The Future of Intelligence Analysis: U.S.-Australia Project on AI and Human Machine Teaming

September 2024

SPECIAL COMPETITIVE STUDIES PROJECT

ASPI
AUSTRALIAN STRATEGIC POLICY INSTITUTE

Contents

Executive Summary

Rapid advances in the development of artificial intelligence (AI) technologies since late 2022, particularly the deployment of Generative AI (GenAI) chatbots powered by large language models (LLMs), have demonstrated the potential for AI to revolutionize how states conduct intelligence work. AI technologies are very likely to continue to rapidly advance given the large amount of investment from the private sector and nation states, with some experts predicting we will see the advent of artificial general intelligence (AGI) – a type of AI that achieves, or surpasses, human-level capacity for learning, perception, and cognitive flexibility – by the end of this decade.[1] Even if this ambitious goal is not fully met, the LLMs available within the next three years will probably far surpass the capabilities of systems we use today and will be able to solve complex problems, take action to collect and sort data, and deliver well-reasoned assessments at scale and at speed.

- The effects of AI likely will be felt at all levels of the intelligence enterprise, including in collection, but the arena that we assess will see the earliest impact will be on the all-source analytic mission because of AI's ability to quickly process large volumes of data and GenAI's ability to produce meaningful insights from them.

Intelligence agencies that are able to effectively and safely incorporate GenAI into their workflows could realize substantial gains in the breadth and depth of their analytic work and significantly speed up the delivery time of critical insights to decision-makers. If integrated into and adapted for intelligence analytic work, currently available GenAI tools would speed up and enhance several stages of the analytic workflow, from the search for and discovery of new data, to conceptualizing analytic products, to applying analytic tradecraft and conducting classification checks.

- Future systems will be even more capable and will be able to shoulder more of the analytic workload; first by autonomously taking care of routine tasks, such as foreign language translation, databasing, and data visualization and eventually by more directly applying intelligence analysis tradecraft to answer policymaker questions and provide unique, value-added insights.

- **While U.S. and Australian Intelligence Communities (ICs) are well-acquainted with AI and have been tracking its development for years, they are taking a cautious approach to deployments.** Their hesitancy is rooted in concerns – well-founded at present – over

[1] Tim Mucci & Cole Stryker, Getting Ready for Artificial General Intelligence, IBM (2024).

some of the technical limitations of existing GenAI systems and the lack of clear legal and policy guidance about how these systems should be used for national security purposes. There is also skepticism about the value added of this technology over highly-trained human analysts with deep subject matter expertise. This has led analytic managers to ban, or severely limit, the use of GenAI, and constrained deployments of GenAI tools to narrow uses, such as document summarization, that are well within the capabilities of current LLMs but will lag far behind what future systems will be able to provide.

- Their hesitancy also reflects a view among analytic practitioners that AI is "just another software tool" that analysts will need to learn how to use and that existing approaches to technology adoption are sufficient. **It is our assessment, however, that future AI capabilities will be so powerful that they will transform the business of intelligence analysis, and that the ICs need to act with greater urgency now to prepare for their arrival and effective deployment, especially in anticipation of adversaries successfully leveraging the power of these tools.**

Australian and U.S. leaders should begin laying the groundwork now for the GenAI future that lies just around the corner. To avoid remaining perpetually behind the curve on the pace of AI technological development, analytic managers should shift their focus away from what GenAI can do today and instead make reasoned bets on what GenAI will be able to deliver within the next 3-5 years. In addition to pressing their home agencies to acquire and integrate AI-related infrastructure (particularly advanced compute capabilities, access to cutting-edge commercially-available GenAI models and algorithms, and secure data storage), we make the following recommendations for U.S. and Australian analytic managers:

1. **Design for Continuous AI Model Improvements.** With the expected exponential growth of LLMs, the ICs cannot only look only to the current technological state-of-play but must also anticipate GenAI's future trajectory over the course of the next five, ten, or twenty years. They must balance quickly and safely deploying these tools while also clearly ensuring the proper integration of the expertise and skills of human analysts. This will include accounting for larger LLMs, expansions in context lengths, and further developments in more sophisticated systems like compound and agentic systems.

2. **Insist on Automating Portions of the Analytic Workflow.** Managers should fully deconstruct all of the key elements of the analytic process with an eye toward using AI capabilities to shrink the amount of time required to deliver insight to policymakers while maintaining stringent standards for quality, accuracy, and analytic tradecraft. Elements that currently have a heavy amount of human redundancy, such as the analytic review process, probably could see some efficiencies.

3. **Build Human-Machine Analytic Teams.** Anticipating the growing power of AI systems, IC leaders should stand up analytic teams that purposefully blend the relative strengths of

humans and machines. This will require establishing the expectations and rules-of-the-road for what humans are responsible for, along with creating new tradecraft standards.

4. **Create AI-Ready Training and Incentive Structures for the Analytic Workforce.** To effectively integrate these systems will require a workforce that is prepared and adept at exploiting these tools to their fullest potential. The ICs will need to invest in digital acumen, both through the recruitment of highly-trained talent and upskilling the existing workforce.

There are opportunities for U.S. and Australian IC leaders to collaborate on the development and responsible deployment of AIs for intelligence analysis. Potential areas for cooperation include articulating ethical and analytic standards for the use of AI systems, exchanging findings from AI testing and evaluation programs, sharing best practices in the management of human-machine teams, and piloting the use of AI to tackle discrete intelligence analysis problems on a shared high-side data cloud.

Scope Note

Conducted through a collaboration between the Special Competitive Studies Project (SCSP) and the Australian Strategic Policy Institute (ASPI), this project seeks to illuminate AI's potential to enhance all-source intelligence analysis. We engaged experts from the national security and emerging technology sectors through a series of workshops held simultaneously in Canberra and Washington. A complete list of contributors can be found at the end of the report.

The inaugural workshop, held in late November 2023, assessed current AI applications, private sector advancements indicative of future potential, and adoption challenges. The second workshop, held in February 2024, developed a series of recommendations for aligning cutting-edge generative AI with analysis needs and supporting the broader organizational transformation needed to harness the potential of AI models for all-source analysis. The outcomes of these workshops, supplemented by a review of relevant literature and expert consultations, form the foundation of this comprehensive report, which presents specific recommendations for strategically implementing AI in the intelligence operations of both countries, targeting near-term, impactful applications.

Introduction

The rapid evolution of artificial intelligence (AI), transitioning from speculative fiction to tangible reality, is underscored by advancements in machine learning (ML) and natural language processing (NLP) as well as the meteoric rise of tools like Gemini and ChatGPT, which boast more than 100 million users.[2] AI-powered machines already excel at games, medical diagnoses, and standardized tests, and specialized AI models now perform tasks in domains like finance, science, marketing, data management, research, game development, and healthcare.[3]

OpenAI's release of ChatGPT in November 2022 – and subsequent releases from not only OpenAI itself (the fourth version, ChatGPT-4o, was released in May 2024), Google (Bard, March 2023, and Gemini, December 2023) and Anthropic (Claude, March 2023) – heralded a new generation of artificial intelligence that offered unprecedented opportunities for users to query and interact with overwhelming volumes of information. These LLM-based generative AI models have a variety of uses, most notably using algorithms to create novel responses to user questions by drawing on the patterns of words detected in the massive amounts of data on which they have been trained. LLMs are likely most familiar to readers, but they are not the only type of (or approach to) generative AI currently available.[4] For this report, however, we focus on LLM-powered generative AI.

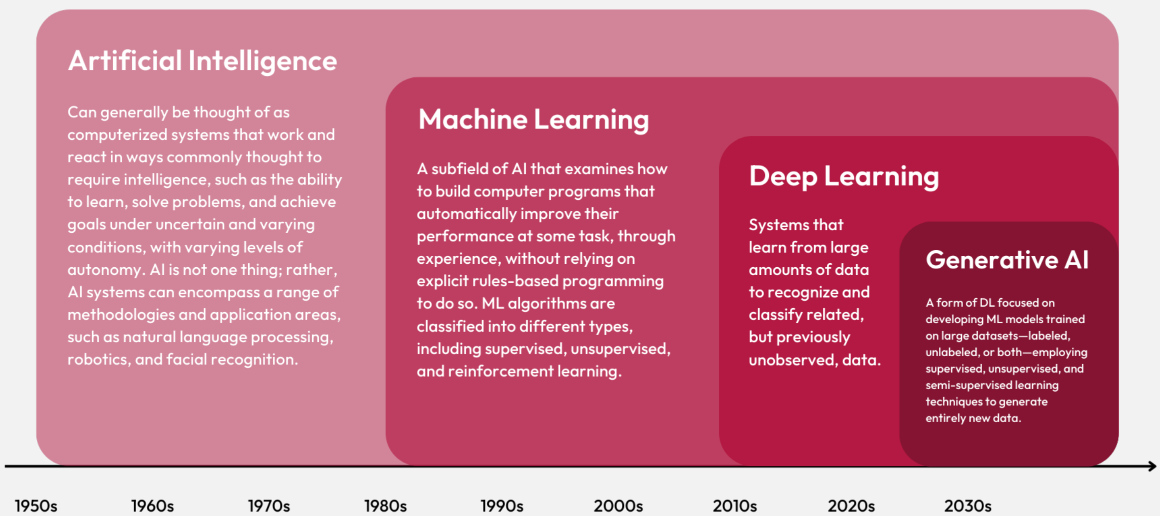
[2] Aisha Malik, OpenAI's ChatGPT Now Has 100 Million Weekly Active Users, Tech Crunch (2023).

