

MP-IDSA

Issue Brief

Generative AI and Military Applications: Is Civil-Military Fusion the Path of Choice?

Nestor Derrick D'Souza

November 12, 2025

Summary

Militaries seeking an early-mover advantage in achieving technological supremacy over adversaries have begun experimenting with Large Language Models/Generative AI to solve military problems. However, the resources required to develop an incipient technology like GAI into a battle-ready product capable of meeting complex military requirements seem prohibitive.

Introduction

ChatGPT has captured the imagination of millions since OpenAI made it available to the public on 30 November 2022.¹ Its attractiveness stems from its simple use. Any person can type a natural-language query and receive a perfectly coherent, comprehensive answer, with further prompts or suggested questions to refine the query. Generative AI, or GAI, can quickly generate coherent answers to questions in natural human language.² It can mean, predict, decide, strategise and execute many vital actions based on the size of its database and the sophistication of its neural networks, which generate patterns, recall information, or correlate the two to predict future occurrences.

GAI has opened a path into the realm now referred to as Cognitive AI. The excitement around GAI rises from its potential to revolutionise numerous fields by performing human-like cognitive functions. Cognitive AI relies on the perceived ability of machines to make decisions based on thinking and, in the near future, to feel like human beings. In the process, machines will be able to make decisions like human beings.³

GAI has emerged as an avenue for exploring the ability of machine intelligence to make decisions in a ‘biologically plausible manner’.⁴ Importantly, machines are likely to do so much faster and with the capability to absorb and process multiple inputs from multiple sources in different forms. In a future conflict, the advantage provided by a machine decision-maker that can arrive at decisions after visualising simultaneous operations across land, sea, air, space, electronic and cognitive domains at a fraction of the time required by a human decision-maker is unimaginable.

The quest in this direction has already begun with what is popularly known as Generative AI, i.e., Large Language Models (LLMs). Such visions of technology-centric future conflict have fascinated militaries around the world, driven by the potential of GAI and LLMs. The military, which can visualise the optimal trajectory for utilising such a tool, may be able to transform warfare itself. Currently, it is a commercial enterprise which leads the way in the research and development of GAI tools.

Military Applications

Militaries are also actively collaborating with commercial technology companies to develop an LLM which can be employed in combat applications. Reports indicate

¹ Varun Mehta, “[ChatGPT - An AI NLP Model](#)”, LTIMindtree POV Report, p. 4.

² Ibid.

³ Isaac Galtzer- Levi et al., “[The Cognitive Capabilities of Generative AI: A Comparative Analysis with Human Benchmarks](#)”, Arxiv, 9 October 2024.

⁴ Tyler Malloy and Cleotilde Gonzalez, “[Applying Generative Artificial Intelligence to Cognitive Models of Decision Making](#)”, *Frontiers in Psychology*, 3 May 2024.

that in July 2023, the Pentagon’s Digital and AI office experimented with five commercially available LLMs. The models were trained on ‘secret level data’ to answer sensitive combat-related questions. The exercises sought to incorporate AI-enabled data across all combat applications, including the deployment of LLMs to generate entirely new courses of action. The report highlights how LLMs have been found helpful in assisting planners in developing military responses to a brewing global crisis, with a minor incident as a trigger in another part of the globe, and the stage quickly moving to the Indo-Pacific region. Such a setting closely resembles the situation in Taiwan.⁵

Intelligence, Surveillance and Reconnaissance (ISR)

The *MIT Technology Review* has highlighted the use of LLMs by two US military officers, who, while on active exercises, employed them to scour open sources to gather intelligence on target countries and prepare intelligence reports. The AI tools they used were developed by Vannevar Labs, a company involved in defence-related AI research.⁶ A *Reuters* report has outlined how ‘six Chinese researchers from three institutions, including two from the People’s Liberation Army’s (PLA) leading research body, the Academy of Military Science (AMS), detailed how they had used an early version of Meta’s Llama as a base for what they call ‘ChatBIT’. The researchers developed a military-focused AI tool to gather and process intelligence and provide accurate, reliable information for operational decision-making.⁷ In October 2024, the Ordnance Science and Research Academy of China (OSRAC) filed a patent that proposed integrating data from OSINT, HUMINT, SIGINT, GEOINT and TECHINT sources to prepare a training set to train the LLM in intelligence tasks, so that it will support every phase of the intelligence cycle and improve decision-making during military operations.⁸

