NATO OTAN

NATO Defense College

Rome | Italy

FUTURE Series

# Decoding the dragon

## A language-based analysis of China's vision of AI warfare


**Anselm Küsters**

**Marius Strubenhoff**


INSIGHT


**8/2026** September 2026

The NATO Defense College was established in 1951 in Paris based on General Dwight D. Eisenhower's suggestion that the Alliance needed an institution that could "develop individuals both on the military and on the civilian side who will have a thorough grasp of the many complicated factors which are involved in creating an adequate defense posture for the North Atlantic Treaty Area." In 1967, it moved to Rome, where it has been located since.

NATO Defense College – Insight

Series editor: Florence Gaub
Copy-editing: Bettina Vestring, Mary Di Martino
The cover image was designed using ChatGPT.

Insight 8/2026
The Future Series

Nota Bene:
The views expressed in this NDC Insight are the responsibility of the author(s) and do not necessarily reflect the opinions of the NATO Defense College, the North Atlantic Treaty Organization, or any of the institutions represented by the contributor(s).

ISSN: 3006-5380 (print)
ISSN: 3006-5399 (online)



Limited copies of this NDC publication are available and may be obtained directly from

NATO Defense College,
Research Division,
Via Giorgio Pelosi 1, 00143
Rome, Italy
m.dimartino@ndc.nato.int

Printed and bound by
lightskyconsulting.com

# Contents

# Acknowledgments

This study was made possible through funding provided by the NATO Defense College. The views and conclusions expressed in this report are those of the authors and do not necessarily reflect the official positions or policies of the NATO Defense College nor of the North Atlantic Treaty Organization.

# The authors

**Anselm Küsters** is Scientific Director for the Nordic Countries and Head of the Department of Digitalisation and AI at the Centre for European Policy (cep). He also works in the field of Digital Humanities as an affiliated researcher at the Max Planck Institute for Legal History and Legal Theory in Frankfurt am Main and as a post-doctoral researcher (Habilitand) at the Humboldt University in Berlin. From April 2025 to March 2026, he held an Interim Professorship for Digital Humanities at the University of Stuttgart. Anselm holds a doctorate from Goethe University Frankfurt am Main (Dr. Phil.) and studied Economic and Social History at the University of Oxford, UK (M.Phil., Distinction), European Law at the University of Würzburg (LL.M.), and History, German Literature, and Economics at the University of Heidelberg (B.A. and B.Sc.). kuesters@cep.eu

**Marius Strubenhoff** works as a policy analyst at the Centre for European Policy, where he reports on EU foreign and security policy issues and French policy towards the EU. Previously, he worked as a foreign and security policy advisor in Berlin and as a lecturer at the University of Sheffield. Marius Strubenhoff was born and raised in Hamburg and has lived in Kyiv, London and Cambridge, among other places. He has been living in Paris since October 2025. He holds a PhD from the University of Cambridge and a Master's from the London School of Economics and Political Science. His commentaries on foreign and security policy has appeared in media outlets such as Frankfurter Allgemeine Zeitung, Table Media, and France 24. strubenhoff@cep.eu


Centres for European Policy Network
FREIBURG | BERLIN | PARIS | ROMA

Centrum für Europäische Politik
FREIBURG | BERLIN
Kaiser-Joseph-Straße 266 |
D-79098 Freiburg
Schiffbauerdamm 40 Räume 4205/06 |
D-10117 Berlin
Tel. + 49 761 38693-0

Centre de Politique Européenne PARIS
17, rue Saint Fiacre | F-75002 Paris
Tel. +33 1 88 80 72 00

Centro Politiche Europee ROMA
Via A. Brunetti, 60 | I-00186 Roma
Tel. +390684388433

The Centrum für Europäische Politik FREIBURG | BERLIN, the Centre de Politique Européenne PARIS and the Centro Politiche Europee ROMA form the Centres for European Policy Network FREIBURG | BERLIN | PARIS | ROMA.

Free of vested interests and party-politically neutral, the Centres for European Policy Network provides analysis and evaluation of European Union policy, aimed at supporting European integration and upholding the principles of a free-market economic system.

1

# Executive summary

**China's pursuit of "intelligentized warfare" has become a central concern for NATO member states**, whose 2022 Strategic Concept identified China as presenting systemic challenges. However, assessments of Beijing's military AI trajectory often rest on anecdote. **To fill this gap, this study conducts a large-scale computational analysis of China's own documentary record from 2000 to 2025**, including:

1. legal and strategic texts providing the formal doctrinal baseline;
2. Chinese AI research publications tracing the technical discourse;
3. Media coverage (*People's Daily*), capturing propaganda and public signalling; and
4. PLA-linked procurement and defence-industrial contracts.

Based on the empirical results and our literature review, four core findings emerge:

**First**, sentiment analysis of state media consistently shows high positivity around military AI. Moreover, AI-military salience and positive tones are increasingly coupled after 2023. **The more a particular text addresses the topic, the more its tone is positive.** Taken together, this indicates a deliberately optimistic propaganda strategy.

**Second, LLM-based classification of AI and military concepts across corpora reveals that surveillance dominates as the primary application of military AI**. However, this concentration should not be interpreted as evidence of a purely defensive Intelligence, Surveillance and Reconnaissance (ISR) posture: Under China's dissipative warfare doctrine (which envisions systematically eroding an adversary's will and capacity to fight through sustained, multidimensional attrition across military, economic and informational dimensions, heavily involving attacks on the homefront, both prior to and during armed conflict) surveillance capabilities are offensive enablers for surprise strikes and decision-cycle compression in contested theatres like the Taiwan Strait. Indeed, there are strong co-occurrence links to targeting and command systems. Near-term PLA investments appear geared towards achieving a battlespace transparency that supports both information dominance and precision strikes. This aligns with the Military Strategy adopted by the Central Military Commission in 2004. References to "autonomous weapons" are often displaced by euphemistic labels such as "autonomous systems" or "unmanned platforms."

**Third, we find a recent shift from generic talk of "AI" towards a more granular discussion of specific techniques related to deep learning and computer vision**. Likewise, there are increasingly concrete references to military applications of AI after 2020. This suggests movement from aspiration to operational integration, particularly in ISR and targeting functions.

**Fourth, Chinese AI governance documents strongly emphasize human control, strategic stability and technology-denial concerns while avoiding any explicit discussion of the interplay between AI and nuclear weapons or demographic compensation**. For AI diplomacy, this offers normative entry points around responsible use and human oversight but suggests that the most sensitive military AI questions are being discussed outside the regulatory texts that are publicly available.

For policymakers, these results imply that **China's immediate focus on military AI lies less in fully autonomous lethal systems than in AI-enabled sensing, targeting and decision support**. These capabilities compress decision times and increase escalation risks even without formally removing humans from the loop. This prioritization creates both risks and diplomatic opportunities. The dominance of surveillance applications suggests that **near-term crisis instability will stem from information asymmetries rather than from rogue autonomous weapons**. At the same time, the combination of positive state narratives and the decline in "meaningful human control" rhetoric after 2021 signals **China's growing confidence and waning commitment to international norms**. Yet Western responses based on worst-case assumptions about autonomous weapons in the narrower sense (as opposed to AI-fueled military intelligence and decision assistance) could fuel an AI arms race. **More promising avenues include targeted transparency regarding nuclear-adjacent AI systems, technical verification mechanisms that leverage hardware-level governance and monitoring of the Chinese discourse** to detect operational shifts before deployment.

## Specific recommendations

**Strengthen total defence as a key aspect of deterrence**. The close relationship between the prominence of AI in the military and optimistic sentiment surfacing in our data, when considered alongside China's dissipative warfare doctrine, supports the claim for the strategic importance of total defence, particularly in terms of civil defence and cyber security. Societal resilience and cyber hardening are assets for deterrence in an information-domain contest, not merely defensive measures.

**Invest in countermeasures targeting China's core AI advantages**. Three priority areas follow directly from our findings in the application domain: First, NATO members should invest in adversarial robustness research to develop AI systems that are resilient to "data poisoning" (injecting/creating malicious data to corrupt pattern recognition), model extraction (reconstructing proprietary models through repeated queries) and attacks that can blind, misdirect or compromise military decision-making systems. Second, they should invest in counter-ISR technologies, including adaptive camouflage, decoy generation and emissions management to degrade potential surveillance-targeting kill chains. Third, they should invest in human-machine teaming architectures that demonstrate how Human-in-the-loop (HITL) systems can match or surpass autonomous systems in decision speed and accuracy, countering the narrative that human control requires accepting operational inferiority.

**Extend and deepen the American-Chinese nuclear AI understanding**. The 2024 understanding should be complemented by agreed definitions of which AI applications in nuclear infrastructure are admissible (e.g. safety analysis and anomaly detection) and which are not. Diplomatic efforts should extend these commitments to other nuclear-armed states.

**Pursue binding international commitments on high-risk applications**. Treaty obligations should target autonomous target selection and "non-reconstructible" black box military AI systems, whose battlefield actions cannot be audited for compliance with international humanitarian law. We propose the term "human-beyond-the-loop" for this category to distinguish it analytically from "human-in-the-loop" and "human-on-the-loop" architectures.

**Build verification capacity**. Our corpus analysis confirms that genuine discourse on the integration of AI in the military is almost entirely absent from open Chinese sources. The significant disparity between China's substantial civilian AI research output and its limited public military discourse means that open-source monitoring alone cannot provide a basis for verification. Effective arms control requires technical measures that are both intrusive and analogous to the inspection protocols of the Chemical Weapons Convention. These measures must include shared standards for AI system documentation, testing protocols for autonomous weapons and cryptographic on-chip governance mechanisms (including location-verification features) that can confirm that the hardware has not been redeployed to unauthorized facilities.

**Form coalitions to promote meaningful CCW regulation**. Pursue binding LAWS commitments through the Review Conference for the United Nations Convention on Certain Conventional Weapons (CCW) while also countering China's selective norm strategy within this forum. Coalition building must extend to non-Western states subject to Chinese influence, including Belt and Road partners, given that China increasingly uses these relationships in its diplomatic strategy to dilute multilateral AI governance efforts. Our findings regarding international law and governance provide the evidential basis for explicitly naming this pattern in diplomatic settings.