[3] David Silver, et al., Mastering the Game of Go Without Human Knowledge, Nature (2017); Machine Learning's Potential to Improve Medical Diagnosis, U.S. Government Accountability Office (2022); Demis Hassabis, AlphaFold Reveals the Structure of the Protein Universe, DeepMind (2022); Introducing BloombergGPT, Bloomberg Professional Services (2023); Ross Taylor, et. al., Galactica, Meta (2023); Daniil A. Boiko, et al., Emergent Autonomous Scientific Research Capabilities of Large Language Models, ArXiv (2023); Copy.ai (last accessed 2024); Data Engine, Scale AI (last accessed 2024); Elicit, Ought (last accessed 2024); Scenario (last accessed 2024); A.J. Ghergich, How Automation Is Transforming Healthcare Jobs, Forbes (2021); and Awesome Generative AI, Github (last accessed 2024).

[4] Retrieval-Augmented AI, for example, uses traditional search methodologies to identify the documents that are most relevant to the users' queries, effectively improving the quality of the response while simultaneously lowering the probability of the AI incorrectly inferring an answer based on the statistical patterns that exist in the underlying data. The concept of RAG AI was introduced in Patrick Lewis, et al., Retrieval-Augmented Generation for Knowledge Intensive Tasks, arXiv (2021). Andrew Ng has advocated for "data-centric AI," which focuses on optimizing the data and metadata to support more sophisticated AI. See Data-Centric AI Resource Hub (last accessed 2024).

What We Mean by "Artificial Intelligence"

This paper explores the potential of Generative AI (GenAI) powered by large language models (LLMs) for intelligence tasks involving unstructured data. While often used interchangeably, ML, Deep Learning (DL), and GenAI are distinct AI subfields with unique capabilities and challenges. ML uses algorithms to interpret data and make predictions, forming AI's foundational layer. DL, a subset of ML, utilizes complex neural networks for tasks like image recognition and natural language processing, handling vast volumes of structured and unstructured data. GenAI, including technologies such as Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs), represents the most advanced subset. GenAI focuses on creating realistic new content like text and images from unstructured data types, requiring the most sophisticated hardware like graphics processing units (GPUs) and tensor processing units (TPUs).



Graphic Source.[5]

Imagine an intelligence analyst who employs GenAI to help forecast Russia's next moves in Ukraine or to unearth illicit Chinese funding in Taiwanese media, uncovering an emerging influence network before Taiwan's elections. She is no longer overwhelmed by data; instead, she employs multiple AI-powered tools to efficiently extract crucial insights with the computational might at her disposal. However, this analyst would not rely solely on AI; she knows that she will need to communicate and thus contextualize those insights. She would critically assess its

[5] Stuart Russell & Peter Norvig, Artificial Intelligence: A Modern Approach, Pearson Education Press at 17-26 (2021); Jeffrey A. Dean, A Golden Decade of Deep Learning: Computing Systems & Applications, Daedalus (2022).

predictions, inject her own tacit knowledge, common sense, and moral compass to steer AI past its inevitable quirks and make nuanced decisions to adapt to surprises, and manage sensitive scenarios or non-routine situations where AI may otherwise fall short.[6] This vision epitomizes the promise of "augmented intelligence" – seamlessly combining human knowledge and creativity with machine scale and precision to create a system greater than the sum of its parts.[7]

The promise of "augmented intelligence" – seamlessly combining human knowledge and creativity with machine scale and precision to create a system greater than the sum of its parts.

For the U.S. and Australian ICs, we argue that the AI avenue with the highest potential impact is human-machine teaming (HMT), which could revolutionize the efficiency, scale, depth, and speed at which analytic insights are generated. AI-HMT promises to elevate analytical capabilities by creating feedback loops that allow analysts and algorithms to benefit from the strengths of the other. In the national intelligence field, pilot projects deploy AI for bespoke analytical functions, experiments, and other discrete tasks, though not yet at scale or integrated across the full analytic workflow.[8]

With continuing breakthroughs, the integration of expansive AI capabilities into the broader craft of intelligence analysis seems imminent, but integrating these tools into intelligence operations presents a unique set of challenges. In a high stakes environment, intelligence services, like those in the United States and Australia, must maintain a very high bar for the quality and accuracy of the assessments they produce; therefore, they have low tolerance for new tools, inaccurate information, or recommendations that conflict with legal or ethical guidelines. In addition, it is important that there be some level of cooperation and coordination between friendly intelligence services when it comes to the deployment and integration of AI tools. If friendly services deploy these tools at different speeds or even deploy different types of tools, it may complicate future collaboration.

---

[6] Ajay Agrawal, et al., Prediction Machines: The Simple Economics of Artificial Intelligence, Harvard Business Review Press at 53–54, 65–69, 102 (2018); Michael Polanyi, The Tacit Dimension, University of Chicago Press at 4 (2009); David Autor, Polanyi's Paradox and the Shape of Employment Growth, National Bureau of Economic Research at 8 (2014).

[7] James Wilson & Paul R. Daugherty, Collaborative Intelligence: Humans and AI Are Joining Forces, Harvard Business Review (2018).

[8] Examples include: NGA's partnership with Impact Observatory to produce AI-generated maps at almost real-time, NGA's Source Maritime Automated Processing System (SMAPS) Program, IARPA's "REASON" Program to develop an intelligence analysis assistant plug-in, and the CIA's deployment of GenAI chatbot. Jeanne Chircop, AI Revolutionizes Mapping Updates, Accuracy, National Geospatial Intelligence Agency (last accessed 2024); NGA Puts Machine Learning to Work to Speed Mission, Further Research, National Geospatial Intelligence Agency (2022); REASON: Rapid Explanation, Analysis and Sourcing Online, Intelligence Advanced Research Projects Activity (last accessed 2024).

The Intelligence Analysis Mission and Expectations of Generative AI

Workshop participants saw opportunities for different types of AI to augment intelligence analysts and, in some cases, automate several parts of their work across the analytic workflow. It should be noted that intelligence services have a long history of using AI – in the form of machine learning algorithms – as important elements of their enterprise information technology stacks. They are also actively testing and experimenting with the current generation of GenAI and other AI models.[9]

We expect to see intelligence services further deploying more sophisticated AI systems into production over the next 12 to 18 months – just as we expect to see LLMs grow and artificial intelligence to become more sophisticated over that same time period. Given that these tools and capabilities will grow at an exponential speed, there is always a risk that ICs will struggle to keep pace. Therefore, over the medium-term, intelligence communities must develop focused, yet flexible, strategies for their implementation.

In the simplest terms, intelligence analysis is intended to discern foreign actors' intentions and actions by warning and informing policymakers of changes in the geostrategic environment that are likely to affect their sense of national interests. It can also characterize what those changes might mean over the near-, mid-, and long-term. The changes can be one-off events (e.g., a bi- or multilateral diplomatic summit, an election, a military acquisition decision) or trends (e.g., rising tensions between two or more countries, the implementation and refinement of a political agenda, a military campaign). Contextualizing the event, trend, and its probable effects in light of available information is a critical subtext of analytic missions.

In order to provide these insights, intelligence analysts work through a cyclical process – the analytic workflow – where new information is synthesized and integrated into analytic products for customers, who in turn provide feedback that guides what new information and insights are

[9] Frank Konkel, The US Intelligence Community is Embracing Generative AI, Government Executive (2024); Brandi Vincent, CIA to Investigate How Generative AI (like ChatGPT) Can Assist Intelligence Agencies, DefenseScoop (2023); Peter Martin & Katrina Manson, CIA Builds Its Own Artificial Intelligence Tool in Rivalry With China, Bloomberg (2023).

required. Artificial intelligence has the potential to automate many parts of this workflow. While analysts are often the key driver pushing through each stage of the cycle, there are other relevant stakeholders. Analysts must liaise with data collectors, including those responsible for open-source and clandestinely acquired information. Similarly, disseminating analyses to customers and consumers occurs through a range of systems and people, from a secure website through to a briefer assigned to support a senior decision-maker for a sustained period of time. As a result, introducing LLMs into the analytic workflow could have spillover effects into the larger intelligence and policymaking apparatuses, especially if various stakeholders have to coordinate their use of technology in order to uphold these relationships.

Core Requirements: Transparency, Explainability, and Accountability

In a typical analytic product, a central argument is bolstered by a small number of strong pieces of evidence. Under the current system, human analysts are largely responsible for manually collating and weighing evidence, which increases the likelihood that a key piece of evidence, either one that adds important nuance or stands in conflict to the central argument, will be missed. The ability of LLMs to hold more data, and change the weight of that data, means that an analyst who is teamed with an AI will be able to draw effortlessly on the most recent, relevant, and reputable supporting information.

The ability of LLMs to hold more data, and change the weight of that data, means that an analyst who is teamed with an AI will be able to draw effortlessly on the most recent, relevant, and reputable supporting information.

Limitations of Existing AI Models for Analysis

GenAI models are constantly gaining in sophistication, with each iteration of a model making crucial improvements over previous versions. At the same time, each iteration also creates new vulnerabilities about which intelligence professionals must be aware. Given the various components that go into these models, they will always have their inherent limitations, which not only necessitate human oversight but also technical safeguards.

- **Early LLMs struggled with factual grounding.** As statistical models focus on sequence predictions rather than factual accuracy, some current LLMs can generate seemingly plausible but wholly invented statements ungrounded in reality. This tendency to "hallucinate" stems from factors like misunderstanding content, limited training data, over-reliance on statistical likelihoods rather than verified evidence sources, and a lack of mechanisms to confirm accuracy. For intelligence analysts, hallucinations could critically misguide high-impact assessments if not caught. Developers are working on mitigation techniques including prompt fine-tuning and algorithms that alert users to possible hallucinations.[10]

- **LLMs have limited reasoning capacities.** Despite advances in natural language processing, most large language models still struggle with complex causal analysis, logical deduction, analogical mapping between scenarios, or mathematically modeling key relationships underlying events, even with the best data available. When policymakers turn to intelligence analysts for assessments, it is vital that analysts explain how they drew their conclusions. Hybrid approaches that combine statistical learning with compositional reasoning, causal diagrams, and other frameworks could better elicit explanatory rationales within AI systems.

