Japan will certainly benefit from existing US-Australia habits and processes of interoperability, but Tokyo also brings a wealth of experience and training in advanced robotics platforms that may be missing in AUKUS Pillar II exercises.

Cross-cultural Training

Another factor includes understanding private sector cultural approaches to AI development. In the TSD member states, private actors retain significant influence over AI development streams, encouraging fears that too many separate AI operating systems will cause misinterpretation at critical moments, as communication across platforms remain isolated. As one expert commented, there is currently no automatic assumption within the private sector that interoperability should be factored in current designs. Increasing awareness about the need for interoperability among AI actors in the private sector is therefore needed. This may need to be managed by the government via a whole of society approach to development. Meanwhile, moving beyond narrow AI applications to machine learning poses important questions about what can be measured.

This point also cuts across the different cultures of the armed forces, which have been hesitant to adopt AI systems more broadly, and because testing and evaluation is likely to be difficult. Upstream data fusion between the different branches of the US armed forces and civilian agencies, for instance, has been problematic because not all forces share their data between each other. Without integration of data streams across the military, autonomous systems training may vary greatly between platforms, with significant implications for accuracy. Across partner countries and militaries, these systems are likely to be more diverse, requiring more training to avoid unexpected outcomes in applications. For instance, autonomous systems cross-trained on alternative data batches may react differently due to complexities in behavioral patterns or insignia, leading to inaccuracies in target acquirement.

This brings attention to linguistic differences, even between cultures and languages as close as those shared by the US and Australia. In Australia, for instance, authorities for drone and AI application require specific commands around line of sight, which is different still from Japan and the US, with implications for rules of engagement.

Traversing different legislative platforms and agency authorities requires greater depth of adjustment and exchange. This is particularly the case for Japan, whose processes and lines of authority diverge considerably from those in Australian and the US. One of the major contributing factors to Japan's slow adjustment to military and specifically advanced technology collaboration has been the highly political and sensitive cultural connotations attached to defense cooperation. Traditionally, and still to this day, the relationship between Japanese academia and the Ministry of Defense has been challenged by sensitivities attached to a pacifist constitution, deterring broader cooperation. Even industrial giants, and particularly legacy companies like Mitsubishi Heavy Industries, Kawasaki Heavy Industries Ltd, Toshiba Corporation, and ShinMaywa Industries Ltd, have been tepid in their responses to government outreach, wary of tarnishing their brand among consumers by joining in collaboration with the military.[51]

Notwithstanding this rigidity/inertia, some subtle but significant efforts have been made to develop stronger defense relationships with changes to how AI applications for security are perceived. The government for instance has replaced the term "dual use" with "multi-use" technologies to imply that more than simply military research is being conducted. The Ministry of Defense has also sought to publicly distance AI dual-use application research between the military and academia by funding projects through a separate agency, the Acquisition, Technology and Logistics Agency (ATLA).

Security

Standards settings

International partnerships illustrate that codifying standards will be key for further interoperability of systems, not just across defense applications, but also in the private sector and among government employees. Standards setting requires a two-step approach, a technical inter-TSD AI-technology sharing framework, and an international-focused governance strategy for AI. Common standards help build trust in AI exchange, development, and security within and across borders. In the technical domain, standards outline the language of design, implementation, legal accountability, employment, and common frameworks for operation.

A key feature of this discussion is the need for diplomatic, political, and logistical unity on AI standards settings. These include issues such as data identification, reliability and safety, data privacy protection, accountability, and fairness. A key point in expert interviews is that organizations like the International Standards Organization need to be staffed and joint programs to lead global governance on AI need to be established. The consensus among stakeholders was that not enough had been done by the US in its approach and, indeed, that it had become difficult to discuss contributions to international standards in the current political climate.

Another initiative emerging is the potential for Privacy Enhancing Technologies (PETs) to increase shared access to public data sets for AI training and testing. This has been recognized at the highest levels of American policy making, with US National Security Adviser Jake Sullivan making the case that PETs offer a promising area "to overcome data privacy challenges while still delivering the value of big data." Some early approaches to establishing stronger security features include:

Ensuring human oversight of AI actions, particularly with regards to the usage of AI in military actions.