# 2

# Introduction

**China's rapid pursuit of "intelligentized warfare" has become a central concern for NATO and its member states**. The Alliance's 2022 Strategic Concept identified China as posing "systemic challenges" to Euro-Atlantic security,[1] and subsequent summit declarations have linked Chinese military AI development to risks to alliance deterrence and critical infrastructure. Notably, the 2024 Washington Summit Declaration saw the Alliance commit to countering AI-enabled threats while accelerating its own adoption of military AI.[2] In October 2023, the Pentagon's annual China Military Power Report identified China as a global leader in AI technology and warned that it aims to overtake the West in AI research and development (R&D) by 2025, becoming the world leader in AI by 2030.[3] This assessment, already voiced by former Google CEO Eric Schmidt in 2017 and (relating to military AI) Senator John McCain's former Chief of Staff Christian Brose in 2020,[4] reflects a broader consensus among Western policymakers that China harbours expansive military ambitions powered by AI technologies. Yet despite these strong concerns, Western policy responses – including NATO's emerging AI governance frameworks[5] and the EU AI Act's exemptions for national security – **often lack systematic evidence about the content of the Chinese military AI discourse and its priorities**. To address this gap, this study conducts a **large-scale computational analysis of China's documentary record from 2000 to 2025**.

This study therefore asks: **What do China's own public records reveal about its actual AI warfare priorities?**[6] This research question is then broken down into operational sub-question:. Above all, what topics do Chinese commentators and researchers mention consistently with respect to the military AI domain? Which AI technologies receive sustained government investment in research, procurement and regulation? How do the priorities articulated by different institutional actors or text sources align or diverge, and what do such divergences reveal about the gap between rhetoric and implementation? Answering these questions is crucial to inform sensible policymaking. Correctly understanding the intentions of a strategic rival should be firmly at the centre of attention. **Overestimating, underestimating and (most importantly) misinterpreting Chinese military AI capabilities and intentions could prove very costly.**[7] As NATO member states consider significant investments in military AI to counter China's vision of intelligentized warfare (for instance, through several CHIPS Acts and export controls), it is crucial to understand China's actual trajectory in order to calibrate appropriate responses. Overestimating Chinese capabilities risks wasteful expenditure and unnecessary escalation. Conversely, underestimating them could lead to strategic

1 See: NATO, NATO 2022 Strategic Concept, Madrid, 29 June 2022, p. 5.

2 The Washington Summit Declaration explicitly reaffirms China as posing systemic challenges to Euro-Atlantic security, expresses concern about China's space capabilities and announces the implementation of a revised Artificial Intelligence Strategy. See: NATO, Washington Summit Declaration, Washington, 10 July 2024.

3 See: Patrick Tucker, "China Will Surpass US in AI Around 2025, Says Google's Erich Schmidt," *Defense One*, November 2017, US DoD, Military and Security Developments Involving the People's Republic of China, 2023, p. 196.

4 See: https://www.defenseone.com/technology/2017/11/google-chief-china-will-surpass-us-ai-around-2025/142214/; as well as: Christian Brose, *The Kill Chain: Defending America in the Future of High-Tech Warfare*, New York: Grand Central Publishing, 2020.

5 Adopted in 2021, NATO's AI Strategy focuses on the responsible development of AI, enhanced interoperability, the vigilant protection of AI technologies and safeguarding against malicious use of AI. See: https://www.nato.int/en/about-us/official-texts-and-resources/official-texts/2021/10/22/summary-of-the-nato-artificial-intelligence-strategy.

6 For a general answer to this question, albeit without a focus on AI, see this recent paper: David C. Kang et al., 'What Does China Want?', *International Security* 50, no. 1 (2025): 46–81, https://doi.org/10.1162/ISEC.a.5.

7 Note the cautionary message of Kevin Rudd, *The Avoidable War: The Dangers of a Catastrophic Conflict between the US and Xi Jinping's China*, New York City: Public Affairs, 2022.

surprise, particularly if China were to achieve breakthroughs in cognitive domain operations that are central to its warfare concept. Misinterpreting its capabilities would lead to setting the wrong response priorities.

Our contribution to this crucial debate is to ground it in systematic, computational analysis, replacing speculation about China's AI warfare vision with evidence extracted from its own documentary record. If policymakers are to navigate an uncertain future, they deserve, at the very least, an accurate computational decoding of the present. To go beyond speculation, we engage in rigorous analysis of observable behaviours across multiple institutional domains. This study aims to conduct such an analysis through the lens of **natural language processing (NLP)** on a large corpus of unstructured but rich text data spanning 2000 to 2025.

The shift towards computational text analysis in political science is driven by necessity and opportunity. It is a necessary development because traditional close-reading methods cannot be applied to corpora comprising millions of documents. It is an opportunity because advances NLP and generative AI now enable the extraction of structured insights from unstructured text with an unprecedented level of precision.[8] This study employs a methodological toolkit that spans established techniques and frontier approaches involving large language models (LLMs) that have been fine-tuned for domain-specific classification.

## Structure

This study consists of six main sections, a glossary and a detailed appendix. The Executive Summary (Section 1) highlights our main findings and their significance for understanding China's evolving approach to intelligentized warfare, followed by an introduction (Section 2).

Readers interested in policy and seeking immediate insights can proceed directly to Section 6 (results), which presents our empirical findings on China's AI technologies, warfare priorities and trends. For researchers interested in replicability and methodology, Sections 3 to 5 provide the necessary background information.

Section 3 reviews the existing literature on military AI development and China's intelligentization doctrine. First introduced by the PLA in July 2019, "intelligentization" builds upon what China called the "informatization" of warfare in its doctrinal reform agenda in the wake of the First Gulf War (1991). Intelligentization refers to the growing application of AI to military purposes and notably entails a prominent cognitive warfare element.[9] Based on this literature review, section 3 derives testable hypotheses about institutional alignment, policy diffusion and technology prioritization.

Section 4 briefly describes the empirical foundation of this study, which consists of four datasets: AI governance documents published in China; Chinese research in the field of machine learning; a large Chinese newspaper dataset; and relevant procurement contracts. More details on these four datasets, including technical information on how they are harmonized and pre-processed, can be found in Appendix (A).

Section 5 explains our LLM-assisted classification procedures. Appendix (B) includes a general primer on the methodology of NLP and LLM-based classification and provides further technical information to enable replication.

Section 6 presents our results, integrating findings across datasets to reveal patterns in China's AI warfare development trajectory, including temporal trends, cross-institutional dynamics and the discrepancies between speculation and implementation. First, the section presents some basic keyword frequencies. Then, it analyses the sentiment of state media discourse, the LLM-based classification of the implications of military AI warfare and hypothesis-driven concept induction via LLM.

Concluding, Section 7 discusses the theoretical and strategic implications for military AI governance, provides policy recommendations and suggests areas for future research. A detailed glossary lists and explains all important concepts and technical terms used throughout this study, while the appendix provides supplementary materials.[10]

## Methods and data

What is the basic conceptual idea behind this research project? Rather than relying on isolated policy pronouncements or ad hoc case studies, as has been often done in the past, our aim is to construct an integrated, empirical portrait by analysing legal frameworks, academic research outputs, military procurement contracts and high-level policy communications.[11] The resulting picture enables us to test hypotheses derived from the qualitative, policy-focused literature on China and described in Section 2 empirically. Our NLP-driven approach enables us

8 For overviews, see: Justin Grimmer et al., Text as Data: *A New Framework for Machine Learning and the Social Sciences* (Princeton University Press, 2022); Ken Benoit, 'Text as Data: An Overview,' in *The SAGE Handbook of Research Methods in Political Science and International Relations,* by Luigi Curini and Robert Franzese (SAGE Publications Ltd, 2020); Matthew Gentzkow et al., 'Text as Data', Journal of Economic Literature 57, no. 3 (2019): 535-74, https://doi.org/10.1257/jel.20181020.
9 Koichiro Takagi, 'New Tech, New Concepts: China's Plans for AI and Cognitive Warfare', *War on the Rocks*, 13 April 2022.
10 The complete Python codebase for reproducing our analyses has been submitted with this study.
11 For the limitations of such an approach, which relies on public documents, see the reflections in the conclusion.

to identify implicit priorities and detect discrepancies between rhetoric and implementation on a scale that would be impossible in a manual review. The empirical foundation for this NLP analysis is based on four datasets, which we describe in detail in Section 3: First, a corpus of Chinese legal and strategic documents on AI and warfare provides the formal doctrinal baseline. Second, an overlay of Chinese AI research publications captures the evolution of the technical-scientific discourse. Third, a large-scale collection of Chinese newspaper articles traces shifts in official propaganda and public signalling around AI and military power. Fourth, a dataset of PLA-related procurement and defence-industrial contracts grounds the analysis in observable military capability development.

Together, these sources facilitate a temporal analysis of pivotal moments at the intersection of China and AI. Such moments include the US *Third Offset Strategy* announced and elaborated in 2014-15, which explicitly prioritized robotics, autonomous systems and big data as the cornerstones of future US military superiority;[12] AlphaGo's victory over Go world champion Lee Sedol in March 2016, which was widely interpreted in China as a transformative moment akin to the launch of the Sputnik satellite;[13] the launch of China's *New Generation AI Development Plan* in July 2017, which established China's ambition to lead the global AI industry by 2030[14] and the *Made in China 2025* industrial policy, signed by Premier Li Keqiang in May 2015, which set the goal of reaching technological self-sufficiency in ten strategic sectors, including robotics and advanced manufacturing.[15]

A persistent challenge in studying Chinese strategic thinking is that existing analyses often reflect interpretations of Chinese rhetoric – by external scholars, e.g. from the US or Europe – rather than the rhetoric itself. This interpretive gap stems partly from limited data availability and the linguistic reality that most Chinese government documents, academic publications and military procurement contracts remain untranslated. **In rapidly developing domains like AI warfare, researchers face an acute temporal constraint: Waiting for official or other human-made translations means analysing yesterday's debates**. In this context, multimodal and multilingual large language models offer a novel methodological pathway. Recent advances in frontier models and specialized Mandarin-language variants enable near-instantaneous processing of Chinese text corpora at scales previously unattainable. As we describe in Section 5, we thus test whether LLMs might support a systematic and scalable engagement with primary Chinese-language sources.

Our approach departs from conventional analyses in three other ways. Firstly, whereas existing studies usually examine single dimensions – either policy rhetoric, research publications or procurement data – we integrate evidence from multiple angles. With this, we aim to identify discrepancies between stated intentions and actual implementation as well as between research, mere media communication and operational deployment. Secondly, we use computational methods to detect patterns that would be invisible to qualitative review alone, extracting structured insights from thousands of Mandarin-language documents. The scale of our datasets makes such methods necessary while enabling empirical precision when decoding how China articulates and pursues its vision of intelligentized warfare. Thirdly, we contextualize China's AI warfare development, acknowledging that current AI investments are rooted in decades of doctrinal evolution towards what the PLA refers to as the "three-izations" (三化): mechanization, informatization and intelligentization.[16]

## Limitations of public information analysis

A key caveat of this study's approach merits emphasis and discussion at the outset. Our **analysis relies on publicly available information** as the classified programmes that are central to China's ambitions in intelligentized warfare remain inaccessible. In this sense, we analyse observable behaviour, such as publications, contracts and regulation, rather than unobservable intentions and behaviour. However, two factors make these public sources insightful for our purposes. First, as a large body of literature has pointed out, public information can also generate political costs: For example, leaders who escalate disputes verbally and then back down may face domestic punishment from constituents who perceive weakness or incompetence.[17] And this is not just true for the population: The drive to foster party oversight over the PLA increased significantly after Xi Jinping took

12 Chuck Hagel, Reagan National Defense Forum Keynote, US Department of Defense, 15 November 2014. For analytical context, see Robert Work, The Third U.S. Offset Strategy and its Implications for Partners and Allies, US Department of Defense, 28 January 2015.