- **LLMs risk pre-existing bias amplification.** For analysis to be of high quality, it must be grounded in an appropriate regional and generational context. Large language models trained on limited societal texts may indirectly propagate and even amplify historical biases. For all-source intelligence application, backwards transmission of disproportionate representations or toxic associations around factors like race, gender, ethnicity and culture could corrode social equity standards vital to public service integrity. Establishing proactive algorithm auditing processes for fairness, inclusion and value alignment tailored to the unique data interoperability and policy notification needs of intelligence communities will help avoid marginalization.

[10] Imama Shezad, Beyond Traditional Fine-Tuning: Exploring Advanced Techniques to Mitigate LLM Hallucinations, Hugging Face (2024); Sebastian Farquhar et al., Detecting Hallucinations in Large Language Models Using Semantic Entropy, Nature (2024).

However, one key challenge that comes with deploying LLMs in intelligence analysis is the opaqueness that comes with LLM outputs. LLMs intrinsically function as “black boxes,” obscuring some of the detailed reasoning that has led to the model's output, which poses a problem for analysts and policymakers alike. Robust accountability and maintaining policymaker and public trust are of utmost importance within the United States and Australian intelligence services, given their unique responsibilities and access to sensitive information. If policymakers cannot understand how and why certain evidence was used, the analysis loses credibility.[11]

Therefore, the ICs must ensure that basic standards for transparency and explainability are designed in conjunction with the deployment of LLMs. For the U.S. IC, these standards must adhere to ODNI requirements, such as ICD 203 (“Analytic Standards”) and ICD 206 (“Sourcing Requirements for Disseminated Analytic Products”).[12] The two ICDs mandate detailed sources and analyst confidence in those sources.

Explainable AI (XAI) is one tool that could ensure that LLM outputs meet these standards. XAI helps in the generation of insights that are justifiable, trustworthy, and foster trust in AI’s use in intelligence. XAI aims to demystify AI decisions by providing two levels of explanation: global explanations that describe the system's overall workings, and local explanations that detail the rationale behind specific decisions. Several research initiatives, such as IARPA's REASON, BENGAL, BETTER, and HIATUS programs, as well as DARPA's XAI program, have been launched to help develop and implement XAI in the intelligence domain.[13] They seek to develop novel technologies that enable intelligence analysts to improve evidence and reasoning in analytic reports, identify and mitigate bias in generative AI systems, improve the accuracy and explainability of information extracted from unstructured text data, and develop explainable models for attributing authorship to anonymous or pseudonymous text data.

In the absence of full explainability, however, ICs may need to reframe the issue to consider and understand AI as a source itself. AI's ability to identify patterns that cannot be manually verified through non-AI analysis poses a dilemma: to deploy AI and risk poor decision-making based on analysis not subject to human verification or to risk a potential intelligence failure by not deploying

---

[11] See Appendix C for further explanation of the difference in perspectives between AI developers and intelligence practitioners.

[12] ICD 203, Analytic Standards, Office of the Director of National Intelligence (2022); ICD 206, Sourcing Requirements for Disseminated Analytic Products, Office of the Director of National Intelligence (2015).

[13] Rapid Explanation, Analysis and Sourcing Online (REASON) Program, Intelligence Advanced Research Projects Activity (last accessed 2024); Bias Effects and Notable Generative AI Limitations (BENGAL) Program, Intelligence Advanced Research Projects Activity (last accessed 2024); Better Extraction from Text Towards Enhanced Retrieval (BETTER), Intelligence Advanced Research Projects Activity (last accessed 2024); Human Interpretable Attribution of Text Using Underlying Structure (HIATUS), Intelligence Advanced Research Projects Activity (last accessed 2024); David Gunning & David W. Aha, DARPA’s Explainable Artificial Intelligence (XAI) Program, AI Magazine at 44-58 (2019).

the AI that might identify certain patterns in the first place.[14] This dilemma raises the central question of the extent to which transparency should be sacrificed for decision advantage. In such cases, it may be necessary to treat AI as a source of intelligence, similar to human sources, and assess its reliability and credibility based on its past performance and the context in which it operates.[15] It might also require human spot-checking of randomly selected inputs or using other sources to corroborate the findings and insights that AI provides. This approach would require the development of new frameworks and methodologies for evaluating AI systems as intelligence sources, considering factors such as their track record, the quality and relevance of their outputs, and their potential biases or limitations.[16]

[14] Cynthia Rudin, Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead, Nature Machine Intelligence at 206-215 (2019).

[15] Ben Buchanan, et al., Automating Cyber Attacks: Hype and Reality, Center for Security and Emerging Technology (2020).

[16] Umang Bhatt, et al., Explainable Machine Learning in Deployment, Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency at 648-657 (2020).

AI Today: Thinking About Applications Across the Analytic Workflow

The ICs' experimentation with AI should focus on identifying opportunities for augmentation, automation, and knowledge-sharing across the entire analytic workflow: triaging new information, conducting research and organizing data, conceptualizing analytic products, drafting analytic products, performing tradecraft checks, and engaging with policy customers. NLP techniques can be used to extract relevant entities and relationships from large volumes of unstructured text data, while computer vision techniques can be applied to identify objects and activities of interest in imagery and video data. These techniques – which were more rudimentary during the time of the Afghan and Iraq wars – are being used to greater success in Ukraine today[17] but further developments are essential.

GenAI has the potential to be a crucial support across the analytic workflow for all-source analysts who primarily work with unstructured texts. However, the participants in our conversations all agreed on a central point: for now, GenAI would not replace human analysts. The models currently have the capability to accept offloaded menial, low-value tasks, but the "expert intuition" of human analysts in identifying novel connections between disparate pieces of information and explaining the meaning and potential implications of those connections was seen as existing beyond current capabilities for AI; however, as the models get more sophisticated and meet increasingly difficult performance benchmarks, they will have the capacity to accept more intensive tasks from the analytic workflow.

Participants . . . agreed on a central point: for now GenAI should not replace human analysts . . . the "expert intuition" of human analysts in identifying novel connections between disparate pieces of information and explaining the meaning and potential implications of those connections was seen as existing beyond current capabilities for AI.

[17] Dr. Charlie Winter, et al., Artificial Intelligence, OSINT and Russia's Information Landscape, Alan Turing Institute, Center for Emerging Technology and Security (2023).

The primary goal of integrating AI into intelligence analysis should be to enhance the efficiency and effectiveness of the analytic process, enabling analysts to generate high-quality, actionable insights more rapidly and comprehensively. This goal encompasses two key objectives: accelerating time to insight and augmenting analytic depth and breadth. Accelerating time to insight involves leveraging AI technologies to help analysts quickly identify, prioritize, and synthesize relevant information from vast and diverse data sources. For example, NLP techniques can significantly reduce the time analysts spend on manual data processing and allow them to focus on higher-level analysis and interpretation.

In addition to accelerating insight generation, AI can also enhance the depth and breadth of analysis by enabling more comprehensive exploration of complex, multi-dimensional data sets and uncovering hidden patterns, trends, and anomalies that may be difficult for humans to detect. For instance, graph analytics techniques can be applied to identify key persons of influence, communities, and information flow patterns within large-scale social networks, providing valuable context for understanding and predicting social dynamics and potential security threats.

Deriving Insights from Large Volumes of Data at Speed

Analysts monitor and triage new information on a daily basis. In doing so, they are looking for information and insights that – in their experience and expert opinion – support an existing line of analysis, suggest a change that could alter the trajectory of an existing line of analysis, or information that – again, in their experience and expert opinion – is interesting and potentially important. The changes require contextualization: analysts are trying to forensically identify and explain the precipitating events and trends and are simultaneously reweighting the information that informs their understanding of the issue(s) they follow, particularly as they revisit historical information that they might have undervalued or overvalued at the time of its arrival.
Triaging information implicitly involves dynamically weighting new information against known information and existing lines of analysis, which when done manually is a time-intensive process.

Both technologists and practitioners saw opportunities to insert AI systems to support human analysts in the basic tasks associated with triaging and monitoring incoming information. The opportunities hinged on GenAI's ability to more quickly detect patterns in large volumes of unstructured text or identify anomalous spikes in reporting that might indicate an emergent situation. ML algorithms could be a replacement for a traditional search as well as a question-answering platform. For example, LLMs could prove to be better suited than an analyst when it comes to consolidating and weighing copious amounts of potentially relevant information without cherry picking based on pre-existing biases. For analysts, identifying and incorporating information that challenges or shifts existing analytic lines is paramount but such information can easily be lost among copious amounts of data. GenAI could support finding this "needle in a haystack."

Practitioners cautioned that too much precision on an AI system's part (i.e., the AI screens out all but the most relevant content) runs counter to the value of triaging information: expertise was seen as a byproduct of triaging information. Data that was only tangentially relevant allowed analysts to broaden their understanding of the dynamics around the issues they followed – and to identify undervalued connections or emergent trends of potential importance.

Conceptualizing Analytic Products

Analytic products can vary highly in length and content, and analytic capacity depends on the resources available across the analytic workflow (e.g., volume and variety of information available, number, type, and experience of analysts available, etc.). While a late breaking, high impact event might require a quick, highly focused analytic product, that may not be an appropriate vehicle for all analytic insights. Similarly, what decisionmakers and policymakers find important and interesting is not always covered by traditional media or from a perspective that speaks to decision-maker needs. Analytic GenAI could assist human analysts by helping track and prioritize product ideas based not just on detecting potentially meaningful changes but on the cyclical nature of some work (e.g., election coverage) or anticipated customer interest (e.g. what sort of information and insights does a policymaker or decision-maker wish they had in the run-up to or during a bilateral meeting?).

In addition to analytic line reviews, LLMs could also be useful in ideation and brainstorming, proposals of new products, and structured analytic techniques.