Common data sharing practices that prioritize privacy, consent, security, and the promotion of responsible data stewardship.

Ensuring AI-enabled products are removed from certain exercises of force, such as in the usage of nuclear munitions.

Cooperating in ensuring the shared best practices of cyber security.

Cyber Security

A third component must be the cybersecurity of AI systems, algorithms, data, and other cyber-related criteria. Due to the range of risks associated with algorithmic tampering, data poisoning, and cyber related malware attacks, the security of cyber and AI systems is essential.[52] Integrated systems with different defensive cyber capabilities will prove difficult to manage, but the potential for attackers to focus on the vulnerabilities of the least sophisticated party is more acute. In this context, the US and Australia share similar cyber competencies and language around procedure, regulatory practices, and information sharing that Japan lacks.

For Japan, cybersecurity remains an achilles heel within its bilateral defense and security cooperation with the US, and by extension in the TSD setting.[53] Stepping up with the challenges, it has vowed to increase its investments to improve cybersecurity. The release of the New National Security Strategy in December 2022 underlines Japan's adoption of active cyber defense that allows the Japanese Self-Defense Force, in principle, to eliminate in advance the possibility of serious cyberattacks—a pre-emptive move that departs from Japan's previous approach in cyber defense. Moving beyond rhetoric, it was reported that the Ground Self-Defense Force will revamp its Signal School into Japan's Ground Self-Defense Force (JGSDF) System Communications and Cyber School to address the country's cyber defense talent shortfall.

Lastly, the Japanese parliament is also making headways in passing legislation that seeks to classify more information as confidential to enable information-sharing among the public and private sectors.[54] If successful, the law will provide security clearances for specific individuals, and organizations, offering a concrete pathway for Japan to share timely threat information with allies like the US,[55] and possibly create opportunities for sharing IP on sensitive technologies like AI, semiconductors, and quantum computing.[56] These efforts if implemented amid Japan's pacifist limitations can lubricate technology cooperation and information-sharing among Tokyo, Canberra, and Washington D.C.

Operationalizing the TSD AI Capability Framework: Policy Recommendations

The following recommendations provide a means to begin the process of deepening AI collaboration between the TSD members. Addressed to all three members, the following suggestions aim for broad adoption, and support a foundational effort to systematize knowledge exchange and operationalize key sectors of society and government to reduce AI shortfalls via a deepening of integration.

1. **Interoperability Awareness Campaigns:** All three states would benefit from awareness campaigns targeting private sector stakeholders and emphasizing the importance of interoperability in AI development. These campaigns should highlight the potential risks associated with siloed AI systems and promote a culture of collaboration and integration. For actors directly involved in military end-use product creation, this awareness is particularly important. Without common operation language and standards, AI algorithms will have difficulty operating outside of data and operational boundaries or across different data sets or operational environments.

   a. **Government-Led Initiatives:** Following on from the above, TSD members should implement a government-led approach to promote interoperability through regulatory frameworks or incentives that encourage private sector cooperation. This approach should involve collaboration across government agencies and industry partners to establish standards and best practices for AI development.

2. **Cross-Cultural AI Training for Military and Government Personnel:** TSD partners should develop specialized training programs for military and government personnel to navigate cross-cultural differences in AI adoption and application. These programs should focus

on understanding cultural nuances, linguistic variations, and legislative differences that impact AI deployment in diverse military contexts. They can also include integrated training on diffuse datasets, which would build resilience into platforms and broaden expertise across systems.

> a. Cultural Sensitivity Training: Provide cultural sensitivity training for military/public service/industry/academic personnel involved in international collaborations, particularly with countries like Japan, where cultural sensitivities may impact defense cooperation. This training should address historical and political factors shaping defense relationships and equip personnel with the skills to navigate cultural differences effectively.

