



DKI APCSS
SECURITY
NEXUS

The DKI APCSS journal for comprehensive security issues throughout the Indo-Pacific region and the world!

Security Nexus Perspective

START HUMAN, END HUMAN: A PRACTICAL FRAMEWORK FOR LARGE LANGUAGE MODELS IN INDO-PACIFIC SECURITY COOPERATION

By Dr. Elizabeth Vaughan Moyer
DKI APCSS Fellow, Major, United States Air Force

Abstract

Security professionals across the Indo-Pacific face an urgent challenge: leveraging large language models (LLMs) for enhanced decision-making while navigating United States-China technology competition, language barriers, and sovereignty concerns. Research demonstrates technologies follow predictable adoption patterns through hype and disillusionment before reaching productivity, yet hesitation during these transitions carries mounting opportunity costs. This perspective proposes a practical framework: start with human judgment, end with human decision, and use LLMs as cognitive tools in between. Evidence suggests Artificial Intelligence can improve skilled worker performance when properly applied within its capabilities, while organizations delaying adoption risk strategic disadvantage. This paper offers practical guidance for processing overwhelming information flows without sacrificing critical thinking or accountability. The framework focuses on immediately actionable approaches for using commercially available LLMs to enhance human analysis in resource-constrained environments. The author demonstrates that security organizations embracing experimentation despite imperfect technology will achieve significant advantages over those waiting for ideal solutions.

Keywords: technology adoption, human-AI partnership, critical thinking

Introduction: The Cost of Waiting

Often, transformative technology follows the same psychological journey. The Gartner Hype Cycle, first introduced by analyst Jackie Fenn in 1995, describes five distinct phases through which technologies progress: technology trigger, peak of inflated expectations, trough of disillusionment, slope of enlightenment, and plateau of productivity (Dedehayir & Steinert, 2016). Those who embrace the new tech, security professionals included, experience this as initial fear, followed by unrealistic enthusiasm, then disappointment, before finally achieving productive integration.

The "trough of disillusionment" represents a critical phase where "interest wanes as experiments and implementations fail to deliver. Producers of the technology shake out or fail. Investments continue only if the surviving providers improve their products to the satisfaction of early adopters" (Gartner, 2023). Yet this phase also presents opportunities for those willing to persist, especially in the era of Artificial Intelligence, or AI. As one technology adoption analysis notes, the trough "can lead to decreased funding, talent acquisition, and momentum, but also presents an opportunity for introspection, refinement, and course correction" (Quantum Zeitgeist, 2025). Today's Indo-Pacific security organizations stand at this critical juncture with large language models or LLMs. Some remain paralyzed in the fear stage, worried about hallucinations, bias, "workslop" and sovereignty (Nemeroff, 2025). Others have crashed into disappointment after expecting AI to solve all problems. Meanwhile, adversaries who push through these stages gain compound advantages.

The stakes are clear: research estimates that generative AI could add the equivalent of \$2.6 trillion to \$4.4 trillion annually across analyzed use cases (McKinsey & Company, 2023). In security terms, this translates to faster intelligence processing, increased pattern recognition, and accelerated decision-making. Organizations that hesitate do not just miss opportunities—they cede advantage to competitors willing to navigate imperfection.

Understanding the Adoption Journey

Why Organizations Hesitate

The technology adoption lifecycle identifies five adopter categories based on psychological characteristics, as originally defined by Rogers (1962) and further elaborated in subsequent research by Carr and others (Rimal & Storey, 2013):

- **Innovators (2.5%):** Risk-oriented individuals with financial resources to absorb failures, tend to be youngest in age, have a high social class, great financial lucidity, and social access to scientific sources
- **Early Adopters (13.5%):** Opinion leaders who are younger in age, have a higher social status, have more financial lucidity, advanced education, and are more socially forward than late adopters

- Early Majority (34%): More conservative but open to new ideas, tend to be more hesitant in the adoption process, have above-average social status, and usually have contact with early adopters
- Late Majority (34%): Skeptical about an innovation, below average social status, very little financial lucidity, often in contact with others in the late majority and early majority
- Laggards (16%): Typically have an aversion to change-agents and tend to be advanced in age, usually focused on traditions, likely to have the lowest social status as well as financial fluidity

In Indo-Pacific security establishments, these manifest as hyper-specialized organizations experimenting with AI, forward-thinking leaders recognizing advantages, large groups waiting for proof, skeptics requiring peer validation, and traditionalists resisting change. A common pitfall is that organizations often mistake prudence for wisdom. Often justified as assessing or mitigating risk, they wait for perfect technology while adversaries gain experience with imperfect tools. As research on the hype cycle model notes, many organizations fail to navigate through the disillusionment phase effectively, missing opportunities to refine and develop capabilities while competitors advance (Dedehayir & Steinert, 2016).