Planning, Simulation and Training

The US Marine Corps School of Advanced Warfighting tasked a team from Scale AI with adapting a planning exercise to examine how LLMs could assist military planning. The team trained an LLM on data related to theatre-level operations. The data comprised open-source intelligence, academic literature, doctrinal publications and other sources. The exercise aimed to examine the design of operations, activities

⁵ Katrina Manson, “[The US Military Is Taking Generative AI Out for a Spin](#)”, Bloomberg, 5 July 2023.

⁶ James O’Donell, “[Generative AI is Learning to Spy for the US Military](#)”, *MIT Technology Review*, 11 April 2025.

⁷ James Pomfret and Jessie Lang, “[Exclusive: Chinese Researchers Develop AI Model for Military Use on Back of Meta’s Llama](#)”, *Reuters*, 1 November 2024.

⁸ Zoe Harver, “[Artificial Eyes: Generative AI in China’s Military Intelligence](#)”, *Recorded Future*, 17 June 2025, p. 9.

and investments at the theatre level to deter an adversary in a competitive scenario short of war. The result was Hermes, an experimental LLM for military planning.⁹

China’s Joint Operations College, National Defence University, has launched a ‘Virtual AI Commander’. This tool aims to assume the role of a real-world Chinese commander in a simulated exercise. In exercises, tactical commanders could submit their plans and options to this AI commander and seek decisions or validation in return.¹⁰ Such tools could enhance combat creativity in real-world combat scenarios, based on lessons learnt in simulation exercises.

Shortening the OODA Loop

The PLA is reportedly deploying LLMs in different services for achieving the goal of ‘synchronised decision-making in a network-centric battlefield’. By integrating AI across domains, the LLMs are being tested in their abilities to reduce delays between data, decision and execution to a ‘fraction of a second for faster sensor-to-shooter loops’. LLMs have enhanced their abilities to employ ‘real-time intelligence’ to guide strike operations. The LLMs could integrate reconnaissance elements with precision-strike capabilities to rapidly close ‘kill chains’. Chinese reports suggest that DeepSeek is capable of such multi-domain integration. The final aim is to develop, as some US analysts call it, a ‘multi-domain kill-web’ that seamlessly integrates sensor capabilities with aircraft and missiles.¹¹

Logistics

Some reports indicate that China has deployed DeepSeek’s models in non-combat environments, such as hospitals and soldier training programmes, suggesting a low-risk environment for experimentation before expanding into high-stakes combat environments.¹² Other use cases of LLMs already employed in logistics management include finding the best paths and transport modes for a given condition, scheduling staff rotations, optimising energy utilisation, predicting equipment maintenance needs, mapping the physical and medical readiness of military personnel, and assisting in designing medical fitness programmes.¹³

Knowledge Management

The US Army Public Affairs drafted the entire press release regarding their own LLM using the same LLM. Named ‘The Army Enterprise Large Language Model

⁹Benjamin Jensen and Dan Tadross, “[**How Large Language Models Can Revolutionize Military Planning**](#)”, *War on the Rocks*, 12 April 2023.

¹⁰Jui Marthe and Chaitanya Giri, “[**Build Defence ‘Indic’ AI-language Models in India**](#)”, Expert Speak, Raisina Debates, Observer Research Foundation, 14 February 2025.

¹¹ Ibid.

¹²Rohith Narayan Stambamkadi, “[**China’s LLM Bet: The Push for AI-Driven Military Dominance**](#)”, Expert Speaks, Raisina Debates, Observer Research Foundation, 25 July 2025.

¹³Sarah Grand-Clement, “[**Artificial Intelligence Beyond Weapons – Application and Impact of AI in the Military Domain**](#)”, UNIDIR Report 2023.