3. Collaborative Platforms for Data Sharing: Each TSD member will need to facilitate collaborative platforms for data sharing among different branches of the military and civilian agencies to improve upstream data fusion and ensure consistency in autonomous systems training. Emphasizing the importance of data interoperability and the need for standardized protocols to enhance accuracy and effectiveness will speed up adoption. This will encourage confidence in national systems that can then be integrated and trialed at the TSD level.

4. Public Diplomacy and Perception Management: Australia, Japan, and the US should engage in public diplomacy efforts to reshape perceptions of AI dual-use applications in defense collaboration; emphasize the multi-use nature of technology; and highlight the benefits of cooperation between the military and academia. This will help foster transparency and trust-building measures to overcome resistance from legacy companies, or sectors of society, and address concerns about brand reputation.

5. Initiating Dialogue to Establish a Unified Ethical Framework for AI Development and Deployment: TSD members should engage in comprehensive dialogue to bridge ethical differences, or at least understanding, and collectively address the rising concerns associated with AI. This dialogue should prioritize the development and adoption of a shared ethical framework encompassing principles, values, and standards to guide responsible state behavior, mitigate malicious activities, and promote trust building. Additionally, collaborative efforts should focus on integrating ethical considerations into AI design and operation, as exemplified by programs like the US Department of Defense's Urban Reconnaissance through Supervised Autonomy program, to address challenges such as algorithmic biases and the explainability problem. Furthermore, there is an urgent need for enhanced research and regulation to navigate the legal implications of AI capabilities, particularly in domains such as conflict technologies, where jurisdictional ambiguities persist.

For securing AI systems and developing data integrity the following recommendations seek to strengthen cyber resilience through interdisciplinary cooperation and unified standards.

6. Global Standardization Efforts: Members should establish a collaborative framework for setting international AI standards, involving organizations like the International Standards Organization (ISO) and joint programs dedicated to global governance on AI. This framework should prioritize technical interoperability, data identification, reliability, privacy protection, accountability, and fairness. They should also encourage active participation and staffing of these organizations to ensure comprehensive representation and effective standardization.

7. Privacy Enhancing Technologies (PETs): Embracing PETs will facilitate shared access to public datasets for AI training and testing while preserving data privacy. The TSD members can promote PETs to overcome data privacy challenges while upholding ethical standards and responsible data stewardship. Exploration of PETs implementation in areas such as AI-driven military actions, data sharing practices, and cyber security should be undertaken. PETs will also enhance transparency and trust.

8. Human Oversight and Ethical Guidelines: The TSD members should implement mechanisms for human oversight of AI actions, particularly in military applications, to ensure accountability and adherence to ethical guidelines. This will mean developing common data sharing practices that prioritize privacy, consent, security, and responsible data management, and establishing protocols to restrict AI involvement in sensitive exercises of force, such as the usage of nuclear munitions, to mitigate risks and uphold ethical standards.

9. Collaborative Cyber Security Measures: Members will need to foster collaboration to strengthen cyber security capabilities and mitigate cyber threats. This can begin by sharing best practices in cyber security, information sharing, and regulatory practices to enhance collective resilience against cyber-attacks and safeguard AI systems, algorithms, and data from malicious activities. Such an approach will underscore the importance of common principles of cyber and AI safety and build trust, ensuring effective collaboration in advanced technological endeavors.

In some cases, collaboration among TSD member states is hindered by fragmentation and bureaucratic barriers. To overcome these challenges and maximize AI's potential, initiatives such as collaborative frameworks, data sharing platforms, government-led innovation programs, and regulatory sandboxes are vital.

10. **Establishing an Ecosystem of AI Innovation and Research:** TSD members should address operational barriers hindering collaboration in AI research and innovation by fostering a unified approach and combining resources. This includes creating a dialogue focused on AI research security to address issues such as intellectual property theft, academic infiltration, and regulatory standards on academic publishing.