13 Tanguy Chouard, 'The Go Files: AO Computer Wraps up 4-1 Victory against Human Champion', *Nature*, 15 March 2016, nature.2016.19575, https://doi.org/10.1038/nature.2016.19575.

14 State Council of the People's Republic of China, New Generation Artificial Intelligence Development Plan, 20 July 2017.

15 State Council of the People's Republic of China, Made in China 2025, 8 May 2015.

16 Sunny Cheung, 'AI Frontlines: China's PLA Evolution vs. NATO's Counterplay', Friedrich Naumann Foundation for Freedom, 24 November 2023, https://www.freiheit.org/taiwan/ai-frontlines-chinas-pla-evolution-vs-natos-counterplay.

17 This is known as audience cost theory. See, e.g.: James D. Fearon, 'Domestic Political Audiences and the Escalation of International Disputes', *American Political Science Review* 88, no. 3 (1994): 577–92, https://doi.org/10.2307/2944796; James D. Fearon, 'Signaling Foreign Policy Interests: Tying Hands versus Sinking Costs', *Journal of Conflict Resolution* 41, no. 1 (1997): 68–90, https://doi.org/10.1177/0022002797041001004; Zenobia T. Chan et al., 'Audience Costs and the Credibility of Commitments', in *International Relations* (Oxford University Press, 2021), https://doi.org/10.1093/obo/9780199743292-0305.

office as General Secretary of the CCP and Chair of the Central Military Commission (CMC) in 2012, culminating in the dismissal of Defence Secretary Li Shangfu and the leaders of the PLA Rocket Force.[18] It continues to this day, with a particularly high-profile dismissal occurring in late January 2026, when General Zhang Youxia and General Liu Zhenli were ousted.[19] This continuing drive to enforce absolute loyalty through purges suggests that the Chinese leadership is acutely aware of the army's stakeholder role. The failed coup against Mao Zedong in 1971 still lingers in perceptions, and the fact that views on security matters tend to be more hardline among PLA leaders suggests that reneging on verbal escalation can be costly for civilian leaders. These costs make public statements informative: Threats issued through official channels signal genuine resolve precisely because retreat may carry penalties. Crucially, the mechanism applies beyond democracies. Autocratic regimes also face constraints when official media outlets articulate positions, as domestic audiences use consistency between words and deeds to evaluate leadership competence. As researchers on the Chinese media system have noted, "the internal media served the process of policy formulation and deliberation, while the open media served the process of explaining the policy once implemented (or at least formulated) and mobilizing the population to support it."[20] Secondly, Chinese leaders are acutely aware that their public statements are being read around the world and therefore have to be carefully calibrated.[21] Sometimes, rhetoric may be a ruse to confound adversaries. More often than not, however, China's top brass will take care that rhetoric is at least to a considerable extent in line with reality (not least in the nuclear realm). The days of the maxim of "biding time and hiding capabilities" of the Deng Xiaoping era are long gone.

**Official rhetoric thus represents more than propaganda; it constitutes strategic communication constrained by domestic accountability structures**. This theoretical foundation justifies a systematic analysis of official text data with the help of Natural Language Processing: If public statements prompt states to take courses of action, then computational analysis of such texts can reveal priorities, track commitment evolution and help forecast behavioural shifts. Moreover, China's political setup, centrally governed by the Chinese Communist Party's leadership, makes statistical analysis of political statements even more robust than comparable analyses of shifting sentiments in Western democracies. Since statements in Chinese outlets are generally more aligned with one another and indeed often follow a centrally set party line, statistical evidence of linguistic change is potentially very insightful.[22] Given China's emphasis on civil-military fusion, we thus think it plausible that open-source texts and activities reflect a significant proportion of its actual capabilities and intentions, particularly in foundational AI research and industry.

18 Chris Buckley, 'China Dismisses Defense Minister Amid Swirl of Speculation', *New York Times*, 24 October 2023.
19 Cate Cadell and Christian Shepherd, 'China fires top general in shocking purge of senior military command', *Washington Post*, 25 January 2026.
20 Joseph Fewsmith et al., *Studying China in the Absence of Access: Rediscovering a Lost Art*, Research Report (Johns Hopkins University School of Advanced International Studies, China Research Center, 2024), 23, https://scgrc.sais.jhu.edu/wp-content/uploads/2024/10/32026_JOHNS_HOPKINS.COVER_SP.pdf.
21 Jackie Wong, 'Bluster vs. Hands-Tying: Evidence from Word Embedding Analysis on China's Rhetoric and Military Escalations', *SSRN Electronic Journal*, ahead of print, 2024, 2, https://doi.org/10.2139/ssrn.4809057.
22 For a paper that documents text-reuse homogeneity across Chinese newspapers and demonstrates that computational text analysis of Chinese media reveals politically meaningful signals, see: Hannah Waight et al., 'The Decade-Long Growth of Government-Authored News Media in China under Xi Jinping', *Proceedings of the National Academy of Sciences* 122, no. 11 (2025): e2408260122, https://doi.org/10.1073/pnas.2408260122.

# 3

# Literature review and hypotheses

To understand China's vision for AI warfare, we must move beyond assessing what China can do and examine what it believes it should do. Our approach is based on the assumption that strategic choices in the military domain reflect technological capabilities as well as institutional beliefs about future conflict, perceived vulnerabilities and theories of victory.[23] Chinese military writings on "intelligentized warfare" provide a rich documentary record of such beliefs.[24] However, systematic analysis of how these visions translate into observable research priorities, procurement patterns and regulatory frameworks remains scarce. This section reviews existing literature on seven key aspects of China's AI warfare vision, deriving testable hypotheses to guide our subsequent empirical analysis. The hypotheses are generated through a structured deductive approach: For each thematic domain, we identify relevant literature and plausible theoretical arguments. We then specify which observable evidence would confirm or refute the related arguments in our corpus and formulate a hypothesis. This approach follows the logic of hypothesis-driven analysis in the Digital Humanities, in which theoretical propositions derived from prior qualitative literature are operationalized into measurable patterns that can be extracted on a large scale from textual data.

1. **Demographic constraints and AI-enabled force reduction**: China faces a well-documented demographic challenge: a rapidly aging population, declining birth rates and a shrinking pool of military-age population.[25] While current personnel concerns of the PLA centre heavily on the quality (health, education, resilience) rather than the quantity of recruits, the real impact of the demographic crisis will be felt only in the coming years and decades.[26] The question of whether China envisions AI as a solution to manpower constraints intersects with broader debates about the role of automation in military modernization. While it seems intuitive that states facing demographic decline have stronger incentives to invest in labour-saving military technologies, this could potentially be counteracted by the fact that personnel-intensive doctrines can create path dependencies resistant to automation. Chinese military statements provide striking evidence for the former view. A senior executive at China's third-largest defence company stated in 2018 that "in future battlegrounds, there will be no people fighting," predicting that by 2025 lethal autonomous weapons would be commonplace.[27] Research from Xi'an Technological University claims that AI-powered systems can evaluate 10,000 battlefield scenarios in 48 seconds, a task that would cost human planners 48 hours.[28] The PLA's focus on unmanned intelligent combat systems and swarm tactics suggests a

23 Here, we are indebted to the emphasis on the importance of the culture of military institutions in AI adoption: Kenneth Payne, *I, Warbot: The Dawn of Artificially Intelligent Conflict*, Oxford Scholarship Online Political Science (Hurst & Company, 2021), https://doi.org/10.1093/oso/9780197611692.001.0001.
24 For an example, see: Kevin Pollpeter and Amanda Kerrigan, The PLA and Intelligent Warfare: A Preliminary Analysis (Arlington, 2021).
25 For an overview, see: ChinaPower, "How Severe are China's Demographic Challenges?", CSIS (Center for Strategic and International Studies), March 2023, https://chinapower.csis.org/china-demographics-challenges/.
26 Mark Cozad and Jennie W. Wenger, *Factors Shaping the Future of China's Military*, Research Report (RAND Corporation, 2025).
27 Gregory C. Allen, *Understanding China's AI Strategy: Clues to Chinese Strategic Thinking on Artificial Intelligence and National Security*, Report (Center for a New American Security (CNAS), 2019), https://www.cnas.org/publications/reports/understanding-chinas-ai-strategy.
28 See: Eduardo Baptista and Fanny Potkin, "How China Could Use DeepSeek and AI for an Era of War," *Reuters*, October 2025, https://www.reuters.com/world/asia-pacific/robot-dogs-ai-drone-swarms-how-china-could-use-deepseek-an-era-war-2025-10-27/.

deliberate strategy to avoid "high personnel losses" while achieving scale through low-cost autonomous platforms.[29]

These observations lead to our first hypothesis: If China genuinely views AI as compensating for demographic constraints, we should observe it systematically prioritizing autonomous systems research, increased procurement of unmanned platforms and policy language explicitly linking intelligentization to force structure optimization. Moreover, temporal analysis should show that the implementation of these priorities is accelerating – a trend that coincides with worsening demographic projections post-2015.[30] Conversely, if demographic concerns are merely rhetorical justification for investments driven by other motivations, we would expect only a weak correlation between demographic indicators and military AI spending patterns.