Workspace’, it provides a platform for service personnel to harness sophisticated artificial intelligence tools to enhance ‘communication, operational efficiency, and drive innovation’.¹⁴ The LLM has reportedly improved data sharing and soldiers’ ability to access data.¹⁵ The Indian Navy is also experimenting with Samvaad.ai, an interactive chatbot that enables efficient data retrieval from vast datasets.¹⁶

Integrating Robots with LLMs

US companies have developed technologies such as Skynode-S boards, miniaturised computers optimised for LLM applications. These systems enable UAVs to lock on to and autonomously navigate to the target while performing other tasks. DeepSeek has also demonstrated integration of an LLM into the Xingji-P60, an autonomous military vehicle, incorporating self-driving software for dual-use applications.¹⁷

Futuristic /Aspirational Applications

Sun Yifeng, a researcher at the PLA’s Information Engineering University and his team are reportedly developing a GPT to ‘predict the behaviour of enemy humans’. Based on Baidu’s *Ernie* and iFlyTek’s *Spark* LLMs, simulation results have assisted decision-making and improved the machine’s knowledge and cognitive levels, as claimed in the article.¹⁸

Limitations of GAI/LLMs

Hallucinations: Responses to queries have factually incorrect information. This happens due to data quality issues, malicious data, or a poor understanding of the question, resulting in poor correlation between words.¹⁹

- **Opacity:** Algorithms used in NLP are opaque. It is difficult or impossible to understand how the neural network has arrived at responses.²⁰
- **Biased and Harmful Data:** Inherent biases inherited from malicious or tainted training data, such as gender or racial biases in facial recognition

¹⁴ [“Army Launches Army Enterprise LLM Workspace, The Revolutionary AI Platform That Wrote This Article”](#), US Army Public Affairs. [\[1\]](#)

¹⁵ Evan Lynch, [“Army’s New AI/LLM Tools Boost Productivity”](#), *Signal*, 1 August 2025.

¹⁶ Cyrus Ghosh, [“Large Language Model Revolution and Implications: Indian Perspective”](#), *ResearchGate*, October 2024.

¹⁷ Rohith Narayan Stambamkadi, [“China’s LLM Bet: The Push for AI-Driven Military Dominance”](#), Expert Speaks, Raisina Debates, Observer Research Foundation, 25 July 2025.

¹⁸ Christopher McFadden, [“China Train AI Generals to Predict ‘Enemy Humans’ in Battle”](#), *Interesting Engineering*, 14 January 2024.

¹⁹ Sarah Grand-Clement, [“Artificial Intelligence Beyond Weapons – Application and Impact of AI in the Military Domain”](#), no. 13.

²⁰ Fiona Fui-Hoon Nah, Ruilin Zheng et al., [“Generative AI and ChatGPT: Application, Challenges and AI-human Collaboration”](#), *Journal of Information Technology Case and Application Research*, Vol. 25, No. 3, 2023, pp. 277–304. available at [\[2\]](#)

systems, can skew responses. Also, biases inspired by geographical, social and cultural factors can distort information.²¹

- **Outdated Data:** Databases which have not been regularly updated can cause the LLM to deliver faulty results based on vintage data.²²
- **Limited Reasoning Capability:** LLMs can correlate but find it challenging to execute causation, leading to issues of reasoning.²³
- **Security of Data:** LLMs are known to leak data through memorisation unintentionally. This risk is compounded by model inversion attacks, where malicious actors could reconstruct classified data from LLM responses.²⁴
- **Vulnerability to Cyberattacks:** Adversarial attacks are designed to target weaknesses in the database by creating such inputs which manipulate or mislead the LLM into producing undesired or harmful outputs. Such attacks can further exacerbate these vulnerabilities, threatening the integrity of decision options generated by the AI tool.
- **Prompt Injection:** Malicious individuals use prompt injection to insert harmful directives into visibly harmless prompts, thus bypassing established safety measures. A user may seek guidance for illegal activities or seek confidential information by posing innocent questions. Such prompts can influence LLMs maliciously or lead to the spread of incorrect or harmful instructions.²⁵
- **Lack of Precision in Natural Language Programming:** Programming language is precise, whereas NLP is ambiguous. This is due to multiple context-based meanings of the exact words and complex grammar. Such ambiguity may lead to errors and inaccuracies in the LLM's responses.²⁶
- **Escalatory Responses:** LLMs exhibit ‘difficult-to-predict escalatory behaviour’ when employed to assist decision-making in a wargame.²⁷ This

²¹ Emanuela Marasco and Thirimachos Bourlai, “[Enhancing Trust in Large Language Models For Streamlined Decision-making in Military Operations](#)”, *Image and Vision Computing, Science Direct*, Vol. 15, 8 May 2025. [2](#)

²² Ibid.