11. **Tri-national Cloud AI Research Repository:** Members will benefit from the development of a tri-national cloud-based repository for AI research data to facilitate shared development and collaboration. This repository should address data categorization challenges and promote the exchange of legally appropriate datasets, enabling cross-cultural training of AI systems and interoperable use across different systems and indicators.

12. **Government-Funded Advanced Technology Programs:** Governments should leverage existing advanced technology programs like DARPA in the US, ATLA in Japan, and ASCA in Australia to advance AI research collaboratively. Emulating successful models such as the US's "pathfinder" programs to invest in ambitious AI research initiatives with strategic significance, fostering collaboration between government, industry, and academia.

13. **Regulatory Sandboxes for AI Projects:** Finally, the partners should establish regulatory sandboxes to facilitate AI innovation by temporarily exempting researchers and technologists from certain regulations, allowing for experimentation and testing of new projects. These sandboxes should involve interdisciplinary teams from research institutes, public service, defense, and industry to address legal, compliance, ethical, and linguistic challenges, ultimately promoting transparency and understanding of regulatory barriers to AI research and collaboration.

Conclusion

Supported by Australia's Department of Defence, this project examined AI's transformative impact to identify implications for the Trilateral Security Dialogue (TSD). Specifically, it assessed the Trilateral states' (Australia, Japan and US) strategic policy toward AI-enabled-capabilities in the defense environment to enhance Australia's strategic partnerships in an emerging multi-domain defense landscape.

To explore practical collaborative pathways for the TSD on AI, this project developed the AI capability framework comprising four foundational elements: (1) Innovation (2) Ethics (3) Interoperability and (4) Security. Ultimately, the proposed AI capability framework is a useful policy tool for defense and security policymakers, and industry practitioners, that aims to generate practical insights and bridge gaps in current capabilities. With recent trends on defense and technology collaboration, including AUKUS, the AI capability framework is a critical tool for Australia's strategic policymaking, specifically in improving its defense planning, development, and acquisition of AI-infused capabilities over the next decade. It serves as a flexible lens to understand the disruptive impacts of AI in the evolving threat-landscape in the Indo-Pacific region from the perspective of the US and Japan. Given the rapid AI development and innovation heralded by new breakthroughs like large language models, the AI capability framework affords policymakers a future-proof tool to navigate the ever-changing tech environment by not losing sight of the fundamental elements that underpin AI such as ethics, interoperability, and security.

As the project concludes, there remains emerging research areas that are worth pursuing in future endeavors. First, testing the viability of the AI capability framework, especially its four foundational components among defense and security policymakers. Second, the political appetite among the TSD members to formally institutionalize an AI or tech-focused working group that avoids duplication with other recent minilateral groupings, and ultimately, the capacity of the TSD to embrace wider private sector involvement as industry and start-up communities play a pivotal role in developing the next-generation of AI's dual-use applications.

Endnotes

1 Victor D. Cha. (2003). The Dilemma of Regional Security in East Asia: Multilateralism Versus Bilateralism. In P. F. Diehl & J. Levgold (eds.) Regional Conflict Management (pp. 104-122). Lanhma, MD: Rowman & Littlefield.

2 Thomas Wilkins. (2023). Middle Power Hedging in the Era of Security/Economic disconnect: Australia, Japan, and the 'Special Strategic Partnership.' International Relations of the Asia-Pacific 23: 93-127. Quote on page 95; William T. Tow. (2019). Minilateral security's relevance to US strategy in the Indo-Pacific: challenges and prospects, The Pacific Review, 32:2, 232-244.