The Opportunity Cost of Delay

Consider two neighboring nations: Nation A begins LLM experimentation in 2024, accepting imperfection while building institutional knowledge. Nation B waits for "mature" technology in 2027. By the time Nation B begins, Nation A has:

- 3 years of prompt engineering expertise
- Established verification protocols and identified potential bias issues
- Identified optimal use cases and areas where LLMs fail
- Trained hundreds of operators and increased AI literacy
- Developed institutional memory and lessons learned

The gap compounds. Research on highly skilled workers found that "GPT-only participants saw a 38% increase in performance compared with the control condition (no access to AI)" and those with additional training showed "42.5% increase in performance" (MIT Sloan, 2023). Nation A's analysts process information faster, identify patterns humans miss, and free human cognition for strategic thinking, while Nation B still manually processes reports.

The Framework: Start Human, End Human

Scope

This paper deliberately limits its scope and does not address a plethora of topics such as autonomous weapons systems or lethal decision-making, disinformation generation or detection, cyber warfare applications, population surveillance systems, strategic AI competition dynamics, or the potential adoption of less expensive small language models. These deserve separate treatment and could be an entire series. This proposed framework focuses solely on using LLMs as cognitive tools for processing information that humans already collect but cannot efficiently analyze.

Theoretical Foundation

Adopting a bookend principle maintains human cognition at decision initiation and conclusion, with LLM processing in between. This is not philosophical—it's a practical workflow maintaining accountability while leveraging computational power. Studies show dramatic productivity gains when AI is properly applied. "One study found that software developers using Microsoft's GitHub Copilot completed tasks 56 percent faster than those not using the tool" (McKinsey & Company, 2023). However, researchers found performance varies significantly: when AI is used within its capabilities, worker performance could improve up to 40%, but "when AI is used outside that boundary to complete a task, worker performance drops by an average of 19 percentage points" (MIT Sloan, 2023). Thus, security professionals must continue to experiment and determine when and where LLMs make sense to utilize.

Starting Human: Problem Definition

Picture this: a Malaysian maritime officer monitoring the Strait of Malacca does not ask an LLM, "What should I do about suspicious vessels?" They start with judgment: "From 47 sensor-flagged contacts, which merit investigation? Prioritize deviations from standard routes, signs of AIS manipulation, correlations with smuggling indicators, and proximity to critical infrastructure." The human supplies context—local patterns, political sensitivities, current intelligence priorities, even intuition—that tools lack by default or without significant training. The model then augments, rather than replaces, expert decision-making.

The Middle: LLM Processing

Once humans frame problems and provide extensive context, LLMs can excel at specific tasks. As prompt engineering guidance emphasizes, humans need to give at a minimum information like: Instruction, describing a specific task you want a model to perform; Context, additional information or context that can guide's a model's response; Input Data, expressed as input or question for a model to respond to and Output Format, the type or format of the output (Prompt Engineering Guide, 2024). Of note, there are several variations of prompt engineering frameworks that are continually evolving, including asking the LLM what data it needs to develop its own prompt and then providing a proposed answer. Returning to the maritime example, the LLM may respond and: 1. Process multiple vessel reports to identify risk indicators, 2. Compare movements to documented patterns 3. Standardize reports from multiple languages and 4. Convert various formats into templates to assist the practitioner.

Ending Human: Decision and Accountability

Most importantly, LLMs output organized information, not decisions. They process best according to human criteria (noting that it can be flawed), not making autonomous judgments. The Malaysian maritime officer must review processed data, apply local knowledge, consider political sensitivities, and determine which vessels to investigate. Accountability remains clear because humans defined parameters, validated logic, applied judgment, and made decisions. Far too often, professionals become lazy, either consciously or unconsciously, at the end of the AI cycle. It is quintessential for humans to be the final check, scrutinizing and utilizing critical thinking.