*H1: China's AI warfare vision explicitly addresses demographic constraints through force reduction, observable in increased emphasis on autonomous systems in research publications post-2015, procurement patterns favouring unmanned platforms or policy documents linking intelligentization to personnel optimization.*

2. **Nuclear deterrence and strategic stability**: The intersection of AI and nuclear weapons poses profound questions about strategic stability. Schelling's classic analysis of deterrence assumes that human decision-makers will balance risks.[31] However, technological advances might undermine second-strike capabilities: AI introduces unprecedented speed in targeting, potentially compressing decision timelines below human reaction thresholds.[32] Observers have even speculated about the end of mutually assured destruction, as the use of AI could simplify "left of launch" attacks[33] on the nuclear Command, Control and Communications infrastructure of adversaries and complicate second-strike capabilities through super-charged AI surveillance technologies targeting submarines as well as land- and air-based systems.[34] The November 2024 Biden-Xi agreement that AI systems should not be given control over the decision to use nuclear weapons suggests Chinese recognition of escalation risks.[35] Yet China's investments in AI-enabled early warning systems, autonomous submarines and "intelligence supremacy" concepts for command decision-making all intersect with nuclear stability. In this way, nuclear strategy is entering a new era that is nonetheless shadowing Cold War strategic considerations surrounding automation, for instance concerning the Soviet "dead hand" and RYaN systems.[36]

Clearly, super-human AI capabilities are threatening conventional-nuclear firebreaks and create pressures for doctrinal adaptation: "If leaders believe that AGI [Artificial General Intelligence[37]] development will create a decisive and irrevocable shift in the balance of power that will leave them at the mercy of enemies committed to their destruction, and if they believe that they can use force to prevent that outcome while avoiding escalation to a general war that could guarantee the same fate, then they might roll the iron dice."[38] If China perceives AI as advantageous for pre-emption or crisis stability, we should observe research emphasis on early warning and time-critical targeting. If China worries about adversary AI undermining its retaliatory capability, we should expect countermeasures in resilience and survivability.

*H2: China's AI warfare vision reflects strategic stability concerns, observable in research on AI-enabled early warning and autonomous second-strike systems or policy advocacy for international AI-nuclear safeguards.*

3. **International law and institutional strategy**: The governance of military AI involves efforts to shape emerging rules in favourable directions. China has actively participated in UN discussions

29 Jiayu Zhang, 'China's Military Employment of Artificial Intelligence and Its Security Implications', *International Affairs Review*, August 2024, https://www.iar-gwu.org/print-archive/blog-post-title-four-xgtap.
30 "The PRC is facing adverse demographic trends, such as an aging population and low-birth rates." Annual Report to Congress: Military and Security Developments Involving the People's Republic of China, p. 149, https://media.defense.gov/2024/Dec/18/2003615520/-1/-1/0/MILITARY-AND-SECURITY-DEVELOPMENTS-INVOLVING-THE-PEOPLES-REPUBLIC-OF-CHINA-2024.PDF.
31 Thomas C. Schelling, *The Strategy of Conflict: With a New Preface*, Nachdr. d. Ausg. 1980 (Harvard Univ. Press, 20).
32 Dan Hendrycks and Adam Khoja, 'AI Deterrence Is Our Best Option', *AI Frontiers*, 22 September 2025, https://ai-frontiers.org/articles/ai-deterrence-is-our-best-option.
33 This term denotes offensive cyber operations, electronic warfare or kinetic operations targeting an adversary's nuclear infrastructure before nuclear weapons are launched. See Glossary.
34 "We introduce the concept of Mutual Assured AI Malfunction (MAIM): a deterrence regime resembling nuclear mutual assured destruction (MAD) where any state's aggressive bid for unilateral AI dominance is met with preventive sabotage by rivals." Dan Hendrycks et al., 'Superintelligence Strategy: Expert Version', version 2, preprint, arXiv, 2025, https://doi.org/10.48550/ARXIV.2503.05628. See also: https://www.europeanleadershipnetwork.org/wp-content/uploads/2023/11/AVC-Final-Report_online-version.pdf.
35 See: https://www.reuters.com/world/biden-xi-agreed-that-humans-not-ai-should-control-nuclear-weapons-white-house-2024-11-16/.
36 John Borrie, 'Cold War Lessons for Automation in Nuclear Weapon Systems', in *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk*, I (Euro-Atlantic Perspectives), ed. Vincent Boulanin (Stockholm International Peace Research Institute, 2019).
37 Commonly defined as AI systems that can match or surpass human cognitive performance in all areas rather than in specific, task-oriented applications.
38 Zachary Burdette and Hiwot Demelash, *Evaluating the Risks of Preventive Attack in the Race for Advanced AI*, Perspective PE-A3691-13 (RAND Corporation, 2025), https://www.rand.org/pubs/perspectives/PEA3691-13.html.

on lethal autonomous weapons systems (LAWS) and the broader REAIM (Responsible AI in the Military Domain) process[39] while resisting any binding constraints that might limit its technological options. China's consistent non-endorsement of the REAIM declarations at summits in The Hague (2023), Seoul (2024) and La Coruña (2026) is a prime example of this pattern of selective engagement.[40] Even as approximately 60 other nations endorsed earlier iterations, China declined to sign successive non-binding "blueprints for action" affirming human responsibility over AI-powered weapons and clear chains of command and control. This mirrors historical patterns: China joined the Nuclear Non-Proliferation Treaty only after acquiring nuclear weapons,[41] and it participates in arms control regimes selectively based on strategic advantage. The 2023 Bletchley Declaration and subsequent AI Safety Summits reveal convergence on process (dialogue) but divergence on substance (acceptable applications).[42] China's dual approach suggests that it is indeed instrumentalizing international law.[43]

According to the rationalist theory of international law, states bind themselves only when doing so serves narrow self-interest. If China views international AI regulations as constraining adversaries more than itself, we should observe advocacy for rules targeting capabilities where China lags (e.g., AI safety standards) while resisting rules constraining areas of strength (e.g., autonomous swarms). A previous analysis found common ground on evaluations and content provenance but divergence on military applications.[44]

*H3: China's approach to military AI governance reflects strategic positioning, observable in advocacy patterns in UN forums correlating with technological strengths/weaknesses, domestic regulations emphasizing different priorities than international proposals or citation networks showing instrumental use of international norms in domestic policy documents.*

4. **Comparative assessment of capabilities**: Capital-intensive military innovations typically favour powerful established actors, while low-cost innovations can disrupt the existing international power balance by favouring upstart newcomers.[45] How AI will impact such dynamics is therefore a pressing question. Secondly, accurate self-assessment of one's own military capabilities makes all the difference between rational adaptation to technological changes in warfare and overstating existing or future threats ("threat inflation"). In this sense, states might construct narratives about relative capabilities that diverge from objective metrics. Chinese writings frequently describe AI as a "leapfrog opportunity" enabling asymmetric advantages over technologically superior adversaries.[46] Specifically, China identifies low-cost autonomous submarines as threatening US carrier groups and drone swarms as saturating advanced air defences, i.e. capabilities that "will not only take advantage of swarm systems' low cost and overwhelming scale, but also be superior to swarms consisting of normal drones."[47] This suggests Chinese confidence in quantity-over-quality approaches enabled by AI.

Simultaneously, Chinese sources acknowledge weaknesses. The emphasis on "talent teams" and "innovative cyber talent ecosystems" implies concerns about human capital constraints relative to the United States.[48] The focus on civil-military fusion and leveraging its massive amount of data for machine learning ("China has control of more than 20 percent of strategic resources in the information era"[49]) suggests awareness that data advantages may compensate for other deficiencies. Western assessments often emphasize Chinese weaknesses in semiconductor fabrication and foundational algorithms, but Chinese documents may reveal different perceived vulnerabilities, perhaps in AI safety or system integration.

*H4: Chinese assessments of relative military AI capabilities exhibit confidence in cost-asymmetric approaches (swarms, autonomous systems), concerns about talent and foundational technology*

39 For this, see: US Department of State, "Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy," Bureau of Arms Control and Proliferation, https://www.state.gov/bureau-of-arms-control-deterrence-and-stability/political-declaration-on-responsible-military-use-of-artificial-intelligence-and-autonomy.
40 See: Victoria Waldersee, "US, China Opt Out of Joint Declaration on AI Use in Military," *Reuters*, 6 February 2026, https://www.reuters.com/business/aerospace-defense/us-china-opt-out-joint-declaration-ai-use-military-2026-02-05/.
41 Talha Burki, 'Crossing the Frontier: The First Global AI Safety Summit', *The Lancet Digital Health* 6, no. 2 (2024): e91–92, https://doi.org/10.1016/S2589-7500(24)00001-3.
42 Oliver Guest and Kevin Wei, 'Bridging the Artificial Intelligence Governance Gap: The United States' and China's Divergent Approaches to Governing General-Purpose Artificial Intelligence', version 1, preprint, arXiv, 2025, https://doi.org/10.48550/ARXIV.2506.03497.
43 It should be noted that the same critique can be made of the definition of autonomous weapons employed by the United Kingdom (though not the United States), see Paul Scharre, *Army of None: Autonomous Weapons and the Future of War*, New York City: W.W. Norton & Company, 2018, 110.
44 Saad Siddiqui et al., 'Promising Topics for U.S.-China Dialogues on AI Risks and Governance', version 1, preprint, arXiv, 2025, https://doi.org/10.48550/ARXIV.2505.07468.
45 Michael Horowitz, *The Diffusion of Military Power: Causes and Consequences for International Politics*, Princeton: Princeton University Press, 2010.
46 Allen, *Understanding China's AI Strategy: Clues to Chinese Strategic Thinking on Artificial Intelligence and National Security*.
47 Zhang, 'China's Military Employment of Artificial Intelligence and Its Security Implications'.
48 Cheung, 'AI Frontlines: China's PLA Evolution vs. NATO's Counterplay'.
49 Zhang, 'China's Military Employment of Artificial Intelligence and Its Security Implications'.

*gaps or emphasis on data and manufacturing scale as compensating advantages, observable through topic prevalence in research publications and procurement priorities.*

5. **Technology denial and economic coercion**: China controls dominant shares of rare earth element production and increasingly leads in battery technology and open-source AI models, i.e. inputs relevant to autonomous systems.[50] One key question with respect to the coming years is therefore: Does China envision technology export controls as a decisive military AI strategy? US semiconductor export restrictions targeting China demonstrate precedent for such weaponization. Chinese policy documents on indigenous innovation and "self-reliance" in critical technologies may reflect both defensive responses to Western restrictions and offensive intentions to create dependencies. However, evidence of an explicit Chinese strategy to "choke off" whole industries or AI capabilities through export bans remains so far limited.

Another explanation for the recent rare earth and magnet restrictions that threaten US and European defence supply chains[51] would be that they are largely defensive. In that case, they would be used as bargaining chips to gain more time, which would allow China to build domestic resilience against Western technology denial through import substitution and alternative supply chains. Since mid-2023, Beijing has progressively tightened controls on critical inputs, beginning with export licence requirements for gallium and germanium (in force from 1 August 2023), extending to seven rare earth elements and permanent magnets under a case-by-case licensing regime in April 2025 and, at the end of 2025, adding further rare earths (namely Holmium, Erbium, Thulium, Europium and Ytterbium), processing equipment and extraterritorial controls on foreign-made magnets containing even trace amounts of Chinese-origin material.[52] Our analysis of policy documents can reveal whether technology control appears as an explicit strategic objective or primarily constitutes a by-product of industrial policy.