Once analysts have new information that they determine impacts existing analytic lines enough to warrant new production, the next step is to outline that new product including determining the central argument, the evidence needed to defend the claim, and the analytic logic. Conceptualizing analytic products can be an intellectually intense process, but LLMs can help streamline many processes. For example, intelligence analysts are trained to be on guard against cognitive and logical fallacies, but at the same time it can become easy for analysts to become wedded to existing analytic lines. As a result, they can become susceptible to examining evidence with a biased view that tips in favor of existing lines. There are already several AI systems that purport to detect fallacies on the market, and incorporating this functionality into the models that support analysts could further improve the quality of analytic products by constantly assessing the robustness of existing analytic lines.[18] In addition to analytic line reviews, LLMs could also be useful in ideation and brainstorming, proposals of new products, and structured analytic techniques.

[18] See e.g., Fallacy Detection, There's An AI for That (last accessed 2024).

Scenario Generation

LLMs can be useful in the ideation and brainstorming stage, including in scenario generation. While intelligence analysts spend significant amounts of time engaged in considering current events and contemporary trends, they are typically also thinking about the future implications of those events and trends. While an analyst might deem a new piece of information as important to customers, it may not be entirely evident what the context of its corresponding analytic product should be. GenAI systems could support this function by providing responses to user queries (e.g., "What are the ten most likely reasons for civil unrest to occur in [country name]?", "What are the top five policy changes likely to occur if the ruling coalition falls out of power in [country name]?").

At this level, the AI model does not need to be as explainable as much as its responses need to be plausible. This functionality could be extended deeper into analytic workflows by integrating scenario generation with search – to surface documents that support or contextualize the scenarios identified by the AI – and modeling and simulation – to allow the analyst to play with variables related to the scenario (e.g., "What if the head of state's approval rating dipped to less than 30 percent?"). Generating data-driven hypotheses and scenarios to help analysts anticipate and prepare for potential future developments.

Structured Analytic Techniques

Structured analytic techniques are an important tool in an analyst's arsenal because they often compel analysts to draw out their logic by putting it through rigorous testing. In the aftermath of the intelligence failures of the early 2000s, the Central Intelligence Agency and the wider U.S. IC looked at ways to improve their analytic tradecraft. Their response, in part, was a more widespread use of structured analytic techniques.[19] Techniques include analysis of competing hypotheses, where several hypotheses for an outcome are simultaneously tested to see which hypothesis has the most consistent evidence, and the key assumptions check, in which taken-as-fact assumptions that underpin judgments are carefully scrutinized.

While the value of these techniques has been challenged, they sometimes address a broader range of potential outcomes and implications than did other analyses.[20] In particular, they allow

[19] Structured analytic techniques give analysts a framework that guides their thinking. These techniques largely fall into three categories: diagnostic (which help analysts critically examine their assumptions or how they got to their central argument), contrarian (which help analysts defend their argument by unpacking counterarguments), and imaginative (which help analyst think through a wider range of options, including unlikely possibilities). For more information, see A Tradecraft Primer: Structured Analytic Techniques for Improving Intelligence Analysis, Central Intelligence Agency (2009). See also A Tradecraft Primer: Basic Structured Analytic Techniques, U.S. Defense Intelligence Agency (2008).

[20] Stephen Artner, et al., Assessing the Value of Structured Analytic Techniques in the U.S. Intelligence Community, RAND Corporation (2017).

analysts to think with longer time horizons; however, they can be time consuming, which is why many analysts might choose to eschew them if they do not deem them necessary. In the context of AI-assisted intelligence analysis, AI systems should incorporate or accommodate standard analytic methodologies and analytic techniques and work with the analyst to determine which (if any) might best support their work.

Creating Analytic Products

There are already several AI systems that provide writing support to users, ranging from GenAI models that will draft complete pieces based solely on the content of their data holdings, from user cues and other inputs, to co-piloting configurations in which an AI system would guide and support the human analyst as they create analytic products. Given the variance in cognitive and work styles, the question is, "How might analysts adjust the type and level of support that an AI model gives them as they draft a product?" This is an area ripe for experimentation. Intelligence agencies should systematically – and continuously – map user requirements to support effective AI deployments that build the case for further adoption.

Sourcing, Classification, and Tradecraft Checks

Analytical products go through two distinct processes: coordination and executive review. Unlike academia where papers are viewed as the work of the author(s), analytic products are ultimately the work of either a particular organization or in some cases, the entire IC. Few analytic products are the work of a single analyst or are limited to a single analytic discipline: a piece looking at domestic politics in a foreign country, for example, is likely to encompass foreign politics, economics, and demography. Coordination across teams, groups, and organizations is intended to ensure that all current and relevant analytic lines are accurately conveyed. It can also serve the purpose of helping IC agencies deconflict to make sure that multiple similar products are not being independently drafted.

After coordination, analytic products then go through one or more layers of editorial review typically performed by supervisors, managers, professional editors, and – depending on the product and the audience for the product – executives. The objective of the review process is to make sure that an analytic product not only represents the most insightful and useful thinking in light of the information available to analysts at the time but is written in such a way that the articulated logic is concise yet clear. Intelligence professionals do not want to put policymakers in a position where analysis can be interpreted in more than one way.

Editing and review tends to be a managerial or professional function for supervisors and managers. The process is not only an opportunity to ensure that the draft analytic product is consistent with IC standards and style guides or put in the context of a line of analysis, but also to develop the analytic and communication skills of their staff. Professional editors tend to focus heavily on stylistics, though they might be the first people to read a piece who do not have deep subject matter expertise on the issue covered in the paper, bringing a fresh perspective to the text. This was where workshop participants first discussed the bureaucratically disruptive nature of GenAI: participants saw editing and review as an inefficient and time-consuming part of the analytic process (even as it was described as a core function) and AI as an opportunity to streamline the analytic production process. Group conversation revolved around what rebalancing workloads might mean for resourcing in the analytic arms of intelligence services.

Publicly available GenAI models can mimic the style of an author and draft and review a document based on the standards and principles set out in a style guide (e.g., Strunk and White's *Elements of Style*, *The Chicago Manual of Style*, CIA's Style Manual and Writers Guide, etc.). For enterprise applications in intelligence settings, a commercial AI might need to be trained on the organization's style guide or articulate in the prompt that a draft is to be reviewed and edited to a specific set of guidelines. Not only would this ensure consistency in insights and articulation of those insights across products, but it would also save hours of manpower for analysts and managers alike.

Customer Engagement and Feedback

Central to this phase of the analytic workflow is understanding national interests in terms of an organization's mission(s) and customer responsibilities and priorities. A core competency of analysts is knowing the audience, the audience's informational needs, and the threshold for warning and informing the audience about changes that are likely to affect their thinking and work.

Analysts typically have a clear sense of the intelligence question they are trying to answer for which customer or customers when they craft an analytic product. While dissemination of analytic products is often depicted as one of the final steps of the intelligence cycle, customer engagement and feedback are important: direct and indirect interactions with customers at all levels of policy- and decision-making processes help ensure that analysts are aware of what their audiences are thinking about and are working on.[21] Analytic independence is further protected through legal and structural mechanisms, such as the ODNI Analytic Ombuds in the United States and Inspector-General of Intelligence and Security oversight and the Office of National Intelligence Act 2018 in Australia.[22]

Workshop participants discussed to what degree AI could, or should, be used to support customers. Technologists tended to see GenAI as a means of enabling deeper customer support

---

[21] The Six Steps in the Intelligence Cycle, Office of the Director of National Intelligence (last accessed 2024).

[22] Objectivity, Office of the Director of National Intelligence (last accessed 2024); No. 155, Office of National Intelligence Act 2018, Commonwealth of Australia (2018).

as a natural outgrowth of allowing customers to search for finished analytic products. Practitioners were wary because of some GenAI behaviors, specifically LLM-fueled hallucinations: they were concerned that while AI responses might be drawn from finished intelligence, the generated response would need to be held to the same standards as human-generated content. This is not to say that human analysts never make mistakes, but the editing and review process is there to serve as a series of checks to ensure that analysis provided to customers is as clear, concise, and accurate as possible. If anything, practitioner concerns were rooted in historical intelligence failures, which feature prominently in their training. The conversation concluded with the recognition that just as analysts will need to be trained on the use of GenAI as a tool, so will customers. This is likely to be a more challenging proposition because of the differences in roles and responsibilities between the two parties.

This paper still assumes a narrow focus on traditional customers of IC products, but GenAI could also be helpful if we were to expand to encompass a broader spectrum of stakeholders. For example, private sector actors, including CEOs of technology companies, play critical roles in geopolitics and intelligence. GenAI could support analysts' efforts to draft or reframe products for these audiences with redactions and a heavier use of commercial sources. Doing so would be in line with intelligence communities' efforts to write for maximum utility, which is currently time consuming because analysts lack the training to write for that audience.

Looking Ahead: The Coming Wave of AI Advancements

A key challenge with AI is the rapid pace at which the technology is being developed: the "current paradigms" of today are quickly eaten by new ones tomorrow. Just ten years ago, no machine could reliably provide language or image recognition at a human level, but now they routinely outperform humans on benchmark tests. Today's cutting edge LLMs (ChatGPT 4o; Llama 3.1; Claude 3) models are more powerful than the first generation of systems that were introduced less than two years ago. These rapid advances in AI capabilities have made it possible to use machines in a wide range of new domains.[23] For example, when you book a flight, it is often an artificial intelligence that decides what you pay, and an AI system assists the pilot in flying you to your destination.[24] AI systems also increasingly determine whether you get a loan, are eligible for welfare, or get hired for a particular job.[25] AI systems help to program the software you use and translate the texts you read, and virtual assistants, operated by speech recognition, have entered many households over the last decade.[26] AI models determine what you see on social media, which products are shown to you in online shops, and what gets recommended to you on online television platforms. Increasingly they are not just recommending the media we consume but based on their capacity to generate images and texts, they are also creating the media we consume.[27]

A key challenge with AI is the rapid pace at which the technology is being developed: the "current paradigms" of today are quickly eaten by new ones tomorrow."