Implementation Challenges and Solutions

As with any technology or tool, security professionals must understand current limitations regarding LLMs.

The Language Problem

Most powerful LLMs show degraded performance in regional languages (Stanford Human-Centered AI Institute, 2024). For example, a report in Tetum or Dhivehi will not process as accurately as English or Mandarin. Furthermore, the LLMs are not attuned to relevant cultural contexts. Therefore, security professionals need to use LLMs for initial translation but mark confidence levels. Consider: High-confidence translations proceed; low-confidence sections require human review. During training, learners must ensure they never make critical decisions based solely on LLM translation of limited-data languages.

The Bias Problem

Every LLM carries training data biases (Gallegos et. al, 2024). Typically trained on an enormous scale of uncurated Internet-based data, LLMs inherit stereotypes, misrepresentations, derogatory and exclusionary language, and other denigrating behaviors that disproportionately affect already-vulnerable and underrepresented populations. Additionally, U.S.-developed models interpret events differently than Chinese models and both will most likely misunderstand Pacific Islander contexts. Hence, when possible, professionals must strive to run analyses through multiple LLMs from different origins. Where divergence appears, human judgment becomes even more important and critical.

The Hallucination Problem

LLMs, via a generative AI chatbot or computer vision tool, perceive patterns or objects that are nonexistent or imperceptible to human observers, producing outputs that are nonsensical or altogether inaccurate (IBM, 2025). These hallucinations occur because LLMs generate responses based on pattern prediction rather than understanding. They "generate plausible content, not to verify its truth," and may produce content that sounds reasonable but is entirely inaccurate (MIT Sloan Teaching & Learning Technologies, 2025). Adding to more concerns, security professionals face the broader challenge of 'AI slop'—low- to mid-quality content created with AI tools, often with little

regard for accuracy. As Nemeroff (2025) explained, this content is “fast, easy, and inexpensive to make” but displaces higher-quality material that could be more helpful. For example, in intelligence analysis, the risk extends beyond individual errors to systemic degradation of information quality, particularly when AI-generated content enters training datasets for future models. An important best practice is to rarely, if ever, use LLMs for factual lookup. Instead, use them for pattern analysis of facts you provide. Common guidance suggests requiring the LLM to verify citations, query information multiple ways to identify inconsistencies, and treat AI outputs as hypotheses requiring confirmation.

The Sovereignty and Security Problem

Cloud-based LLMs mean data processed on foreign servers. For classified information, sensitive or controlled unclassified information, this is unacceptable. Beyond simple security concerns, this represents a fundamental sovereignty challenge. When a defense analyst in Bangkok queries GPT-5, their data travels to U.S. servers; when Tonga uses Claude, prompts route through Anthropic's infrastructure, and aggregated patterns can reveal intelligence priorities, operational focus, and strategic vulnerabilities. The problem compounds under legal frameworks like the U.S. CLOUD Act or China's National Intelligence Law, which may compel data disclosure from providers operating in their jurisdictions (Bradford, Bunn & Gharagozlou, 2024). Indo-Pacific nations must navigate this through various approaches—Singapore is developing local AI capabilities while maintaining selective openness (Ministry of Communications and Information, 2023), India is pushing data localization requirements (Rajmohan, 2025), and smaller nations face the stark choice between accepting sovereignty risks or forgoing capabilities entirely. Some nations are exploring regional partnerships for shared secure infrastructure, but this introduces complexity around access controls and operational security. While technical solutions like federated learning or homomorphic encryption exist, most nations must develop pragmatic classification tiers that balance capability with sovereignty, understanding that in an era where information processing partially determines security outcomes, digital dependency represents a new form of strategic vulnerability.