*H5: China's technology denial strategy in military AI is primarily defensive rather than offensive, observable in policy emphasis on import substitution and self-reliance rather than export restriction, research investments in domestic alternatives to Western-controlled technologies and the absence or limitation of discussions of leveraging supply chain dominance for military advantage.*

6. **Offense-defence balance in AI-enabled warfare**: Technological changes have continuously changed how easy or difficult it is to attack other countries and actors as well as to defend against such incursions. Warfare can transform from stalemate to fast movement and victory within short periods of time, as happened between the end of the First World War and the beginning of the Second World War. Does China believe that AI will improve the chances of attackers or defenders in the future? The offense-defence balance literature examines how technology shifts advantages between attackers and defenders, with implications for crisis stability.[53] China's military modernization has emphasized anti-access/area-denial (A2/AD) capabilities consistent with defensive objectives.[54] Yet AI-enabled capabilities, particularly autonomous swarms and AI-assisted targeting, may restore offensive advantages through speed, saturation and surprise.

Chinese military writings on compressing OODA (observe-orient-decide-act) loops emphasize offensive potential: "Situational awareness capabilities strengthened by AI will not only accelerate the traditional OODA loop but also make a 'storm of steel' fired by autonomous vehicles and intelligent weapons possible."[55] This language suggests Chinese perceptions that AI restores offensive momentum after a period of defensive advantage. However, there is also a simultaneous emphasis on air defence, coastal security and resilience. In short, there is a mixed assessment based on the literature.

*H6: China's AI warfare vision encompasses both offensive (swarm saturation, compressed targeting timelines) and defensive (enhanced ISR, layered air defence) elements, with temporal evolution from defensive priorities (2000-2015) towards offensive emphasis (2016-2025), observable in research topic shifts and procurement patterns.*

7. **Human control and autonomous weapons**: The principle of meaningful human control over lethal autonomous weapons has become central to international debates, yet its definition

50 For China's open source lead, see Dashveenjit Kaur, 'Exclusive: Why are Chinese AI models dominating open-source as Western labs step back?', AI News, 9 February 2026.

51 See: Gracelyn Baskaran, "China's New Rare Earth and Magnet Restrictions Threaten US Defense Supply Chains," CSIS (Center for Strategic and International Studies), 9 October 2025, https://www.csis.org/analysis/chinas-new-rare-earth-and-magnet-restrictions-threaten-us-defense-supply-chains.

52 See, e.g.: Resilinc, https://resilinc.ai/blog/germanium-and-gallium-china-export-restrictions/; Keith Bradsher, 'Step by Step, How China Seized Control of Critical Minerals,' *The New York Times*, October 2025, https://www.nytimes.com/2025/10/27/business/china-rare-earth-export-controls.html.

53 Robert Jervis, 'Cooperation under the Security Dilemma', *World Politics* 30, no. 2 (1978): 167–214, https://doi.org/10.2307/2009958; Keir A. Lieber, 'Grasping the Technological Peace: The Offense-Defense Balance and International Security', *International Security* 25, no. 1 (2000): 71–104.

54 Oriana Mastro, 'China's Antiaccess-Area Denial (A2/AD) Capabilities: Is the U.S. Rebalancing Enough?', in *American Strategy and Purpose: Reflections on Foreign Policy and National Security in an Era of Change*, ed. William H. III Natter and Jason Brooks (CENSA, 2014).

55 Zhang, 'China's Military Employment of Artificial Intelligence and Its Security Implications'.

remains contested. A widely used classification, formalized in the deliberations of the CCW Group of Governmental Experts (GGE) on LAWS, distinguishes three categories of human control over autonomous weapons systems:[56]

> Human-in-the-loop (HITL): A human operator must positively authorize each specific action (e.g. select a target and initiate an attack). The system cannot engage a target without positive human authorisation.

> Human-on-the-loop (HOTL): The system autonomously selects and engages targets but operates under the supervision of a human operator who can intervene and abort the attack if necessary (veto power).

> Human-out-of-the-loop (HOOTL): The system selects and engages targets without any human intervention or possibility of override once activated.

How do Chinese military writings engage with the question of human primacy in warfare? One analysis notes that "Chinese discussions about keeping a human in the loop in technical writings remain limited to non-existent," while simultaneously observing that "the PLA's way of fighting is rooted deeply in a strong strategic culture arguing that human beings are the ultimate determination of war."[57] This tension between cultural norms and technological imperatives creates doctrinal ambiguity. The resolution may lie in how implicitly or explicitly risks are acknowledged when discussing human control. If texts emphasize HITL/HOTL mechanisms, they may be signalling an awareness of AI reliability risks or escalation dangers. Conversely, if texts celebrate "full autonomy" or "unmanned combat" with consistently high sentiment and without reference to human oversight, it suggests a dismissal of these risks in favour of operational efficacy. Application-specific variation is likely: Full autonomy may be deemed acceptable for certain defensive systems (e.g., counter-swarm), while human control is retained for strategic systems, in line with recent US-China understandings on nuclear AI. Alternatively, China may interpret "human control" as high-level political oversight rather than tactical intervention, allowing it to claim compliance with norms while pushing the envelope on tactical autonomy.

*H7: If Chinese discourse emphasizes "human-in-the-loop" mechanisms, it implicitly acknowledges AI reliability risks and escalation dangers, reflected in more tempered or negative sentiment scores for fully autonomous applications. Conversely, if autonomy is discussed without reference to human control mechanisms and with consistently high sentiment, it indicates prioritizing operational speed over risk mitigation, diverging from Western meaningful human control frameworks.*

Overall, these hypotheses provide the analytical framework for our empirical investigation. Each hypothesis generates observable implications testable through our four datasets: regulatory documents revealing official positions, research publications showing technical priorities, procurement contracts demonstrating actual investments and policy communications indicating doctrinal evolution. The following sections describe these data sources and the NLP methods we employ to extract systematic evidence on China's intelligentized warfare vision.

56 CCW Group of Governmental Experts on LAWS, Guiding Principles Affirmed by the GGE, Annex III to CCW/MSP/2019/9 (13 December 2019). This goes back to a 2012 Human Rights Watch report by Bonnie Docherty (*Losing Humanity: The Case Against Killer Robots*), as summarized here: https://www.armyupress.army.mil/Journals/Military-Review/English-Edition-Archives/May-June-2017/Pros-and-Cons-of-Autonomous-Weapons-Systems/.

57 Zhang, 'China's Military Employment of Artificial Intelligence and Its Security Implications'.

# 4

## Text data on Chinese approaches to AI

In order to understand China's approach to AI warfare, we systematically analyse how stated intentions manifest as observable institutional behaviour across four domains: regulatory frameworks, academic research, news communication and military procurement.[58] Together, these sources enable triangulation by comparing what the Chinese authorities signal directly (regulations) and indirectly (propaganda in state media), what researchers investigate (publications) and what the People's Liberation Army actually purchases (contracts). All four datasets were harmonized to create a unified corpus of over 171,000 documents spanning the period from 2000 to 2025. Figure 1 illustrates the temporal distribution across the datasets. Further technical details on provenance, pre-processing and merging are provided in Appendix A.

**AGORA AI governance documents**: The AI Governance and Regulatory Archive (AGORA), maintained by Georgetown University's CSET, provides our primary insight into China's formal regulatory frameworks. From the full database of over 950 AI governance documents, we identified 26 Chinese texts with AI relevance, five of which have explicit military AI content: These are China's 2016 CCW position paper on lethal autonomous weapons, the 2017 New Generation AI Development Plan, the 2021 Position Paper on Regulating Military Applications of AI, the 2023 Global Security Initiative Concept Paper and Ambassador Zhang Jun's 2023 remarks to

**Joint corpus coverage by year and dataset**

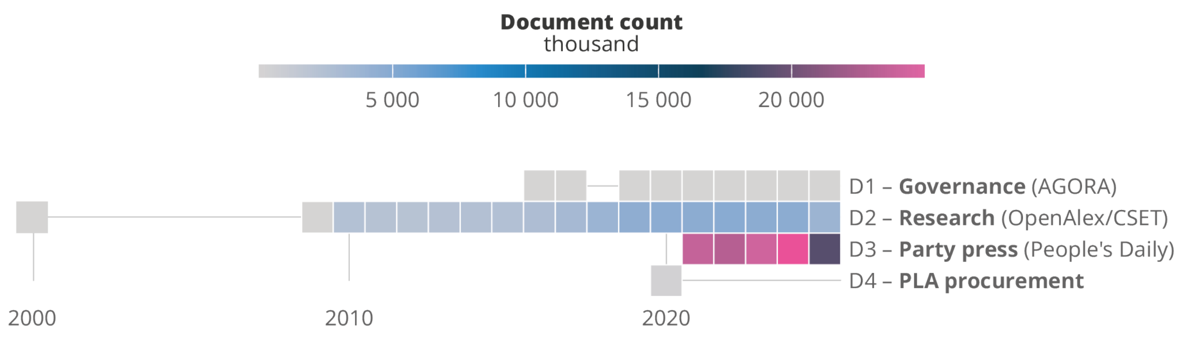


Figure 1: Joint corpus coverage by year and dataset.
*Source: Authors' generation.*

58 Policy documents and leadership speeches would have provided another lens on China's AI warfare vision, revealing how official discourse frames intelligentized warfare priorities. For instance, a recent analysis leveraged the Chinese Adaptive Policy Communication Corpus (CAPC-CG), a collection of 3.3 million paragraphs from Chinese laws, regulations, ministerial rules and high-level policy statements spanning 1949 to 2023. Bolun Sun et al., 'Creation of the Chinese Adaptive Policy Communication Corpus', arXiv:2510.08986, preprint, arXiv, 10 October 2025, https://doi.org/10.48550/arXiv.2510.08986. Unfortunately, at the time of writing, this dataset was not yet publicly available.

the UN Security Council. Preliminary term-frequency analysis shows a strong focus on ethics, security and data governance, while military domains are primarily framed through broader dual-use and security categories.

**CSET OpenAlex scholarly literature**: This dataset comprises 55,943 Chinese AI-related research papers, identified via CSET's Emerging Technology Overlay for OpenAlex and covering the period from 2000 to 2025. Computer vision dominates the subfield distribution (49.4%), followed by robotics (12.5%) and natural language processing (NLP) (6.2%). AI safety accounts for under 2%, suggesting that China's research portfolio focuses heavily on operationally relevant perception and control tasks. A keyword-based heuristic identified 1,407 papers (2.5%) as being relevant to military AI, covering topics such as UAV swarms, target detection and autonomous path planning. The number of military AI publications grew steadily from 52 papers in 2016 to 181 in 2024.