The models we are likely to see three years from now will be even more powerful, with some industry experts even predicting the advent of

[23] Max Roser, The Brief History of Artificial Intelligence: The World Has Changed Fast — What Might Be Next?, Our World In Data (2022).

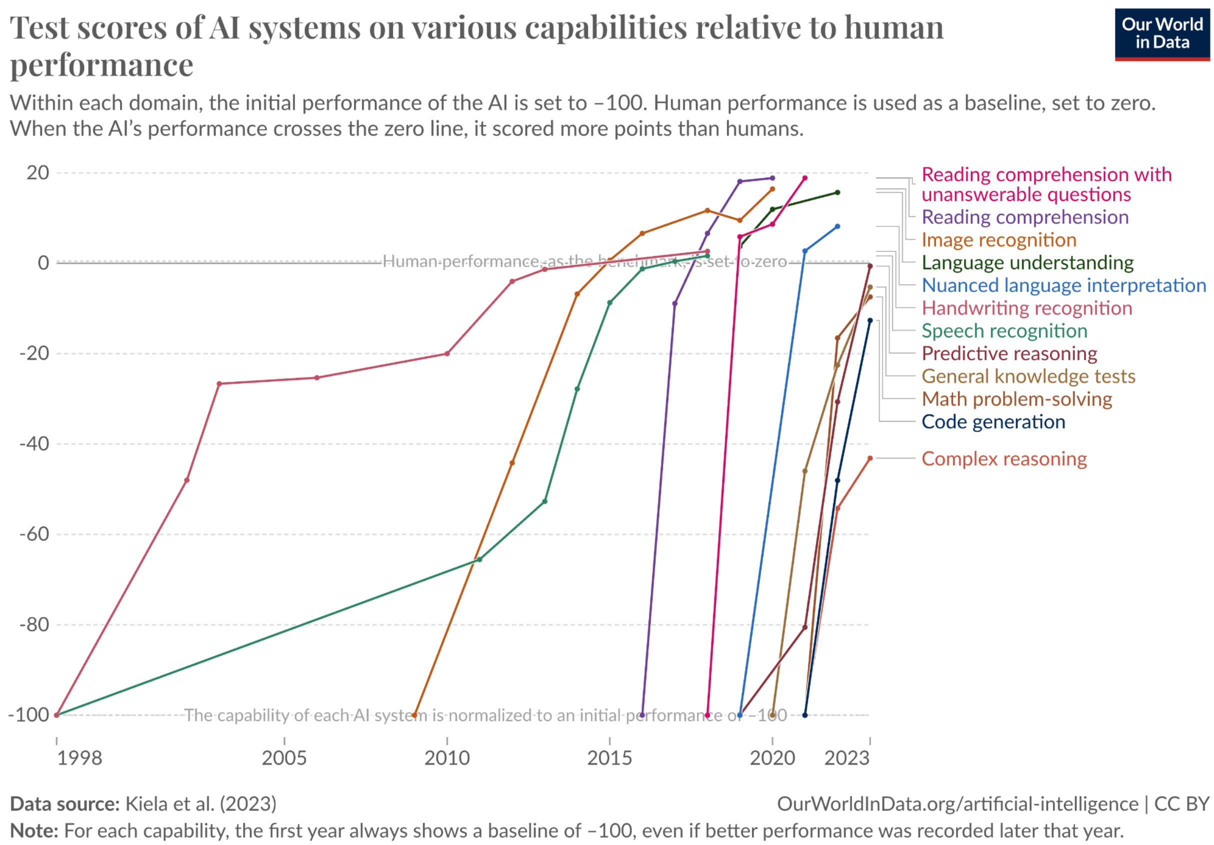
[24] Angus Whitley, AI Knows How Much You're Willing to Pay for Flights Before You Do, Bloomberg (2022); Artificial Intelligence: Capitalizing on the Value of Data, Airbus (last accessed 2024).

[25] Aaron Klein, Reducing Bias in AI-Based Financial Services, Brookings Institute (2020); Michele Gilman, AI Algorithms Intended to Root Out Welfare Fraud Often End Up Punishing the Poor Instead, The Conversation (2020); Jeffrey Dastin, Insight - Amazon Scraps Secret AI Recruiting Tool That Showed Bias Against Women, Reuters (2018).

[26] Thomas Dohmke, GitHub Copilot X: The AI-Powered Developer Experience, GitHub Blog (2024); Will Knight, I Tested a Next-Gen AI Assistant. It Will Blow You Away, Wired (2024).

[27] Matthew Hutson, Robo-Writers: The Rise and Risks of Language-Generating AI, Nature (2021).

AGI by as early as 2027.[28] While precisely forecasting the progress of AI presents innate difficulties, technologists who participated in the workshop urged the ICs to anticipate further rapid advances in GenAI capabilities and to make preparations now for the type of systems that will be available in three to five years, not what they see today. In their view, designing actions based on the GenAI available now risks letting the ICs fall further behind, ceding advantage to foreign rivals and reducing the ability of ICs to support U.S. and Australian policymakers in the future.



Graphic Source.[29]

Workshop participants identified several likely technical advancements in GenAI over the next 3-5 years that will impact the ICs' analytic mission:

---

[28] Leopold Aschenbrenner, Situational Awareness: The Decade Ahead (2024). But see also, What Is Artificial General Intelligence?, McKinsey & Company (2024) ("Artificial general intelligence (AGI) is a theoretical AI system with capabilities that rival those of a human. Many researchers believe we are still decades, if not centuries, away from achieving AGI.").

[29] Douwe Kiela, et al. – with minor processing by Our World in Data. Test Scores of AI Systems on Various Capabilities Relative to Human Performance, Our World in Data (2023).

- **Larger LLMs.** LLMs are typically measured in terms of the number of parameters – the number of weights learned during training – it possesses. ChatGPT contained 1.5 billion parameters when it was released in November 2022. In contrast, ChatGPT 3 had 175 billion parameters when it was released in June 2020, and while there is no public disclosure of ChatGPT 4's size, it is estimated to be around 1.8 trillion.[30] As research into LLMs continues, next-generation models with trillions of parameters could provide several benefits, including accommodating exponentially more information and training data. Larger LLMs are likely to have several positive effects for intelligence analysts including: accommodating more information; an improved ability to understand user queries and generate more sophisticated responses to those queries; and the ability to perform more advanced analytic tasks, such as sentiment analysis and predictive analysis.

- **Greater Context Length.** Ongoing innovation continues to expand feasible context length for informing LLMs, with some labs announcing capacities for hundreds of pages. Google recently announced that its Gemini Pro will have a 128,000 token context window,[31] allowing it to accommodate over 400 pages of input. If that research is borne out, intelligence analysts may someday have granular control over model behavior by providing extensive background context, like defining a detailed persona or supplying collections of documents as framing. The ability to reference detailed and specific case files or extensive profiles could support uniquely customized LLMs for specialized applications. However, effectively processing such large context lengths poses significant software and hardware scaling challenges that are yet to be resolved.

- **External Memory and Access to External Information.** Incorporating credible external sources of information could hugely reinforce internal assessments by linking to expanded evidence beyond what any standalone model contains. Seamlessly integrating externally available information across firewalls, air gaps, and access controls in highly secure environments poses complex engineering challenges still requiring pioneering breakthroughs. Future hardware and software innovations may one day allow AI systems like LLMs to optimize secure retrieval speeds across vast supplementary datastores and knowledge bases, while upholding strict access protocols. If key obstacles around retrieval latency, connectivity infrastructure, and the maintenance of governance policies timely and relevant to the technology at hand are solved over time, purpose-built model architectures could eventually access select repositories of domain knowledge and

[30] Bernard Marr, A Short History of ChatGPT: How We Got To Where We Are Today, Forbes (2023); Maximilian Schreiner, GPT-4 Architecture, Datasets, Costs, and More Leaked, The Decoder (2023).

[31] Sundar Pichai & Demis Hassabis, Our Next-Generation Model: Gemini 1.5, Google (2024)

document libraries, to tap relevant real-time information. Such models would offer greatly enhanced contextualization, explanation, and ability to rapidly "ground" analysis against trusted external information sources. But making this a broadly usable reality depends on finding solutions for walled-off systems to judiciously leverage data beyond their own content.

- **Grounded Search.** Grounded search systems that connect internal memories with real-time access to curated external databases and credible internet sources could one day enable intelligence analysts to query and synthesize diverse supplementary information beyond their internal datasets. Future systems that supplement internal assessments with relevant external facts and statistics in a secure and efficient manner may eventually improve IC analysis of complex geopolitical situations. Filtering out irrelevant tangents and evaluating the legitimacy of external sources with the degree of confidence needed to make intelligence assessments remains an area of active research.

- **Better Reinforcement Learning with Human Feedback.** Interactive human-in-the-loop reinforcement offers promise for analytical tools that could continually self-improve assessments of shifting real-world scenarios based on qualitative analyst critiques over time. Research is exploring how to gauge feedback consistency and mitigate potential unintended skews. If solutions emerge, capabilities like tuning language suggestions based on stylistic edits or flagged mischaracterizations could prove transformative.

- **Copiloting.** Copilot configurations have the potential to someday significantly boost productivity by combining analyst judgment with automated drafting for finished analytic products. Key challenges in this space include determining the optimal and preferred divisions of labor between the AI and the analyst and developing the mechanisms to correctly identify inaccuracies and offer corrections in a user-friendly manner. If realized, copiloting could increase analyst bandwidth by reducing the time needed to produce the first drafts of analytic products.