Real-World Potential Application

Maritime Domain Awareness (Philippines)

As one potential example, the Philippines faces substantial maritime governance capacity gaps, with shortfalls in maritime domain awareness and additional challenges like limited surveillance capabilities, inadequate assets including patrol vessels and aircraft, and underdeveloped information-sharing mechanisms (Asia Maritime Transparency Initiative, 2023). Philippine maritime analysts confront overwhelming operational realities: more than 200 vessels transit their waters daily, generating reports in multiple languages, while limited analyst teams manually process this information (Rabasa & Chalk, 2012). The "start human, end human" framework could transform this bottleneck. Human analysts define priorities based on current intelligence requirements, focusing on vessels near contested zones or those exhibiting anomalous behavior. LLMs then process incoming reports, flag patterns such as AIS manipulation or suspicious vessel clustering, standardize reports

from regional stations, and provide initial translations with confidence scores. Human analysts apply irreplaceable local knowledge, understanding seasonal fishing patterns, recognizing legitimate versus suspicious activities, considering other nuanced sensitivities, and finally deciding which vessels warrant investigation. Processing time is reduced, enabling the same small team to effectively monitor a domain that would otherwise require dozens of analysts.

Critical Considerations

The China Factor

The U.S.-China technology competition creates uncomfortable realities for security professionals. Major LLMs come from competing powers—OpenAI's GPT and Anthropic's Claude from the U.S., Baidu's ERNIE and Alibaba's models from China. For Indo-Pacific security organizations, this creates operational dilemmas: using U.S.-developed LLMs may embed Western analytical frameworks into intelligence assessments, while Chinese systems carry their own geopolitical perspectives shaped by different information environments. The technology competition extends beyond bias concerns into strategic dependency—reliance on either nation's AI infrastructure creates potential vulnerabilities if access is restricted during crises or if adversaries exploit known system limitations. As U.S.-China strategic competition intensifies across the technological domain, with emerging technologies playing a decisive role in the contest between the two superpowers over the future of global power (Stokes, Sullivan & Greene, 2023), security organizations must navigate these dependencies carefully. Practitioners must exhibit pragmatic diversification using unique systems for different purposes while building indigenous capabilities. This means employing multiple LLMs for comparative analysis on sensitive assessments, understanding each system's embedded biases, and investing in domestic AI expertise rather than accepting permanent dependence on external powers—recognizing that in an era of strategic competition, technological sovereignty matters as much as territorial sovereignty.

Building Prompt Engineering Skills

Prompt engineering is a relatively new discipline for developing and optimizing prompts to efficiently use language models and requires practice, experimentation, and most importantly, patience. Research indicates effective prompting requires structured communication, not coding (Saravia, 2022). The CO-STAR framework provides guidance: Context (background), Objective (task), Style (writing style), Tone (attitude), Audience (readers), Response (format) (The Modern Scientist, 2024). Interestingly, some of the most proficient professionals to interact with large language models end up being people who work with a lot of other people, like educators or human resources, because they understand communication nuances.

Managing the Motivation Paradox

Recent research reveals a critical challenge: "While gen AI collaboration boosts immediate task performance, it can undermine workers' intrinsic motivation and increase feelings of boredom when they turn to tasks in which they do not have this technological assistance" (Liu et al., 2025). This dependency is dangerous in security operations where system degradation is expected during crises.

Teams must maintain manual capabilities through regular non-AI exercises. For intelligence analysts, this means the cognitive muscle memory for pattern recognition, analytical reasoning, and threat assessment without computational assistance can atrophy if not deliberately exercised—precisely when adversaries may target AI infrastructure or when network failures eliminate access to cloud-based LLMs during critical operations. Security organizations should consider scheduling routine "analog days" where analysts conduct assessments using only traditional methods, ensuring that when LLM systems fail during a cyber attack or when operating in denied environments, teams retain the fundamental analytical capabilities that technology was meant to augment, not replace.