***People's Daily* news corpus**: Official state media provides the clearest window into China's deliberate rhetorical choices. Unlike pluralistic media environments, China's centralized Publicity Department issues directives that produce highly coordinated messaging. Recent analysis shows that, between 2012 and 2022, party newspapers published scripted propaganda under central directive on 90% of days.[59] Our dataset comprises 114,879 *People's Daily* articles spanning the time from January 2021 to September 2025. Filtering for AI- and military-related keywords reveals that 71,752 articles (62.5%) reference relevant terms, reflecting the degree to which AI and security concepts have become mainstream in broader development rhetoric.

**PLA procurement contracts**: Finally, procurement data grounds the analysis in observable military commitment. This dataset comprises 343 AI-related equipment contracts awarded by the PLA and state-owned defence enterprises between April and November 2020 and was compiled by CSET. Keyword analysis of the titles of these contracts reveals three dominant clusters: command and control and decision support systems (220 contracts); AI/ML software and analytics (146 contracts); and autonomous or unmanned systems (136 contracts).[60] The largest contract, valued at ¥21.13 million, is for a UAV data access and management system. This exemplifies the dataset's central finding: As early as 2020, the PLA prioritized not only autonomous platforms but also the data and ecosystems that enable their operational use.

59 Waight et al., 'The Decade-Long Growth of Government-Authored News Media in China under Xi Jinping'.
60 The co-occurrence with command and control and decision support systems suggests a similarity with NATO's Political-Military Assisted Decision Making programme, highlighted in NATO, 'Harnessing Artificial Intelligence: Allied Command Transformation at the Forefront of NATO Innovation', 16 April 2025, https://www.act.nato.int/article/harnessing-artificial-intelligence/.

5

# Methodological approach

It is impossible to decode China's AI warfare priorities from over 170,000 documents spanning regulatory texts, research papers, state media and procurement contracts in English and Mandarin Chinese through manual reading alone. We therefore employ computational text analysis, specifically a Natural Language Processing (NLP) tool pipeline that can systematically read, classify and compare large volumes of text at a speed and scale that no human research team could match. Readers unfamiliar with these methods will find a non-technical introduction in Appendix B, where we describe the key methods, how they work and how the field has developed over the past few years. Here, we explain what our own NLP pipeline does and the rationale behind each analytical choice.

**Choosing the right tools**: At the core of our analysis are Small Language Models (SLMs), the same class of AI systems that underlie tools such as ChatGPT.[61] These models can read and interpret text across languages, including Mandarin Chinese. This multilingual capability is essential, as most of our corpus is in Chinese. Waiting for human translations would mean analysing yesterday's debates in a rapidly evolving field. However, not all language models are equally suitable for analysing Chinese military strategy. Chinese-developed models such as DeepSeek and Alibaba's Qwen family perform well on Mandarin text, but they have been shown to systematically refuse or distort responses on politically sensitive topics. Recent academic research has documented refusal rates of up to 60% on prompts touching on subjects that the Chinese government deems sensitive.[62] This censorship is attributed to both direct government regulation, specifically the 2023 "Interim Measures for the Management of Generative Artificial Intelligence Services,"[63] which mandate algorithm registration and content filtering aligned with "socialist core values" and indirect influence through Chinese training data. Using such a model to analyse Chinese military AI discourse would be akin to asking a government spokesperson to independently evaluate their own government's strategy. We therefore exclusively use Western-developed, openly available models – specifically Mistral AI's Mistral-3B[64]

61 Smaller models refer to architectures with fewer parameters, typically in the 7-13 billion range rather than 70+ billion, which reduce memory footprint and accelerate inference on well-defined classification tasks. Fali Wang et al., 'A Comprehensive Survey of Small Language Models in the Era of Large Language Models: Techniques, Enhancements, Applications, Collaboration with LLMs, and Trustworthiness', version 1, preprint, arXiv, 2024, https://doi.org/10.48550/ARXIV.2411.03350.question answering, and reasoning, facilitating various tasks and domains. Despite their proficiency in various tasks, LLMs like LaPM 540B and Llama-3.1 405B face limitations due to large parameter sizes and computational demands, often requiring cloud API use which raises privacy concerns, limits real-time applications on edge devices, and increases fine-tuning costs. Additionally, LLMs often underperform in specialized domains such as healthcare and law due to insufficient domain-specific knowledge, necessitating specialized models. Therefore, Small Language Models (SLMs Distilled models are created through knowledge distillation, a training process where a compact "student" model learns to approximate the outputs of a larger "teacher" model, compressing learned representations while preserving much of the original capability. Victor Sanh et al., 'DistilBERT, a Distilled Version of BERT: Smaller, Faster, Cheaper and Lighter', version 4, preprint, arXiv, 2019, https://doi.org/10.48550/ARXIV.1910.01108; Dan Busbridge et al., 'Distillation Scaling Laws', version 1, preprint, arXiv, 2025, https://doi.org/10.48550/ARXIV.2502.08606. Both strategies allow researchers to process hundreds of thousands of documents locally. Peter Belcak et al., 'Small Language Models Are the Future of Agentic AI', version 1, preprint, arXiv, 2025, https://doi.org/10.48550/ARXIV.2506.02153.

62 Jennifer Pan and Xu Xu, 'Government-Imposed Censorship in Large Language Models', Stanford University and Princeton University, 2025, https://xu-xu.net/xuxu/llmcensorship.pdf.

63 See, for an English translation: https://www.chinalawtranslate.com/en/generative-ai-interim/#gsc.tab=0.

64 See: https://docs.mistral.ai/models/ministral-3-8b-25-12.

and Meta's Llama 3.1[65] – which, while not entirely unbiased,[66] do not systematically suppress discussion of military and strategic topics central to our analysis.[67] Both models are run locally on our own hardware rather than via commercial cloud services. This eliminates per-query costs, protects data and ensures that results can be independently reproduced by others.

Our pipeline comprises four sequential analytical stages, each of which is designed to address a specific question concerning China's military AI discourse.

**Stage 1: What words appear, and how often?** Before deploying any AI tools, we establish a descriptive baseline by tracking the frequency of key terms (both in English and their Mandarin equivalent) across all four datasets over time. This can be thought of as a systematic word count: We determine how frequently terms such as "machine learning," "surveillance," "autonomous weapons" and "command systems" appear each year, and whether terms relating to AI technology and military applications increasingly appear together in the same documents. This stage requires no AI inference and produces the temporal frequency charts presented in Section 6. It also generates a focused sub-corpus of the roughly 14,000 documents that mention both AI and military concepts, which feeds the more intensive analyses in Stages 2 and 3.

**Stage 2: Is the tone optimistic or cautious?** This stage examines the emotional tone of Chinese state media coverage of military AI. Using a Chinese-language sentiment tool called SnowNLP, each *People's Daily* article is given a score between 0 (most negative) and 1 (most positive).[68] This enables us to determine whether China frames military AI as an exciting opportunity, a necessary precaution or a potential risk, and whether this framing has shifted over time. The tool is simple yet efficient,[69] enabling us to analyse over 100,000 articles and track changes in the official tone. Applied to all Chinese-language texts in our news corpus, this stage produces sentiment scores that reveal whether state media frame military AI development in primarily optimistic (technological opportunity) or cautionary (risk management) terms and whether sentiment varies systematically across time periods.

**Stage 3: Which specific technologies and applications are discussed?** This stage is the most technically intensive. For each document in our focused sub-corpus, we ask the SLM to identify which AI technologies (e.g. deep learning, computer vision and natural language processing) and military applications (e.g. surveillance, targeting, autonomous weapons and command systems) are mentioned.[70] Rather than simply asking whether a document is "about military AI", i.e. a binary yes/no, our approach extracts granular information about which capabilities are being coupled with which military functions. The result is a detailed map of China's military AI priorities as expressed across institutional domains. Details on inference are given in Appendix C, and the full prompting scheme is detailed in Appendix D.

**Stage 4: Do Chinese texts support or contradict our hypotheses?** In this final stage, LLooM (Large Language Model-assisted Opinion Mining), which was developed at Stanford University, is used to systematically test our seven theoretical hypotheses against the corpus.[71] LLooM is a concept induction framework that combines automated text analysis with human-defined criteria.[72] Unlike purely exploratory topic modelling,[73] LLooM allows researchers to operationalize theoretical constructs as explicit concept prompts and score every document against precise inclusion criteria. This process transforms the seven theoretical propositions presented in Section 3 into measurable patterns derived from the texts themselves rather than from our prior assumptions. The output comprises quantitative prevalence scores indicating the proportion of documents that strongly

65 See: https://huggingface.co/meta-llama/Llama-3.1-8B. The distilled version "llama3.1:8b-instruct-q4_K_M" can be found here: https://ollama.com/library/llama3.1:8b-instruct-q4_K_M.

66 Reinforcement learning from human feedback likely reflects Western annotator preferences and content moderation norms. For these biases, see: Shannon Lodoen and Alexi Orchard, 'Ethics and Persuasion in Reinforcement Learning from Human Feedback: A Procedural Rhetorical Approach', version 1, preprint, arXiv, 2025, https://doi.org/10.48550/ARXIV.2505.09576.

67 For document classification tasks, we use Mistral AI's novel Ministral-3:8b model, which demonstrates superior performance on Chinese text categorization compared to alternatives in the same parameter class. For discourse interpretation and hypothesis evaluation, we deploy Meta's Llama 3.1 8B Instruct model. This model produces reliably structured JSON output, which is essential for LLooM's computational pipeline (see below), while offering strong multilingual performance on Mandarin text. Both models were developed outside the direct jurisdiction of authoritarian content controls.

68 See: https://github.com/isnowfy/snownlp.

69 This approach proves comparable in performance to deep learning sentiment models for determining overall valence, with the crucial advantage of computational speed suitable for large corpora. Jennifer Pan et al., 'Disguised Repression: Targeting Opponents with Non-Political Crimes to Undermine Dissent', *SSRN Electronic Journal*, ahead of print, 2023, https://doi.org/10.2139/ssrn.4488481.

70 Using the Ministral-3:8b model deployed locally via Ollama, we conduct batch inference on a stratified sample of our corpus documents that contain both AI and military terminology. Each document receives a structured prompt requesting four specific outputs: (1) whether the text mentions AI/ML technologies; (2) if affirmative, a list of specific AI technologies mentioned (e.g., machine learning, neural networks, computer vision, natural language processing); whether the text refers to military or defence applications; and (4) if affirmative, a list of specific military applications (e.g., surveillance, targeting, autonomous systems, command systems, autonomous weapons).