3 Wilkins. Middle Power Hedging in the Era of Security/Economic disconnect.

4 Adam Bartley and Aiden Warren. (2022). Wither the Whole of Government? The Trump Administration, National Security, and the Indo-Pacific Strategy. Journal for Peace and Justice Studies 31, No. 1: 20-45; John Nilsson-Wright. (2017). Creative Minilateralism in a Changing Asia: Opportunities for Security Convergence and Cooperation Between Australia and Japan. Research Paper, Chatham House, July. Accessed via https://www.chathamhouse.org/sites/default/files/images/2017-07-28-Minilateralism.pdf

5 Thomas S. Wilkins. (2019). Security in the Asia-Pacific: The Dynamics of Alignment. Boulder: Lynne Rienner Publishers.

6 Tow Minilateral security's relevance to US strategy in the Indo-Pacific; see also, Robert O. Keohane. (1986). Reciprocity in International Relations. International Organization 40, No. 1: 1-27; Robert O. Keohane. (1990). Multilateralism: An Agenda for Research. International Journal 45, No. 4: 731-764; John G. Ruggie. (1993). Multilateralism Matters: The Theory and Praxis of an International Form. New York: Columbia University Press.

7 Mark Manantan. (2023). The ASEAN-Quad Partnership in Undersea Cables: Building Inclusion, Sustainability, and Regional Connectivity. ANU National Security College. Accessed via https://nsc.anu.edu.au/national-security-college/content-centre/research/asean-quad-partnership-undersea-cables-building

8 Lavina Lee. (2020). Assessing the Quad: Prospects and Limitations of Quadrilateral Cooperation for Advancing Australia's Interests. Lowy Institute Analysis. Accessed via https://www.researchgate.net/profile/Lavina-Lee-2/publication/341568596_Assessing_the_Quad_Prospects_and_Limitations_of_Quadrilateral_Cooperation_for_Advancing_Australias_Interests_Assessing_the_Quad/links/5ec7b04f299bf1c09ad29bba/Assessing-the-Quad-Prospects-and-Limitations-of-Quadrilateral-Cooperation-for-Advancing-Australias-Interests-Assessing-the-Quad.pdf; Dhruva Jaishankar and Tanvi Madan. (2021). How the Quad Can Match the Hype. Foreign Affairs, 15 April. Accessed via https://repec.viet-studies.com/kinhte/HowQuadMatchHype_FA.pdf

9 Lisa Curtis, Jacob Stokes, Joshua Fitt, and Andrew Adams. (2022). Operationalizing the Quad. Center for New American Century (CNAS) Indo-Pacific Security Program. Accessed via http://files.cnas.org.s3.amazonaws.com/CNAS+Report-IPS-Quad_Final.pdf; Vibhanshu Shekhar. (2022). Rise of Quad as a "Premier Regional Grouping": Harmonizing the Optics of Balancing and Normativism. Journal of Global Strategic Studies 2, No. 2: 37-60.

10 Frederck Kliem. (2020). Why Quasi-Alliances Will Persist in the Indo-Pacific? The Fall and Rise of the Quad. Journal of Asian Security and International Affairs, 7(3), 271-304.

11 Husanjot Chahal, Ngor Luong, Sara Abdulla, and Margarita Konaev. (2022). Quad AI: Assessing AI-Related Collaboration Between the United States, Australia, India, and Japan. Issue Brief, Center for Security and Emerging Technology. Quote on page 5. Accessed via https://cset.georgetown.edu/publication/quad-ai/

12 Guy Boekenstein. (2023). The Quad's Role in Tech Diplomacy. The Strategist, Australian Strategic Policy Institute, 15 February. Accessed via https://www.aspistrategist.org.au/the-quads-role-in-tech-diplomacy/. For statements see "Quad Principles on Technology Design, Development, Governance and Use, White House press release, 24 September , 2021, https://www.whitehouse.gov/briefing-room/statements-releases/2021/09/24/quad-principles-on-technology-design-development-governance-and-use/; "Fact Sheet: Quad Summit," White House press release, 12 March, 2021, https://www.whitehouse.gov/ briefing-room/statements-releases/2021/03/12/factsheet-quad-summit/.