²³ “[The Strengths and Limitations of Large Language Models](#)”, Eric D. Brown, 3 June 2024.

²⁴ Emanuela Marasco and Thirimachos Bourlai, “Enhancing Trust in Large Language Models For Streamlined Decision-making in Military Operations”, [no. 21](#).

²⁵ Justin Lavadia, “[Military Fortifies Frontline Against LLM AI](#)”, *Artificial Intelligence News*, 12 February 2025.

²⁶ Jamie Freestone, “[Are Large Language Models Safe for Military Use](#)”, Land Power Forum, Australian Army Research Centre, 6 February 2025.

²⁷ Juan Pablo Rivera, Gabriel Mukobi et al., “[Escalation Risks from LLMs in Military and Diplomatic Contexts](#)”, Policy Brief, HAI Policy & Society, Stanford University, May 2024.

raises questions about the reliability of the decision support framework provided by LLMs.

Further, a team of Google researchers tested LLMs currently in use on parameters of verbal comprehension, perceptual reasoning and working memory. The models' results are mixed at best. All models excelled at one parameter, achieving near-human-level results, while failing miserably in another. For example, all models performed well at recalling information from working memory but poorly at perceptual reasoning and processing visual information. Models performed well when tested only for memory or verbal comprehension, but poorly when perception tests were also included.²⁸ This indicates that the vision of an all-encompassing machine brain ready for deployment in real combat scenarios remains a distant objective.

Civil–Military Fusion: Choice or Compulsion?

The lure of owning an effective GAI/LLM seems strong enough for militaries to ignore limitations or possibilities of failure. However, no military has tread the path of committing individually to a technology that has shown only mixed results. Civil–Military fusion emerges as the only safe path.

On 5 March 2025, the DIU awarded a contract to Scale AI to develop comprehensive LLM-based AI solutions under the Thunderforge program for planning and decision-making at the theatre level.²⁹ On 16 June 2025, the US Department of Defense (DoD), through the CDAO, awarded a US\$ 200 million contract to OpenAI to develop a GAI model to ‘develop prototype frontier AI capabilities to address critical national security challenges in both warfighting and enterprise domains’.³⁰ Scale AI’s Donovan model claims to be the first LLM to work with classified military data for decision-making solutions within the XVIII Airborne Corps of the US Army.³¹ The generative tools used by two Marine Corps officers in developing an OSINT platform were created by the earlier-mentioned Vannevar Labs. In November 2024, this company was awarded a production contract worth up to US\$ 99 million by the Defense Innovation Unit to bring this intelligence analysis technology to more military units.³²

²⁸ Isaac Galtzer-Levi et al., “[The Cognitive Capabilities of Generative AI: A Comparative Analysis with Human Benchmarks](#)”, no. 3.

²⁹ “[DIU’s Thunderforge Project to Integrate Commercial AI-Powered Decision-Making for Operational and Theater-Level Planning](#)”, Project Spotlight, Defense Innovation Unit, 5 March 2025.

³⁰ “[Open AI Wins \\$200Million Contract with US Military](#)”, *The Hindu*, 17 June 2025.

³¹ Kathryn Harris, “[From Prototype to Production – Unlocking Mission Ready AI](#)”, *ScaleAI Blog*, 17 September 2025.

³² James O’Donell, “[Generative AI is Learning to Spy for the US Military](#)”, no. 6.