71 Michelle S. Lam et al., 'Concept Induction: Analysing Unstructured Text with High-Level Concepts Using LLooM', *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA), CHI '24, 2024, https://doi.org/10.1145/3613904.3642830.

72 We deployed LLooM with a locally hosted Llama 3.1 8B model with 4-bit quantization for all text processing operations, including distillation, synthesis and scoring, replacing the default OpenAI API. We used the all-MiniLM-L6-v2 sentence transformer model for embedding generation, as recommended by the LLooM developers. The all-MiniLM-L6-v2 model has 22 million parameters compared to the 8 billion of Llama 3.1 and is optimized for semantic similarity tasks. It produces efficient multilingual embeddings that are suitable for clustering Mandarin-language documents. For technical details, see: https://stanfordhci.github.io/lloom/about/custom-models.html.

73 David M. Blei et al., 'Latent Dirichlet Allocation', *The Journal of Machine Learning Research* 3 (2003): 993–1022.

match each hypothesis, as well as qualitative concept summaries that synthesize how each hypothesis manifests in Chinese discourse. In practical terms, LLooM achieves in hours what would take a team of human coders weeks to complete manually while producing auditable and reproducible results.

We restrict the LLooM analysis to the 26 AGORA governance documents rather than the full joint corpus. This decision is based on two factors. Firstly, applying the full LLooM pipeline to 171,149 documents would be computationally prohibitive as the workflow requires a large number of sequential LLM calls at each stage. Secondly, and more importantly, exploratory testing revealed that merging corpora of different types into a single LLooM run results in inconsistent seed selection. The model recognizes the different documentary contexts and fails to generate overarching hypotheses that adequately capture the distinct characteristics of each dataset. Restricting each run to a single, internally consistent corpus resolves this problem. Future work employing LLooM across heterogeneous document collections should account for this limitation.

As the exploratory concepts identified by LLooM primarily reflect general governance frameworks, while our seven hypotheses target more specific strategic questions, a second phase was necessary. We operationalized each hypothesis as a custom concept prompt with explicit decision criteria.[74] We designed our prompts in line with established best practices for LLM-based text classification, assigning the model an explicit evaluator role, providing binary decision criteria with concrete examples and structuring the prompts to minimize ambiguity while maintaining sensitivity to domain-specific language. For instance, the prompt for Hypothesis 1 (H1) marks a document as positive if it explicitly links AI, automation or unmanned systems to military manpower issues such as demographic decline, recruitment shortfalls, reducing casualties or optimizing force structure. It marks it as negative if AI is mentioned only in generic modernization terms or in civilian labour contexts without any connection to military force structure. Complete prompt specifications are provided in Appendix E. LLooM then scored each document against all seven hypothesis concepts using a five-point scale (0, 0.25, 0.5, 0.75, 1.0), where scores above 0.75 indicate strong evidential support for the hypothesis.

A practical note on safety guardrails: During configuration, the models repeatedly refused to process queries framed around "AI-enabled warfare", citing built-in safety restrictions. We resolved this issue by reframing the instructions to focus on "long-term visions for AI" rather than explicit warfare terminology. This approach enabled us to proceed with the analysis without compromising its substantive focus. Our experience highlights a broader challenge for computational security studies: The same safety measures that make LLMs reliable consumer tools can impede legitimate academic research. Future methodological guidelines for the field must therefore explicitly address this tension.

74 In technical terms, we configure LLooM with hypothesis-derived seed concepts corresponding to each of our seven theoretical dimensions: demographic force reduction (H1), nuclear stability (H2), international law strategy (H3), asymmetric capabilities (H4), technology denial (H5), offensive momentum (H6) and human control ambiguity (H7). These seeds guide the distillation phase, biasing concept discovery towards theoretically relevant themes while allowing the model to refine definitions based on actual corpus language.

# 6

# Empirical results

## 6.1. Stage 1: Keyword frequencies

Before applying computationally intensive, SLM-based classification, we establish a descriptive, empirical baseline by tracking the frequency with which key AI and military terms appear across our four datasets over time. This first stage serves two purposes: It provides a transparent, accessible overview of the main themes in China's AI-related military discourse, and it identifies documents with sufficient content to warrant further analysis in subsequent stages.

The core instrument for this stage is a bilingual keyword dictionary, i.e. a manually **curated list of Chinese terms related to AI technologies, military applications and defence governance**. Rather than attempting to cover every technical variant exhaustively, the dictionary is designed for so-called high recall: It captures any document that meaningfully engages with AI in a military or security context, even if the terminology varies across institutional registers. It includes foundational AI terms such as 人工智能 ("artificial intelligence"), 深度学习 ("deep learning") and 智能 ("intelligent/smart") as well as military and security terms such as 军事 ("military"), 国防 ("national defence") and 作战 ("combat operations") and application-specific terms such as 无人机 ("drone/UAV"), 侦察 ("reconnaissance") and 目标识别 ("target recognition"). To enable meaningful comparisons to be made across datasets of vastly different sizes, we normalize term counts by total document volume within each dataset. This ensures that a term appearing fre-

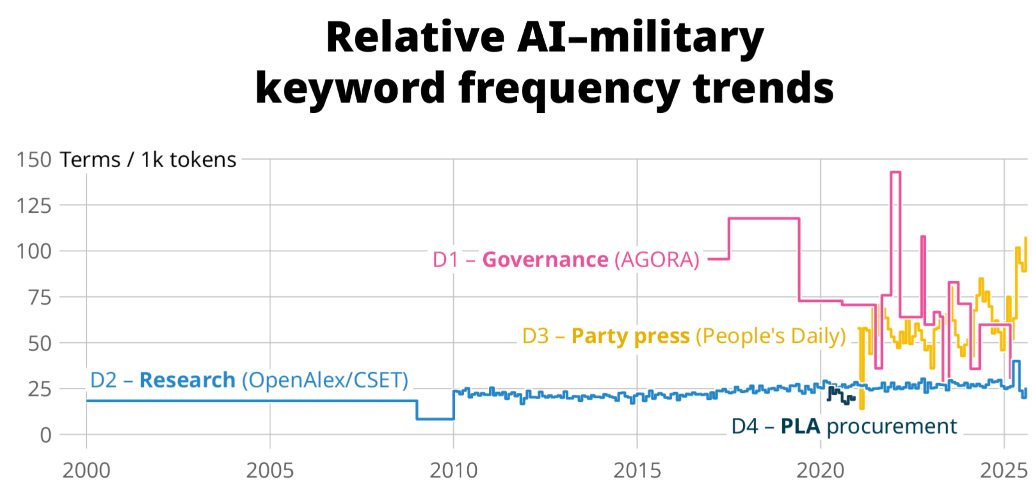


Figure 2: Relative AI–military keyword frequency trends

*Source: Authors' generation.*

quently in a small corpus of governance documents can be directly compared to the same term's footprint in the much larger news corpus.

The resulting time-series plots (Figure 2) show when and where the discourse on AI and the military intensified. In the *People's Daily* dataset, the proliferation of military and intelligentization language after 2021 stands out. Here, "AI" (人工智能), "intelligent" (智能) and "military" (军事) appear persistently week after week, frequently in the context of industry partnerships, high-profile defence exercises or efforts to signal China's technological modernity. Legal and governance texts exhibit a more punctuated pattern: Each spike often corresponds to the publication of a white paper or ministerial guidance, where terms like 管理 ("regulation" / "management") and 人工智能 ("AI") are coupled with policy shifts. Research papers are generally less security-oriented. Here, the most common drivers are terms like "AI," "machine learning" and "robot," used predominantly for technical contributions rather than signalling military posture or state intent. In contrast, PLA procurement data, while limited to 2020, is highly specific: Keywords such as "drone," "UAV" and "targeting" dominate, reflecting the operational priorities for autonomous systems, surveillance and battlefield decision support.

For validation, we also conduct a brief qualitative review of matched texts, which underscores these quantitative patterns. In the regulatory corpus, a rise in the AI-military dictionary typically links to concrete governance actions such as new labelling requirements for military AI or strategic chapters on defence readiness. In research papers, the term "robot" predominantly marks application domains (like manufacturing) unless cross-linked with "military." *People's Daily*, by contrast, mobilizes the dictionary's terms in routine coverage of drills and "AI-powered" upgrades to existing forces, positioning intelligentization as both a narrative of modernization and a deterrence signal. By contrast, procurement records embed our keywords directly within operational specifications, such as bids for patrol and reconnaissance drones, contracts for automated threat-detection systems and procurement notices for targeting modules and command platforms, with virtually no civilian or administrative applications. While the research corpus routinely uses terms such as "autonomous" or "surveillance" to describe civilian manufacturing or logistics applications, the procurement records reflect hardware and software that has been acquired explicitly for warfighting functions. This overview, and the peaks in Figure 2, gives a first glimpse into Chinese AI–military discourse and its key themes and contexts.

Before applying the more complex computational analyses described in the subsequent subsections, we narrow the full corpus of 171,149 documents to a focused working set by retaining only those containing at least one term from our AI-military dictionary. This step serves two purposes: It ensures that the analysis is based on substantively relevant material, and it excludes content, such as *People's Daily* articles on agricultural subsidies or cultural policy, that would otherwise introduce irrelevant information. Filtering reduced the corpus by roughly half (49.7%), to 86,038 documents. Retention rates differ by source type. Regulatory texts (AGORA) were retained in full (100%), as AI-military terminology is inherent to their subject matter; likewise, PLA procurement records passed the filter at 90%. Research publications were retained at a rate of 83.1%, reflecting their mix of military and civilian AI applications. *People's Daily* showed the lowest retention rate at 34.1%, as it is the only general-purpose source in the corpus (it covers the full breadth of party-state communication). The final sub-corpus contains terminology-rich documents, with an average of 4.72 AI-military keyword mentions each (median: 3). The texts with the highest concentration of keywords are not newspaper articles, but legal instruments: the Chinese Military Facilities Protection Law contains 352 mentions of keywords, and regulatory documents on unmanned aerial vehicles contain 201 mentions. This confirms that the filter successfully captures the normative and operational core of Chinese AI-military discourse.