- **Reflexion.** Reflexion is an AI architecture that uses feedback and self-reflection to improve the quality of its responses to user queries. For instance, a system using the Reflexion architecture will use its initial response to formulate complementary questions that will improve the completeness and accuracy of its response to a user query. More advanced future architectures could proactively anticipate and mitigate limitations or inconsistent behaviors by conjecturing diverse scenarios.[32]

[32] Noah Shin, et al., Reflexion: Language Agents with Verbal Reinforcement Learning, Conference on Neural Information Processing Systems (2023).

- **Modeling and Simulation.** In the future, integrated modeling and simulation capacities could enable intelligence analysts to quickly and easily explore an array of hypothetical scenarios to better understand the implications of likely and possible outcomes. This would allow them to better inform policy planning by assessing plausibility before events unfold. Computational forecasting aids might someday run conjectured decision chains based on analyst- or agency-developed protocols for emerging threat conditions. For instance, by algorithmically simulating a spectrum of potential overseas crisis responses, key indicators like cost, mobility rates, and public sentiment could one day project under each contingency to accelerate preparedness by uncovering blindspots.

- **AIs that Perform Together as a "Community of Experts."** As the architectures underlying generative AIs become more specialized and sophisticated, so-called "conductor" AI systems could be developed that orchestrate the activities of several specialized AI models to enable complex tasks across multiple domains. For example, a human operator could task a "conductor" AI to assess the likely future trajectory of a foreign conflict by querying a suite of sub-AIs that are specialists in military capabilities and doctrine, diplomacy, foreign leader behavior, and food security to arrive at an expert assessment.

- **Compound Systems.** Compound systems – defined by their use of multiple interacting components – would expand upon a single model to one that can execute a variety of tasks by adding components such as planning, detection, and data labeling. For example, a Retrieval-Augmented Generation (RAG) can be particularly useful to analysts who are often simultaneously retrieving information through relevant documents while generating their own texts. By having several components, compound systems can specialize in several specific tasks while remaining more flexible, resilient, and scalable than LLMs; however, the downside will be the larger upfront investment that is required to develop them.[33]

- **"AI Agents."** AI developers are creating AI systems that can interact with their environment, collect data, and use the data to perform self-determined tasks to meet predetermined goals.[34] Humans set the goals, but an AI agent independently chooses the best actions it needs to perform to achieve them. For example, an analyst might one day task their AI to execute a collection strategy to increase the amount of data available on a foreign military facility of interest. The AI agent would then autonomously identify the critical information gaps and then issue tailored requirements to other IC units responsible

[33] Matei Zaharia, et al., The Shift from Models to Compound AI Systems, Berkeley Artificial Intelligence Research (2024); Maithra Raghu, et. al., Does One Large Model Rule Them All?, Maithra Raghu Blog (2023); Tehseen Zia, A Silent Evolution in AI: The Rise of Compound AI Systems Beyond Traditional AI Models, Unite.AI (2024).

[34] See e.g., What are AI Agents?, AWS (last accessed 2024).

for managing overhead satellite imagery and signals intelligence systems and draft cables to the embassy in the target country to guide human collectors.

Recommended Actions

Powerful GenAI technology is likely to evolve within the next three years and Australian and U.S. intelligence leaders will need to make preparations now for this new reality. Several participants urged the ICs to take a more strategic approach to AI adoption with the broader aim of using AI to transform the IC analytic culture away from traditional, reactive methods to a faster, more anticipatory, data-driven future. Beyond acquiring and experimenting with LLMs, which already is underway, the two ICs should use this time to take stock of their analytic workflows and begin using AI to make each stage of the analytic process more efficient, as outlined above.

Several participants urged the ICs to take a more strategic approach to AI adoption with the broader aim of using AI to transform the IC analytic culture away from traditional, reactive methods to a faster, more anticipatory, data-driven future.

To avoid remaining perpetually behind the curve on the pace of AI technological development, analytic managers should shift their focus from what GenAI can do today and instead make reasoned bets on what GenAI will be able to deliver within the next 3-5 years. In addition to pressing their home agencies to acquire and integrate AI-related infrastructure (particularly advanced compute capabilities, access to cutting-edge commercially available GenAI models and algorithms, and secure data storage), we recommend that U.S. and Australian analytic managers undertake the following five actions:

1. Design for Continuous AI Model Improvements
2. Start Automating Portions of the Analytic Workflow Now
3. Build Human-Machine Analytic Teams
4. Create AI-Ready Training and Incentive Structures for the Analytic Workforce
5. Collaborate to Develop Shared U.S.-Australian Standards and Evaluation Metrics

Action 1: Design for Continuous AI Model Improvements

With the expected exponential growth of LLMs, the ICs cannot only look only to the current technological state-of-play but must also anticipate GenAI's future trajectory over the course of the next five, ten, or twenty years. We expect GenAI models will be updated on a near-continuous basis, with successive iterations adding important new capabilities. The ICs should adapt their technology acquisition processes – and budgets – to enable for multiple annual updates. This will include accounting for larger LLMs, expansions in context lengths, and further developments in more sophisticated systems like compound and agentic systems.

Moving beyond experimentation means investing in strategic, systemic data and technology management to support interoperable AI technologies, within and across ICs. The second workshop identified key aspects of a strategic "AI-ready" approach to data, including that it would: assess whether data sources are legitimate and credible, define a minimum data retention period, assess data for bias, weight it according to accuracy, and ensure it conforms to established data formats and standards. Where AI models are making high risk or high impact assessments or providing complex decision support, AI-ready agencies would also ensure that training data and settings are archived and available to test for accountability, bias, and decision-making repeatability.

IC agencies should continue to advance toward federated data architectures – and do more to leverage open-source information – to take full advantage of AI tools.[35] Judiciously deploying LLMs and AI models in secure, high-side cloud environments that have seamless access to relevant classified data could create quality outputs, which would then incentivise intelligence organizations to take a strategic, enterprise-wide approach to data management. With the "right" data representations and compute power, AI could transform knowledge systems within the ICs, transitioning from traditional, reactive methods to a more anticipatory, data-driven approach that moves beyond a reliance on historical data and instead focuses on predictions and real-time insights as the basis for strategic decision-making.

Action 2: Start Automating Portions of the Analytic Workflow Now

Participants argued that the ICs would find that several of their legacy organizational structures and behaviors, and analytic tradecraft, would likely be disrupted as AI technologies mature, systems become more capable, and consumer demand for AI-powered intelligence solutions rises. Participants recognized that cultural changes tend to happen slowly in intelligence organizations due to classified work environments and an aversion to risk, but nonetheless urged the ICs to prepare themselves for potentially significant cultural change brought on by AI. One key to successfully navigating AI integration would be to take a systemic approach to developing new AI-enabled analytical tradecraft, one that actively designs for interoperability and functionality.

As a starting point, participants suggested that Australian and U.S. analytic leaders deconstruct all of the key elements of their analytic workflows with an eye toward identifying areas that AI automation might be usefully applied to achieve greater efficiencies. Analytic managers should then select a few areas for initial focus and insist that units begin utilizing AI to realize early gains

[35] The 2023 National Intelligence Strategy notes that "a Community-wide, data-centric approach based on common standards is crucial to realizing the full promise of new capabilities." National Intelligence Strategy 2023, Office of the Director of National Intelligence at 9 (2023).

and to build their organizational acumen with the technology. An individual analyst's daily routine for "checking traffic" to keep up with events, for example, likely would also experience significant change. Workshop representatives from private sector firms advised IC leaders to look beyond likely impacts on analysts' personal workflows and be prepared to explore how AI might transform organizational behavior. Some noted that their experience taught them that workflow phases that currently have a heavy amount of human redundancy could be shortened or eliminated by using machines. If AI enables authors to perform sourcing and analytic self-checks, then presumably the need for multiple layers of management review of written work would be reduced as would the amount of time for delivery of insights to customers. Such advances may not fit well with existing production processes and structures and may require new ones. One participant from the private sector warned that the role of first- and second-line analytic supervisors will eventually need to be re-thought as a core component of their current job function – performing review – would become less necessary.

Action 3: Build Human-Machine Analytic Teams

Workshop participants from private sector firms that are already well along on the journey toward AI integration advised the ICs to create hybrid teams of human beings and machines to take full advantage of AI. They underscored the urgency of undertaking this task now, while AI is still in its infancy, so that the two ICs can get a head start on making the many procedural, structural, and cultural changes that will be necessary for successful adaptation. The concept of HMT revolves around forging a relationship – one made up of at least three equally important elements: the humans, the machines, and the interactions and interdependencies between them. Building trustworthy AI that is transparent, interpretable, reliable, and exhibits other characteristics and capabilities that enable trust is an essential part of creating effective human-machine teams. But so is having a good understanding of the human element in this relationship.[36]

Workshop discussions centered on how to foster greater trust in AI systems by human analysts. Achieving truly advanced HMTs with AIs acting alongside analysts in seamless "copilot" arrangements depends on building institutional knowledge, trust, and capabilities today, through incremental – but steady – progress toward more automation that prioritizes usability and transparent and equitable partnerships between human and machine. The ICs would do well to study human analysts' needs and mission priorities in future AI design and integration.[37] And analysts should be afforded the space and time to explore the capabilities of AI tools in a 'safe' testbed environment, such as on low-side unclassified networks. Participants argued that integrating AI tools into analysts' workflows in this way holds the best chance for increasing the

[36] Margarita Konaev & Husanjoy Chahal, Building Trust in Human-Machine Teams, Brookings Institute (2021).