13 Urmika Deb. (2024). India in the Quad: Insider or Outlier? The strategist, Australian Strategic Policy Institute, 2 April. Accessed via https://www.aspistrategist.org.au/india-in-the-quad-insider-or-outlier/

14 Kate Sullivan de Estrada. (2023). India and Order Transition in the Indo-Pacific: Resisting the Quad as a "Security Community." The Pacific Review 36, No. 2: 378-405. See also Derek Grossman. (2018). India is the Weakest Link in the Quad. Foreign Policy, 23 July. Accessed via https://foreignpolicy.com/2018/07/23/india-is-the-weakest-link-in-the-quad/; Chet Lee. (2021). India: The Quad's Weakest Link. The Diplomat, 19 October. Accessed via https://thediplomat.com/2021/10/india-the-quads-weakest-link/

15 Nikkei Asia. (2022). Japan Drops Plan to Send Ukraine Refugees Aid Supplies Via India. 26 April. Accessed via https://asia.nikkei.com/Politics/Ukraine-war/Japan-drops-plan-to-send-Ukraine-refugees-aid-supplies-via-India

16 Chahal et alQuad AI: Assessing AI-Related Collaboration Between the United Staes, Australia, India, and Japan.

17 Amélie Chalivet. (2022). India's Place in the Quad in Light of AUKUS. Policy Brief, Network for Strategic Analysis, 29 April. Accessed via https://ras-nsa.ca/indias-place-in-the-quad-in-light-of-aukus/

18 Jada Fraser and Mohammad Soliman. (2023). The Quad, AUKUS, and the I2U2 Formats: Major Lessons from Minilaterals. Orbis 67, No. 3: 411-419.

19 Austin Wyatt, James Ryseff, Elisa Yoshiara, Benjamin Boudreaux, Marigold Black, and James Black. (2024). Towards AUKUS Collaboration on Responsible Military Artificial Intelligence: Co-Design and Co-Development of AI Among the United States, the UK and Australia. Santa Monica, CA: RAND Corporation. Accessed via https://www.rand.org/pubs/research_reports/RRA3079-1.html.

20 Louisa Brooke-Holland. (2024). AUKUS Pillar 2: Advanced Capabilities. Research Briefing, House of Commons, Parliament, United Kingdom, 8 March. Accessed via https://researchbriefings.files.parliament.uk/documents/CBP-9842/CBP-9842.pdf

21 US Department of Defense, AUKUS Defence Ministers Meeting Joint Statement, 1 December 2023

22 Tomohiko Satake. (2023). The Rise of Minilateralism in the Indo-Pacific: The Quad and AUKUS. Japan Spotlight, March/April.

23 Austin Wyatt, James Ryseff, Elisa Yoshiara, Benjamin Bourdeaux, Marigold Black, and James Black. (2024) Towards AUKUS Collaboration on Responsible Military Artificial Intelligence. Research Report. RAND Australia, Accessed via https://www.rand.org/pubs/research_reports/RRA3079-1.html; William Greenwalt and Tom Corben (2023). Breaking the Barriers: Reforming US Export Controls to Realise the Potential of AUKUS, United States Studies Centre at the University of Sydney.

24 Ben Westcott. (2024). Japan Won't be Invited to Formally Join AUKUS, Australian PM Says, Bloomberg, 9 April. Accessed via https://www.bloomberg.com/news/articles/2024-04-09/japan-won-t-be-invited-to-formally-join-aukus-australia-pm-says

25 Tilting Horizons. (2023). the Integrated Review and the Indo-Pacific. Parliament of the United Kingdom House of

Commons Foreign Affairs Committee, 30 August. Access via https://publications.parliament.uk/pa/cm5803/cmselect/cmfaff/172/report.html

26 Arron Mehta. (2024). Aussie PM Says no Japan as AUKUS Member, But Pillar II on Table. Breaking Defense, 8 April. Accessed via https://breakingdefense.com/2024/04/as-aukus-opens-discussions-with-japan-officials-confident-pillar-i-will-continue-unabated/

27 Ryosuke Hanada. (2023). No, Japan is Not Ready for AUKUS. The Strategist, Australian Strategic Policy Institute. 7 December. Accessed via https://www.aspistrategist.org.au/no-japan-is-not-ready-for-aukus/