From Experimentation to Integration

Successful LLM adoption requires deliberate sequencing across three levels. First, individual analysts should experiment with unclassified open-source analysis, documenting which prompt strategies work and which fail to build institutional knowledge. Second, organizations must establish sandboxes for risk-free team experimentation, identifying patterns that justify formal protocols and training (Mühlroth & Grottke, 2025). Organizations need to invest in verification systems *before* expanding beyond sandboxes—deploying unvalidated capabilities during crises introduces unacceptable risk. Successful AI adoption requires "embracing experimentation while setting strategy," capturing "lessons and data from these experiments—both successes and failures—to inform medium-to-longer-term AI strategies" (McDonagh-Smith, 2024). Third, regional cooperation like the ASEAN Defence Ministers' Meeting-Plus and ASEAN Regional Forum enables smaller nations to share techniques and address language gaps collectively, building on these platforms' track record in maritime security and counterterrorism (Tan, 2019). These levels are interdependent: individual competence scales to organizational capability only with proper infrastructure, while regional cooperation fails without domestic expertise to absorb shared knowledge. For resource-constrained Indo-Pacific security organizations, this phased approach—experiment individually, institutionalize organizationally, collaborate regionally—provides a realistic path forward while maintaining the non-AI capabilities essential when systems fail during crises.

Conclusion: Embracing Imperfect Tools

The Indo-Pacific's security challenges will not wait for perfect technology. Every day of hesitation widens the gaps between organizations embracing experimentation and those awaiting ideal solutions. The bookend principle—start human, end human—is far from perfect, but it's implementable today.

Technologies predictably traverse cycles of hype and disillusionment before reaching productivity. As research notes, successful navigation requires recognizing that interest wanes as experiments and implementations fail to deliver, but that more instances of how technology can benefit the enterprise start to crystallize for those who persist (Gartner, 2023). For overworked security professionals from Manila to Male, the question is not whether to use LLMs but how to use them effectively while maintaining critical thinking. In a region where resources are limited but challenges are vast, even imperfect enhancement might determine whether threats are identified in time.

The cost of waiting for perfect AI is accepting imperfect security today. That's a price the Indo-Pacific cannot afford.

References

Anthropic. (2024, November 8). Anthropic awarded \$200M DOD agreement for AI capabilities. <https://www.anthropic.com/news/anthropic-and-the-department-of-defense-to-advance-responsible-ai-in-defense-operations>

Asia Maritime Transparency Initiative. (2023, December 18). Assessing the Philippines' maritime governance capacity: Priorities and challenges. Center for Strategic and International Studies. <https://amti.csis.org/assessing-the-philippines-maritime-governance-capacity-priorities-and-challenges/>

Boeing. (2025, September 17). Boeing Defense, Space & Security partners with Palantir to accelerate AI adoption across defense, classified programs [Press release]. <https://investors.boeing.com/investors/news/press-release-details/2025/Boeing-Defense-Space--Security-Partners-with-Palantir-to-Accelerate-AI-Adoption-Across-Defense-Classified-Programs/default.aspx>

Bradford, A., Bunn, M., & Gharagozlou, A. (2024). The CLOUD Act and transatlantic trust. *Center for Strategic and International Studies (CSIS)*. <https://www.csis.org/analysis/cloud-act-and-transatlantic-trust>

Damji, J. S. (2024, March 14). Best prompt techniques for best LLM responses. *The Modern Scientist*. <https://medium.com/the-modern-scientist/best-prompt-techniques-for-best-llm-responses-24d2ff4f6bca>

Dedehayir, O., & Steinert, M. (2016). The hype cycle model: A review and future directions. *Technological Forecasting and Social Change*, 108, 28-41. <https://www.sciencedirect.com/science/article/abs/pii/S0040162516300270>

Doshi, A., & Hauser, O. (2024, July 12). Generative AI enhances individual creativity but reduces the collective diversity of novel content. *Science Advances*, 10(28). <https://www.science.org/doi/10.1126/sciadv.adn5290>

Gartner. (2023). Gartner hype cycle research methodology. <https://www.gartner.com/en/research/methodologies/gartner-hype-cycle>

IBM. (2025). What are AI hallucinations? <https://www.ibm.com/think/topics/ai-hallucinations>

Kellogg, K., et al. (2023, October 19). How generative AI can boost highly skilled workers' productivity. *MIT Sloan Management Review*. <https://mitsloan.mit.edu/ideas-made-to-matter/how-generative-ai-can-boost-highly-skilled-workers-productivity>