[37] Brian Katz, The Analytic Edge: Leveraging Emerging Technologies to Transform Intelligence Analysis, Center for Strategic and International Studies (2020).

rate of AI adoption and ensuring AI tools meet mission requirements and tradecraft standards.[38] Concurrently, ICs should work with government authorities to obtain the appropriate permissions and resources to acquire and deploy recommended tools effectively.[39]

- In this, there is no "one size fits all" approach to AI adoption. Leaders should look to an agile, risk-based approach and graduated explainability standards that account for variability in user needs and use-case sensitivities, as well as the AI systems and tools applied.[40]

Retaining human control and oversight of AI-augmented teams was viewed by workshop members as essential to gaining initial acceptance with the analytic workforce.[41] AI systems will need to integrate seamlessly into existing analyst workflows, rather than introduce unfamiliar processes. The ideal system would operate almost invisibly in the background, proactively delivering relevant insights and drafting and editing products that appear as if written by a colleague. Such applications could generate enthusiasm and drive voluntary adoption more effectively than mandated tools that seem extraneous to an analyst's core tasks. In such cases, identifying and promoting a flagship AI project can act as a beacon, showcasing the potential and driving broader cultural acceptance and enthusiasm for AI integration.

- Early experimentation should focus on developing intuitive and user-friendly interfaces that enable analysts to interact with and contribute to the maturation of AI systems in natural and efficient ways. ICs should embrace a variety of avenues to develop human-AI collaboration, such as input-process-emergent state-output-input (IPEOI) models, input-moderator-output-input (IMOI) models, and joint cognitive systems.[42] These models provide a framework for understanding the roles and interactions between humans and AI in the analytic process, emphasizing the iterative nature of collaboration, and the moderating role of humans.

---

[38] Anna Knack, et al., Human-Machine Teaming in Intelligence Analysis: Requirements for Developing Trust in Machine Learning Systems, Alan Turing Institute: Centre for Emerging Technology and Security, (2022).

[39] The AIM Initiative: A Strategy for Augmenting Intelligence Using Machines, Office of the Director of National Intelligence (2024).

[40] Patricia McDermott, et al., Human-Machine Teaming Systems Engineering Guide, MITRE (2018).

[41] Anna Knack, et al., Human-Machine Teaming in Intelligence Analysis: Requirements for Developing Trust in Machine Learning Systems, Alan Turing Institute: Centre for Emerging Technology and Security, (2022).

[42] Breana M. Carter-Browne, et al., There is No "AI" in Teams: A Multidisciplinary Framework for AIs to Work in Human Teams, Applied Research Laboratory for Intelligence and Security (2021). Also referred to as input-process-output (IPO) models. See also Daniel R. Ilgen, et al., Teams in Organizations: From Input-Process-Output Models to IMOI Models, Annual Review of Psychology at 517-543 (2005); Steve W. J. Kozlowski & Daniel R. Ilgen, Enhancing the Effectiveness of Work Groups and Teams, Psychological Science in the Public Interest at 77–124 (2006); Laurent Karsenty & Patrick Brézillon, Cooperative Problem Solving and Explanation, Joint Cognitive Systems, Cooperative Systems and Decision Support Systems: A Cooperation in Context, Proceedings of the European Conference on Cognitive Science at 129-139 (1997).

Because HMTs are fundamentally different from traditional human-only teams, the two ICs will need to redesign and align agencies' legal, policy, and governance frameworks to account for them. Equally important will be the task of creating the new analytic tradecraft that HMTs will employ. Workshop contributors encouraged the ICs to take a flexible approach to developing these frameworks that can adapt as the technology evolves and the scope of what AI "teammates" are entrusted to handle grows. But work should begin right away so that within 3-5 years time, analysts, analytic managers and IC executives, and intelligence customers will be able to share a common understanding of how strong analytic HMTs perform and how accountability for results is to be apportioned between humans and machines.

Action 4: Create AI-Ready Training and Incentive Structures for the Analytic Workforce

To effectively integrate these systems will require a workforce that is prepared and adept at exploiting these tools to their fullest potential. To date, AI integration has tended heavily towards testing, evaluation, and experimentation. Workshop participants noted that, because they ultimately will be held accountable for assessments, analysts are naturally reluctant to trust AI outputs unless they are able to see all the data used to underpin judgments made by AI tools.[43] The two ICs' track record on rolling out new technologies to the analytic workforce (which too often rely on insufficient amounts of training as the sole crux of the adoption strategy) has created skepticism and "tool fatigue" across the workforces in both countries. Analytic practitioners warned that approaching AI deployments in a similar manner would be rejected by line analysts as overly-burdensome, thus placing an unnecessary brake on adoption. Equally, workshop discussions noted that an overreliance on training frequently avoided addressing deeper integration challenges, assuming that tools were inherently well-designed and that users lacked skills or willingness to leverage them effectively, when the crux of the issue often lies in mismatched tool design rather than human inadequacy.

The ICs will need to invest more in digital acumen, both through the recruitment of highly-trained talent and the upskilling of the existing workforce.[44] The intelligence professionals who participated in the workshops readily acknowledged that U.S. and Australian agencies currently lack sufficient staff with in-depth expertise on AI and that a growing number of private sector industries, such as finance and insurance, are much better at recruiting fresh AI talent. All participants agreed that trusted partnerships between intelligence agencies and other government agencies, industry, non-profit, and academic stakeholders held enormous potential

[43] See also Anna Knack, et al., Human-Machine Teaming in Intelligence Analysis: Requirements for Developing Trust in Machine Learning Systems, Alan Turing Institute: Centre for Emerging Technology and Security, (2022).

[44] Anna Knack, et al., Human-Machine Teaming in Intelligence Analysis: Requirements for Developing Trust in Machine Learning Systems, Alan Turing Institute: Centre for Emerging Technology and Security (2022).

for upskilling the IC workforce through a continuous process of development, deployment, and iteration. Workshop participants advocated for hands-on training with customized systems to allow unskilled users to experience streamlined productivity firsthand, which would spark both enthusiasm and capability. Over time, this would encourage analysts to take ownership of their HMT's performance, including contributions from AI.

All participants agreed that trusted partnerships between intelligence agencies and other government agencies, industry, non-profit, and academic stakeholders held enormous potential for upskilling the IC workforce through a continuous

Practitioners recognized the need to keep the talent pool engaged and focused through challenges and non-financial rewards. This meant seeking changes to traditional government incentive systems and mere "output production," ensuring that analysts saw a link between their analysis and meaningful decision-making. IC analytic leaders should incentivize and reward analysts and units that lean forward to climb the learning curve, and IC leadership must be prepared to accept a degree of risk and the occasional failure.[45]

Action 5: Collaborate to Develop a Shared U.S.-Australian Analytic AI Roadmap

Workshop participants noted a disparity between U.S. and Australian technical depth on AI, which could create a gap between the two intelligence communities' analytic capabilities in the years ahead. Building and maintaining vibrant and healthy alliances – undergirded by strong intelligence ties – is key to overcoming this gap and bolstering deterrence in a global strategic environment in which democracies are likely to face stronger competition from the world's leading autocracies – notably the People's Republic of China, Russia, Iran, and North Korea.

The two ICs should institute a regular bilateral dialogue on AI capabilities and deployments, a portion of which should be led by analytic leaders from both sides to focus on AI for analysis. A key first step would be for Australian and U.S. analytic leaders to clearly articulate a set of goals for how they intend to incorporate AI. These goals should be ambitious, take into account where AI technologies are likely to be in 3-5 years, and should include a roadmap of collaborative initiatives to keep the two analytic communities roughly on par with each other. In addition to establishing regular bilateral exchanges, Canberra and Washington should avail themselves of the recently launched AUKUS "Pillar Two" trilateral dialogue with the UK on advanced capabilities to further synchronize on AI deployments.

[45] Brian Katz, The Analytic Edge: Leveraging Emerging Technologies to Transform Intelligence Analysis, Centre for Strategic and International Studies (2020).

The bilateral U.S.-Australian AI integration roadmap should focus, in the first instance, on key definitions and activities that will help to implement current AI technologies into the analytical workflow and foster user acceptance, while laying the foundations for organizational change. The roadmap should:

- **Conduct a rapid initial assessment of agencies' "AI readiness" on data governance and technology management standards, policies and processes**, with a view to developing and aligning data governance frameworks and metadata management systems across the ICs. Where AI is currently deployed, capture how augmented workflows succeed and fail to refine best practices while building institutional know-how.
- **Define cross-IC prerequisites for AI adoption**, coupled with guidance on sustainability considerations and the extent to which AI model outputs can be generalized. This should include defined but graduated risk thresholds and explainability requirements that account for variability in user needs and analytical tasks, giving agencies and analysts the flexibility and confidence to select appropriate AI tools for different use cases.
- **Develop formal and transparent ethical and analytic tradecraft standards for HMTs**. This should focus on articulating ethical guardrails and thresholds, and explaining how they will relate to existing guidelines such as the United States' October 2023 Executive Order on the Responsible Use of AI and the forthcoming U.S. National Security Memorandum on AI for National Security.[46] Agencies should also update and augment existing analytic tradecraft standards, such as the U.S. Intelligence Community Directive 203 to articulate how HMT will be expected to perform, including the extent to which human analysts will be held accountable for AI actions.[47] Participants recognized that standards probably would vary, but urged pressing ahead with collaboration so each side can sharpen its thinking on AI use and develop mutual understanding of how each IC plans to proceed.
- **Commit to fielding analytic HMTs and sharing lessons learned.** One idea for fostering cross-collaboration would be to launch a pilot project to employ an AI-augmented HMT to tackle a discrete intelligence analysis challenge of common concern. The ICs' analytic leadership should establish a regular dialogue to evaluate their respective HMTs and share lessons learned. Goals for this exchange should be to:

[46] Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence, The White House (2023); Alexandra Kelley, Biden to Receive AI National Security Memo Outlining Forbidden Uses, Opportunities for Innovation, NextGov/FCW (2024).

[47] ICD 203, Analytic Standards, Office of the Director of National Intelligence (2022). See also AI Ethics Framework for the Intelligence Community, Office of the Director of National Intelligence (2020).