28 The Japan Times. (2024). Protecting Information is the Key to National Security. 1 March, accessed via https://www.japantimes.co.jp/editorials/2024/03/01/protecting-information-national-security/

29 Tsuruoka Michito. (2024). Why AUKUS Will Not Become JAUKUS. The Diplomat, May 13. Accessed via https://thediplomat.com/2024/05/why-aukus-will-not-become-jaukus/

30 John Blaxland. (2024). Japan's AUKUS Cooperation Limited, Formal Joining Unlikely. Mirage News April 9. Accessed via https://www.miragenews.com/japans-aukus-cooperation-limited-formal-joining-1211095/

31 Australian Embassy Tokyo, Japan. (2023). Questions and answers session - Ambassador-Designate Hayhurst at Japan National Press Club, 13 April, Accessed via https://japan.embassy.gov.au/tkyo/transcript_jnpc_qa.html

32 Hayley Channer. (2022). Trilateral – Not Quad – Is the Best Chance for Indo-Pacific Defense. The Diplomat, 16 June. Accessed via

33 Thomas Wilkins. (2024). U.S.-Japan-Australia Trilateralism: The Inner Core of Regional Order Building and Deterrence in the Indo-Pacific. Asia Policy 19, No. 2: 159-185, quote on pp 176-177.

34 Thomas Wilkins, Kyoto Hatakeyama, Miwa Hirono, and H. D. P. Envall. (2024). Australia, Japan and the New Web of Indo-Pacific Minilateralism. East Asia Forum, 21 February. Accessed via https://eastasiaforum.org/2024/02/21/australia-japan-and-the-new-web-of-indo-pacific-minilateralism/

35 Bo Erickson and Kathryn Watson. (2024). Biden Announces New Steps to Deepen Military Ties Between the U.S. and Japan. CBS News, 10 April. Accessed via https://www.cbsnews.com/news/biden-kishida-to-announce-ramped-up-military-partnership/

36 Jim Garamone. (2022). "Austin Holds Trilateral Meetings with Indo-Pacific Allies," U.S. Department of Defense, June 11. Accessed cia https://www.defense.gov/News/News-Stories/Article/Article/3059988/ austin-holds-trilateral-meetings-with-indo-pacific-allies.

37 "Fact Sheet: Indo-Pacific Strategy of the United States," White House, February 11, 2022 https://www.whitehouse.gov/briefing-room/ speeches-remarks/2022/02/11/fact-sheet-indo-pacific-strategy-of-the-united-states; Yoshimasa Hayashi, "Adoption of the New 'National Security Strategy (NSS),' " Ministry of Foreign Affairs (Japan), December 16, 2022 https://www.mofa.go.jp/press/release/press4e_003192.html; Anthony Albanese and Richard Marles, "Release of the Defence Strategic Review," Defence Ministers (Australia), April 24, 2023. Accessed via https://www.minister.defence.gov.au/media-releases/2023-04-24/release-defence-strategic-review

38 Wilkins. U.S.-Japan-Australia Trilateralism.

39 Michael J. Green, Christopher B. Johnstone, Peter Dean, Nicholas Szechenyi, Tom Corben, and Shizuka Takada. (2024). Operationalising Japan's Security Role in Asia: A Survey of Experts in Japan, the United States and Australia. United States Study Cetre report, April 30. Accessed via https://www.ussc.edu.au/operationalising-japans-security-role-in-asia-a-survey-of-experts-in-japan-the-united-states-and-australia

40 Kei Koga. (2022). A New Strategic Minilateralism in the Indo-Pacific. Asia Policy, 17(4), 27–34.

41 Mark Manantan. (2021). The Cyber AI Nexus: Implications for the US-Japan Cybersecurity Alliance. Pacific Forum. Accessed via https://pacforum.org/wp-content/uploads/2021/11/PacForum_Report_Final_Single_Page.pdf

42 Ibid.

43 Ian Klaus and Simon Curtis. (2024). The New Corridor Competition Between Washington and Beijing. The Carnegie Endowment for International Peace, 4 April. Accessed via https://carnegieendowment.org/posts/2024/04/the-new-corridor-competition-between-washington-and-beijing