- Liu, Y., Wu, S., Chen, S., & Xie, X. Y. (2025, May 13). Research: Gen AI makes people more productive—and less motivated. *Harvard Business Review*. <https://hbr.org/2025/05/research-gen-ai-makes-people-more-productive-and-less-motivated>
- McDonagh-Smith, P. (2024, April 3). Leading the AI-driven organization. *MIT Sloan Management Review*. <https://mitsloan.mit.edu/ideas-made-to-matter/leading-ai-driven-organization>
- McKinsey & Company. (2023, June 14). The economic potential of generative AI: The next productivity frontier. <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier>
- Ministry of Communications and Information. (2023). *National AI Strategy 2.0: AI for the public good, for Singapore and the world*. Smart Nation Singapore. <https://www.smartnation.gov.sg/files/publications/national-ai-strategy.pdf>
- MIT Sloan. (2023, October 19). *How generative AI can boost highly skilled workers' productivity*. <https://mitsloan.mit.edu/ideas-made-to-matter/how-generative-ai-can-boost-highly-skilled-workers-productivity>
- MIT Sloan Teaching & Learning Technologies. (2025, June 30). *When AI gets it wrong: Addressing AI hallucinations and bias*. <https://mitsloanedtech.mit.edu/ai/basics/addressing-ai-hallucinations-and-bias/>
- Mühlroth, C., & Grottko, M. (2025, May 22). AI initiatives don't fail—Organizations do: Why companies need AI experimentation sandboxes and pathways. *California Management Review*. <https://cmr.berkeley.edu/2025/05/ai-initiatives-don-t-fail-organizations-do-why-companies-need-ai-experimentation-sandboxes-and-pathways/>
- Nemeroff, A. (2025, September 2). What is AI slop? A technologist explains this new and largely unwelcome form of online content. *The Conversation*. <https://theconversation.com/what-is-ai-slop-a-technologist-explains-this-new-and-largely-unwelcome-form-of-online-content-256554>
- Prompt Engineering Guide. (2024). Prompt engineering guide. <https://www.promptingguide.ai/>
- Quantum Zeitgeist. (2025, January 14). What is the trough of disillusionment phase? <https://quantumzeitgeist.com/what-is-the-trough-of-disillusionment-phase/>
- Rabasa, A., & Chalk, P. (2012). *Non-traditional threats and maritime domain awareness in the tri-border area of Southeast Asia: The Coast Watch System of the Philippines* (OP-372-OSD). RAND Corporation. https://www.rand.org/pubs/occasional_papers/OP372.html
- Rajmohan, K. (2025, January 23). Data localization: India's tryst with data sovereignty. *TechPolicy.Press*. <https://www.techpolicy.press/data-localization-indias-tryst-with-data-sovereignty/>
- Rimal, R. N., & Storey, J. D. (2013). Exploring audience segmentation: Investigating adopter categories to diffuse an innovation to prevent famine in rural Mozambique. *Journal of Health Communication*, 18(4), 369-383. <https://pmc.ncbi.nlm.nih.gov/articles/PMC3772073/>
- Rogers, E. M. (1962). *Diffusion of innovations*. Free Press.

Saravia, E. (2022, December). *Prompt engineering guide*. <https://github.com/dair-ai/Prompt-Engineering-Guide>

Stanford Human-Centered AI Institute. (2024). Mind the (language) gap: Mapping the challenges of LLM development in low-resource language contexts. <https://hai.stanford.edu/policy/mind-the-language-gap-mapping-the-challenges-of-llm-development-in-low-resource-language-contexts>

Stokes, J., Sullivan, A., & Greene, N. (2023). *U.S.-China competition and military AI: How Washington can manage strategic risks amid rivalry with Beijing*. Center for a New American Security. <https://www.cnas.org/publications/reports/u-s-china-competition-and-military-ai>

Tan, S. S. (2019). Is ASEAN finally getting multilateralism right? From ARF to ADMM+. *Asian Studies Review*, 44(1), 43-61. <https://doi.org/10.1080/10357823.2019.1691502>



The views expressed in this article are those of the author and do not reflect the official policy or position of DKI APCSS, the Department of Defense, or the U.S. Government. The appearance of external hyperlinks does not constitute endorsement by the United States Department of Defense (DoD) of the linked websites, or the information, products, or services contained therein. DoD does not exercise any editorial, security, or other control over the information you may find at these sites.

December 2025