- Systematically map analytic workflows to identify areas for automation and to identify best-of-breed AI tools and systems that enable human analysts to interact with AI systems in natural and efficient ways.
- Exchange lessons learned from AI training programs to identify which approaches work best for building analysts' understanding and trust in the technology.
- Exchange accountability mechanisms and procedures for rigorously stress testing AI systems.

Workshop participants agreed that the ICs will not be able to take full advantage of AI technologies without committed partners from industry and academia that are highly attuned to the intelligence operating environment, the ICs' missions, strategic directions, and operational objectives.

U.S. and Australian IC leaders should foster persistent joint exploration and co-creation of AI capabilities between their respective private sectors where possible. As part of the roadmap, Australian and U.S. experts should establish a private sector partner forum to regularly update IC leaders on the latest AI innovations, provide technical advice, exchange findings from AI testing and evaluation programs, and share best practices in the management of human-machine teams. Should the two ICs identify critical gaps between U.S. and Australian AI capabilities, the private sector forum could propose discrete solutions or serve as a conduit for directing technical expertise to whichever IC needs assistance.

Appendix A: Factors That Influence Generative AI Model Performance

LLMs rely on several interrelated components which determine how the model creates an output and the content and quality of that output. Tweaking any of these components even slightly can have significant impacts on a given model's performance. The four major factors that influence the performance of today's LLM-powered generative AIs are:

- **The model's architecture.** The architecture refers to how an LLM is designed so that it can interpret and respond to user queries, and it determines how an LLM processes inputs, learns during training, and generates outputs. While model architectures tend to be the domain of engineers and developers, the architecture also contains 'attention mechanisms' that determine how the LLM has been instructed to weigh (i.e., assign a degree of importance to) information. The weights assigned to information are likely to be of significant importance to intelligence analysts and will require close coordination between technologists and practitioners.

- **The data.** The breadth, depth, and quality of the information contained in the training of an LLM determine its ability to generate quality responses to user queries. Unlike the LLM that powers ChatGPT (or any of the other major publicly available LLMs geared toward serving generalists), LLMs used in support of the analytic mission will also need to contain information from sources that human analysts use, such as: speeches, transcripts, and news articles from intergovernmental, governmental, and nongovernmental organizations; news outlets; and social media, as well as analytic-specific rules sets such as style guides that account for acronyms and intel-specific terms and jargon. How the data is structured is likely to influence how sophisticated a question an LLM can respond to, as well as the sophistication of its answers.[48]

- **The analytic methodologies and algorithms used to train an LLM.** The design of the analytic methodologies used to train the algorithms in an LLM is key to how effectively the model will learn to detect meaningful patterns in the data. An LLM that has been

[48] Lauren C. Williams, The CIA's Data-Challenge AI Imperative, Defense One (2023).

optimized for generalists, for example, might fail to meet the information needs of subject matter experts because generalists might be satisfied with an answer that seems superficial to the expert. At the same time, not using off-the-shelf frontier LLMs will be cost prohibitive both from a time and financial perspective.

- **Prompts.** ICs will need to fine tune prompting using established analytic standards and methodologies currently used to train intelligence analysts throughout the analytic process and repositories of finished analytic products. Users employ prompts to interact with the LLM. Unlike Boolean queries, which follow a strict set of rules defining how a user queries information, prompts allow the user to interact with the LLM using natural language. Prompts can be used to aid a variety of text-based workflows, including low- and high precision tasks (e.g., scenario generation and question answering, respectively) as well as framing the perspective or context for a generated response (e.g., how would you interpret this event if you were a named politician). The LLM's ability to respond to any given prompt will be dependent on its data, architecture, and training.

Appendix B: Differing Perspectives on AI's Potential

The variety and velocity of work being done on artificial intelligence, and the differing analytic disciplines and customer/consumer perspectives, often complicate conversations about how AI might be applied in analytic settings. During our workshops, the conversation was wide-ranging, touching on everything from potential applications of AI for tactical and operational military planning through to how AI might be used in support of national policymakers. A critical subtext of these conversions was whether the speaker was a technologist or a practitioner. Both positions were equally valid and, taken together, produced a landscape of opportunities and challenges; however, both groups highlighted the transformative potential of GenAI through the benefits of increased efficiency and data consolidation.

- *Technologists* is our shorthand for individuals who may have had direct or indirect experiences with national security institutions in a technical or technological capacity and were focused on the technological potential of AI to transform intelligence analysis. In some cases, they were former government executives, and in other cases they represented vendors who support and service analytic missions. As a cohort, they were more optimistic and enthusiastic about AI and its ability to support an array of tasks across analytic missions, including serving as the driving force behind unnamed transformative analytic processes and products.

- *Practitioners* is our shorthand for individuals who have worked as analysts or alongside (or in support of) analysts. While they often were not as familiar with projected and likely AI developments, they were aware that the rise of GenAI had implications for analysts as well as the professional, cultural, and bureaucratic dynamics likely to influence AI's development and implementation in analytic work environments. In most cases, practitioners adopted a more conservative approach that emphasized the need for guardrails and gradual adoption.

Regardless of whether a participant was a technologist or practitioner, all thought that AI is poised to alter analytic workflows in ways that would allow human analysts to spend more time on value-added activities, such as those that required expert intuition. The best approach to exploiting ML and AI in intelligence organizations was the close coupling of practitioners and technologists: it was clear that intelligence services have experience and insight that can push ML and AI in compelling directions and we hope to continue to glimpse their use cases and successes as they experiment with and implement the current generation of advanced ML and AI technologies.

	Technologists' Expectations	Practitioners' Expectations
Performing Search & Discovery	AI will revolutionize how intelligence agencies extract meaning from the global datasphere.	AI could help analysts prioritize actionable intelligence from large datasets and reduce the scanning workload.
Producing Analytic Insights	AI will allow intelligence services to generate more valuable assessments, across a broader set of targets, at greater speed and including a wider array of sources.	AI could help analysts generate ideas, test their hypotheses, and perform simple, low-risk analytic functions.
Engaging Intelligence Customers	AI could transform how customers interact with intelligence agencies, allowing for more "self-service" and embedding customer requirements into the intelligence cycle.	AI may be useful for tracking rapidly-evolving customer requirements and alerting analysts when new assessments break with pre-existing analytic lines, requiring notification to customers.
Work Processes	AI will significantly revolutionize the analytic process, adding speed and more data-driven analysis. Human-centered quality and security assurance activities will be automated, spurring a cultural shift in how analytic enterprises are staffed and led.	If analysts' concerns are taken into account, AI could reduce administrative tasks and shorten timelines for editing and review.

Contributors

The authors are grateful for the energetic support they received from the leadership and staff of the Australian Strategic Policy Institute and the Special Competitive Studies Project and from the Australian and U.S. Intelligence Communities. We wish to acknowledge the assistance of intelligence professionals from Australia's Office of National Intelligence, Department of Home Affairs, Australian Transaction Reports Analysis Centre, and from several U.S. agencies, including the Office of the Director for National Intelligence, Central Intelligence Agency, National Geospatial Intelligence Agency, and Intelligence Advanced Research Projects Activity, all of whom gave generously of their time and expertise. The below list of non-government experts contributed vital perspectives and technical knowledge. We are especially indebted to Mr. Dennis Gleeson, who drafted portions of this report.

Gavin Artz
Commonwealth Scientific and Industrial Research Organisation

Dr. Debi Ashenden
University of New South Wales Institute for Cyber Security

Dr. Zena Assad
Australian National University

Aileen Black
Groq

Paul Brown
Groq

Destinie Carbone
Omega Minds

Robert Cardillo
Former Director, National Geospatial-Intelligence Agency

Melissa Carraway
University of Maryland Applied Research Laboratory for Intelligence and Security

James Cooper
Agent Oriented Software

James de Lorimier
Center for Advanced Defense Studies

Dr. Richard de Rozario
University of Melbourne

Marina Favaro
Anthropic

Cameron Fox
Athena

Dennis Gleeson
Former Central Intelligence Agency

Bob Gourley
OODA LLC

Brett Greenshields
Australian Department of Home Affairs

Michael Groen
Former Director, Joint Artificial Intelligence Center, U.S. Department of Defense

Ada Guan
Rich Data Co.

Heather Horniack
Scale AI

Kuba Kabacinski
Athena

Geoff Kahn
Palantir

Greg Levesque
Strider Technologies

Dr. Andrew Lucas
Agent Oriented Software

Dr. Simon Lucey
Australian Institute for Machine Learning

Dr. Andrew S. Mara
Johns Hopkins University Applied Physics Laboratory

James McCool
Former National Geospatial Intelligence Agency

Kathy Nicholson
Australian Institute for Machine Learning

G. Andrew Otterbacher
Scale AI

Dr. Gareth Parker
Commonwealth Scientific and Industrial Research Organisation

Dr. Barton L. Paulhamus
Johns Hopkins University Applied Physics Laboratory

Dr. Jane Pinelis
Johns Hopkins University Applied Physics Laboratory

Chris Poulter
OSINT Combine

Krista Rasmussen
Center for Advanced Defense Studies

Jonathan Ross
Groq

Dr. Philip Sage
Johns Hopkins University Applied Physics Laboratory

Rohan Samaraweera
Australian Department of Home Affairs

Glen Schafer
Trusted Autonomous Systems

Michael Sellitto
Anthropic

Dr. Sarah Shoker
OpenAI

Dr. Jason Signolet
Fivecast

Chitra Sivanandam
Living Dino

Dean Souleles
Former Director of the AI Augmented Mission Office, ODNI

Margaret A. Stromecki
MITRE

Mike Wyatt
OpenAI

David Tynan
Commonwealth Scientific and Industrial Research Organisation

Kristin Wood
August Interactive

SPECIAL COMPETITIVE STUDIES PROJECT

ASPI AUSTRALIAN STRATEGIC POLICY INSTITUTE