44 AI gaps were determined using the Australian Strategic Policy Institute's Critical Technology Tracker. Dated 30 May, 2024. Accessed via https://techtracker.aspi.org.au/

45 Georgia Butler. (2024). Australian and US Governments' Cloud Act Agreement for Sharing Data Comes into Force. Data Center Dynamics, News, 2 February. Accessed via https://www.datacenterdynamics.com/en/news/australian-and-us-governments-cloud-act-agreement-for-sharing-data-comes-into-force/

46 Allen, G. C. (2023). Six Questions Every DOD AI and Autonomy Program Manager Needs to Be Prepared to Answer, Center for Strategic and International Studies, https://www.csis.org/analysis/six-questions-every-dod-ai-and-autonomy-program-manager-needs-be-prepared-answer

47 Regulatory Sandbox. Australian Border Force. Accessed 31 May, 2024 via https://www.abf.gov.au/about-us/what-we-do/regulatory-sandbox

48 Stew Magnuson. (2023). Governments Work Behind Scenes to Pave Way for AUKUS Success. National Defense, 24 October. Accessed via https://www.nationaldefensemagazine.org/articles/2023/10/24/governments-work-behind-scenes-to-pave-way-for-aukus-success

49 OECD, (2023). "Regulatory Sandboxes in Artificial Intelligence", OECD Digital Economy Papers No. 356, OECD Publishing, Paris, https://doi.org/10.1787/8f80a0e6-en

50 Nestor Maslej, Loredana Fattorini, Raymond Perrault, Vanessa Parli, Anka Reuel, Erik Brynjolfsson, John Etchemendy, Katrina Ligett, Terah Lyons, James Manyika, Juan Carlos Niebles, Yoav Shoham, Russell Wald, and Jack Clark, "The AI Index 2024 Annual Report," AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA, April 2024. Josephine Wolff. (2020). How to Improve Cybersecurity for Artificial Intelligence. Brookings Report, series: AI Governance. Accessed via https://www.brookings.edu/research/how-to-improve-cybersecurity-for-artificial-intelligence/

51 Henry Ridgewell. (2023). Japan Struggles to Booste Defense Industry Amid China's Military Ambitions. VOA News, August 10. Accessed via https://www.voanews.com/a/japan-struggles-to-boost-defense-industry-amid-china-s-military-ambitions/7219689.html

52 Wyatt Hoffman. (2022) AI and the Future of Cyber Competition, Center for Security and Emerging Technology, 2021.Jackett, Jennifer, Laying the Foundations for AUKUS: Strengthening Australia's High-Tech Ecosystem in Support of Advanced Capabilities, United States Studies Centre at the University of Sydney.

53 Kaori Kaneko, Tim Kelly, and John Geddie. (2024). The Glitch in Japan's Plans to Bolster U.S. Defence. Reuters, 27 April. Accessed via https://www.reuters.com/world/glitch-japans-plans-bolster-us-defence-2024-04-26/

54 Reuters. (2024). Japan Proposes Law to Classify More Information as Confidential. 28 February. Accessed via https://www.reuters.com/world/asia-pacific/japan-proposes-law-classify-more-information-confidential-2024-02-27/

55 Mark Manantan. (2024). US-Japan: Advancing Cybersecurity and Resiliency in the Age of Uncertainty. Pacific Forum. Accessed via https://pacforum.org/wp-content/uploads/2024/02/EN-Pacific-Forum-Layout-January-2024-Pass-Pages_Feb7.pdf

56 Yuki Fujita. (2024). Japan Eyes New Security Clearances to Aid Overseas Tech Cooperation. Nikkei Asia, 28 February. Accessed via https://asia.nikkei.com/Politics/Japan-eyes-new-security-clearances-to-aid-overseas-tech-cooperation

PACIFIC FORUM

INTERNATIONAL

pacforum.org | pacificforum@pacforum.org