The US Marine Corps School of Advanced Warfighting tasked a team from Scale AI with adapting a theatre-level planning exercise called Hermes.³³ The Chinese PLA has actively engaged indigenous civil technology companies in this regard.³⁴ They have also involved LLMs like Llama from US technology giants such as Meta and Google.³⁵ The experiments discussed earlier strongly suggest that militaries have selected an open-source LLM and fine-tuned it by training the baseline LLM on additional text deemed to represent a particular domain adequately. After successful ‘fine-tuning’, the resulting models have exhibited greater understanding of the target domain and, therefore, potentially enhanced performance on domain-specific tasks. This result seems attractive to the militaries experimenting with them.³⁶

All the existing use cases are based on LLMs from a civil technology enterprise. At the current stage of GAI/LLMs, militaries seem hesitant to develop military LLMs without collaborating with civilian technology giants. The reason for this approach appears to be the enormous effort required to create an LLM. Frontier LLMs of the likes of ChatGPT4.5, Claude 4 Sonnet, Gemini, Llama 4 or Grok 4, are trained on trillions of parameters, require humongous databases, and consume vast resources in terms of computing power, funds, time and AI engineering talent. Only commercial technology behemoths like OpenAI, Meta, Google and X seem to possess the required wherewithal. Their single-minded focus on developing commercially successful AI tools and platforms is the key to their products' success.

Militaries may find it prohibitive to develop their own LLMs, as the US military has realised through its CamoGPT and NIPRGPT projects.³⁷ The time required to create an LLM is also a factor. By the time a new military LLM takes shape, the established LLMs of the day would be miles ahead, technologically and in sophistication of use.³⁸ Observers also apprehend that haste in pursuance of GAI/LLMs to solve specific military problems—for example, intelligence gathering from OSINT or retrieving particular answers to queries from unclassified databases—will lead to the development of ‘narrow use-case applications’ rather than technology that is based on a comprehensive understanding of warfare.³⁹ The civil-military fusion approach seems to offer militaries a safe avenue for experimenting with LLMs without the

³³ Benjamin Jensen and Dan Tadross, “[How Large Language Models Can Revolutionize Military Planning](#)”, no. 9.

³⁴ Zoe Harver et al., “[Artificial Eyes: Generative AI in China’s Military Intelligence](#)”, Recorded Future, 17 June 2025.

³⁵ James Pomfret and Jessie Lang, “[Exclusive: Chinese Researchers Develop AI Model for Military Use on Back of Meta’s Llama](#)”, Reuters, 1 November 2024.

³⁶ Daniel C Ruiz and John Sell, “[Fine-Tuning and Evaluating Open-Source Large Language Models for the Army Domain](#)”, 27 October 2024.

³⁷ Zachary Szewczyk, “[The Army Needs Frontier Models](#)”, Military Review Online Exclusive, August 2025, pp. 1–5.

³⁸ Ibid.

³⁹ Benjamin Jensen, “[Building a New Brain: Transforming Military Schoolhouses into AI Battle Labs](#)”, War on the Rocks, 28 August 2025.

enormous effort of developing one from scratch or the risk of losing a vast amount of funds if the project is abandoned for any reason.

Conclusion

The path of choice—or compulsion—to a robust Military LLM seems to be civil-military fusion. The advanced militaries of the world, namely the US military and the Chinese PLA, are actively engaging with commercial technology companies to integrate GAI/LLMs into military applications. The scale, pace and variety of this engagement clearly indicate that the technology competition in the GAI/LLM space is becoming more vigorous and is aimed at outmanoeuvring the competition to the point where the other needs to play catch-up. While LLMs have already been applied to ‘low-hanging fruits’ of military applications, entrusting them with sensitive classified data and tasking them in decision-making in active combat operations, may take some more time as the LLMs get fortified against their organic limitations and achieve greater perceptive and reasoning capabilities to match or surpass the good old human brain.

About the Author

Col. Nestor Derrick D'Souza is Research Fellow at the Manohar Parrikar Institute for Defence Studies and Analyses, New Delhi.

Manohar Parrikar Institute for Defence Studies and Analyses is a non-partisan, autonomous body dedicated to objective research and policy relevant studies on all aspects of defence and security. Its mission is to promote national and international security through the generation and dissemination of knowledge on defence and security-related issues.

Disclaimer: Views expressed in Manohar Parrikar IDSA's publications and on its website are those of the authors and do not necessarily reflect the views of the Manohar Parrikar IDSA or the Government of India.

© Manohar Parrikar Institute for Defence Studies and Analyses (MP-IDSA) 2025