Of the hypotheses tested in this study, Hypothesis 7 – which predicts a decline in the emphasis placed on human control over autonomous systems in Chinese AI-military discourse – can be analysed using keyword frequency alone. This means that it does not require the SLM-based coding applied to the remaining hypotheses. We therefore close this section by systematically tracking human-control terminology across all *People's Daily* articles and 55,943 research papers. Among Chinese AI research publications, human-in-the-loop (HITL) terminology dominates, appearing in 139 documents (60.4% of human control references) as compared to 60 mentions of human-out-of-the-loop/full autonomy (26.1%) and 31 mentions of human-on-the-loop supervisory control (13.5%). This 2.3:1 ratio of HITL to full autonomy suggests that Chinese researchers, like their Western counterparts, recognize technical and safety imperatives for maintaining human involvement in autonomous systems, particularly for high-stakes applications. Temporal trends, depicted in Figure 3, show sustained HITL discussion from 2010 onwards, peaking in 2020. Notably, "meaningful human control" as a normative concept appears in only five research papers, indicating limited engagement with the international legal and ethical frameworks emerging from UN deliberations on lethal autonomous weapons systems (LAWS).

Interestingly, *People's Daily* exhibits the opposite pattern: Human-out-of-the-loop/full autonomy terminology (完全自主, 自主决策, 无需人工) appears in 540 documents (58.3%), decisively outnumbering HITL references (人在回路, 人工控制) at 373 documents (40.3%). Human-on-the-loop supervisory language is nearly absent (13 documents, 1.4%). This ratio favouring full autonomy over human involvement stands in sharp

## Human control terminology over time

Number of documents mentioning each concept, by corpus

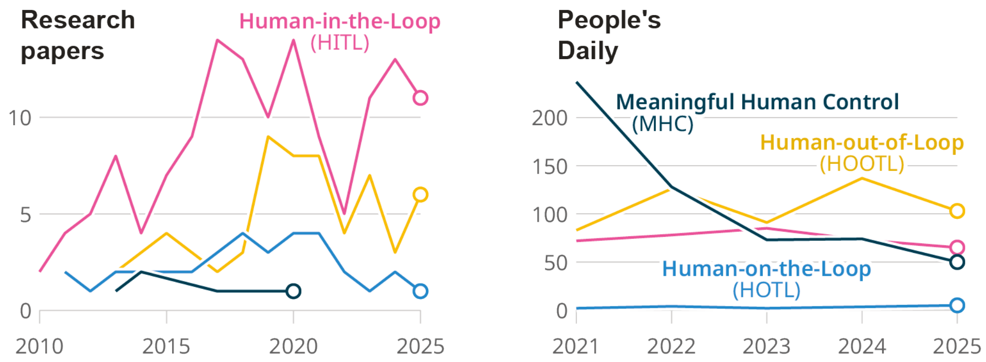


Figure 3: Human control terminology over time.
*Source: Authors' generation.*

contrast to the research corpus and suggests that state media actively promote autonomous systems as desirable and achievable, consistent with the propaganda strategy of projecting confidence that we will later identify in Section 6.1 using sentiment analysis. This difference might suggest that the scientific community in China is less confident than the official language about autonomous weapons. Even more revealing is the treatment of "meaningful human control" (人的控制, 有效控制, 人类主导): While researchers rarely invoke this concept (5 mentions), *People's Daily* uses it extensively (562 documents, 0.49% of corpus). However, temporal analysis exposes this as a declining priority. References peaked in 2021, coinciding with intensified international debates over autonomous weapons regulation following the 2021 UN Group of Governmental Experts (GGE) discussions, then fell sharply. This trajectory suggests that "meaningful human control" functioned primarily as diplomatic vocabulary during peak international scrutiny but has been progressively de-emphasized as China's autonomous systems capabilities matured and as Western regulatory momentum stalled.

From the data, we can also see that both corpora show critical shifts around 2020-2022. In research, HITL mentions peaked in 2020 and then declined, while full autonomy references surged in 2020. In news, meaningful human control peaked in 2021 before its precipitous decline, while full autonomy references grew between 2021 and 2024. This synchronized shift across both corpora during 2020-2022 coincides with the PLA's accelerated "intelligentization" drive, observed combat performance of autonomous systems during the Second Nagorno-Karabakh War (2020) and early Russo-Ukrainian War (2022) and diminishing prospects for binding international restrictions on autonomous weapons following repeated failures to advance LAWS negotiations at the UN. Two qualitative examples illustrate how Mandarin-language news and commentary discuss AI and unmanned military systems. In late 2025, the PLA's 77th Group Army conducted exercises integrating unmanned drones, vehicles and boats into a cohesive battle plan involving complex coordination. The headline was: "Drones, robots and unmanned surface vessels 'team up' for an attack."[75] Chinese media also recently reported that, over the last 15 years, the J-20 stealth fighter had evolved into a centrepiece of stealth technology, integrated with AWACS and EW aircraft as well as unmanned systems for system-level combat.[76] The idea was to create "a comprehensive detect–jam–strike–control" chain, with future development focusing on "AI empowerment, engine upgrades and enhanced avionics to maintain air superiority in complex near-real-combat environments."

What does this imply for our hypothesis 7? The keyword evidence strongly supports the argument that China expresses permissive language in technical publications and increasingly in state media, with discourse evolving towards reduced human involvement over time. However, the asymmetry between research (HITL-dominant) and news (autonomy-dominant) suggests a more nuanced picture than simple domain-dependence. This asymmetry may point to a structural divide in Chinese AI-military

75 See: Emergency Technology Observatory, https://scout.eto.tech/?id=5025.
76 See: Emergency Technology Observatory, https://scout.eto.tech/?id=5290.

communication. In technical publications, researchers adopt a cautious approach to human oversight, treating human-in-the-loop requirements as genuine engineering constraints (in line with the global academic discourse). By contrast, state media consistently project confidence in autonomous military capabilities, demonstrating strategic resolve to domestic and international audiences rather than reflecting technical assessments. Most significantly, keyword analysis reveals a growing divergence between China's international and domestic positions on autonomous weapons governance. "Meaningful human control"' – the normative cornerstone of UN CCW negotiations on lethal autonomous weapons systems, which requires humans to retain accountability for targeting decisions – virtually disappears from *People's Daily* after 2021, while references to full autonomy increase. While China continues to engage in LAWS discussions in Geneva, the domestic discourse has quietly shed the normative vocabulary on which those talks rest. This suggests a gradual rhetorical abandonment of international norms. China participates in the process while progressively decoupling its domestic messaging from the framework that the process assumes.

## 6.2. Stage 2: Sentiment analysis of state media discourse

Sentiment analysis of our *People's Daily* corpus reveals an extraordinarily optimistic and homogeneous affective framing of military AI discourse in Chinese state media. Using SnowNLP scores, where 0 is the most negative and 1 the most positive, the distribution exhibits a pronounced ceiling effect (a statistical phenomenon in which scores cluster at the upper boundary of the scale). The mean sentiment is 0.96 (SD=0.19), and three-quarters of all articles receive a perfect score (Figure 4). Less than one percent of articles falls below the neutral 0.5 threshold. This ceiling effect is consistent with People's Daily's propagandistic role: Military modernization and AI are framed as sources of national pride, technological achievement and strategic opportunity, with virtually no space for cautionary or critical narratives.

At the article level, we find only a negligible positive correlation between sentiment and AI-military keyword density (Pearson $r = 0.017$, $p < 0.001$),[77] where salience is measured as the number of AI-military terms per 1,000 tokens using the dictionary defined in the previ-

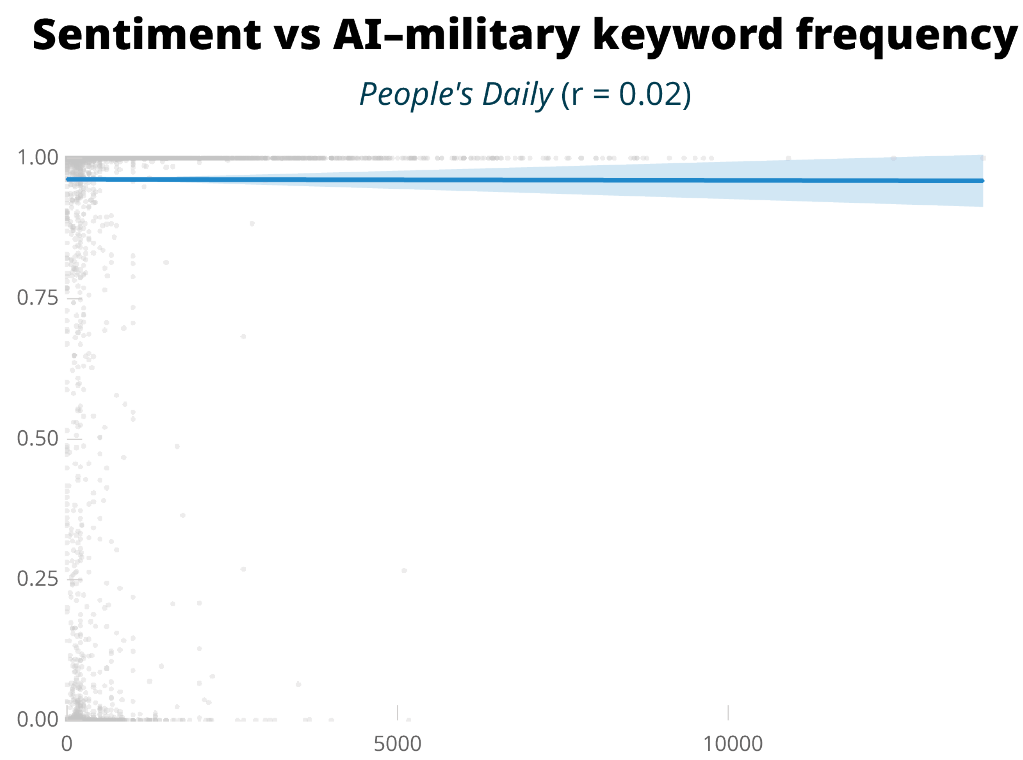


Figure 4: Sentiment vs AI–Military Keyword Frequency. 0=negative; 1=positive
*Source: Authors' generation.*

77 For explanations of this measure, see Glossary.

ous sub-section. Pan and Xu's analysis of China-based LLMs, mentioned earlier, demonstrates empirically that government censorship manifests as asymmetric sentiment divergence in LLM outputs: Topics that the regime seeks to denounce exhibit markedly lower sentiment in Chinese models when compared to non-Chinese models, while commended topics show no cross-model differences.[78] Applied to our news corpora, this logic would predict that articles with high military AI salience, which might potentially raise or indicate issues of vulnerability or AI risk, should be framed more negatively if censors regard such discussion as sensitive. Instead, our near-zero correlation indicates that **simply mentioning military AI more frequently does not trigger any systematic sentiment shift. State media maintain a uniformly positive tone regardless of how intensively AI or military topics are discussed**.

Lexical analysis of word-level sentiment associations reveals how this positivity is constructed. Among the most positively associated terms in our sub-corpus of AI-military-news are ideological expressions such as 恩格斯 (Engels), 卡尔 (Karl), 奋发有为 (striving for

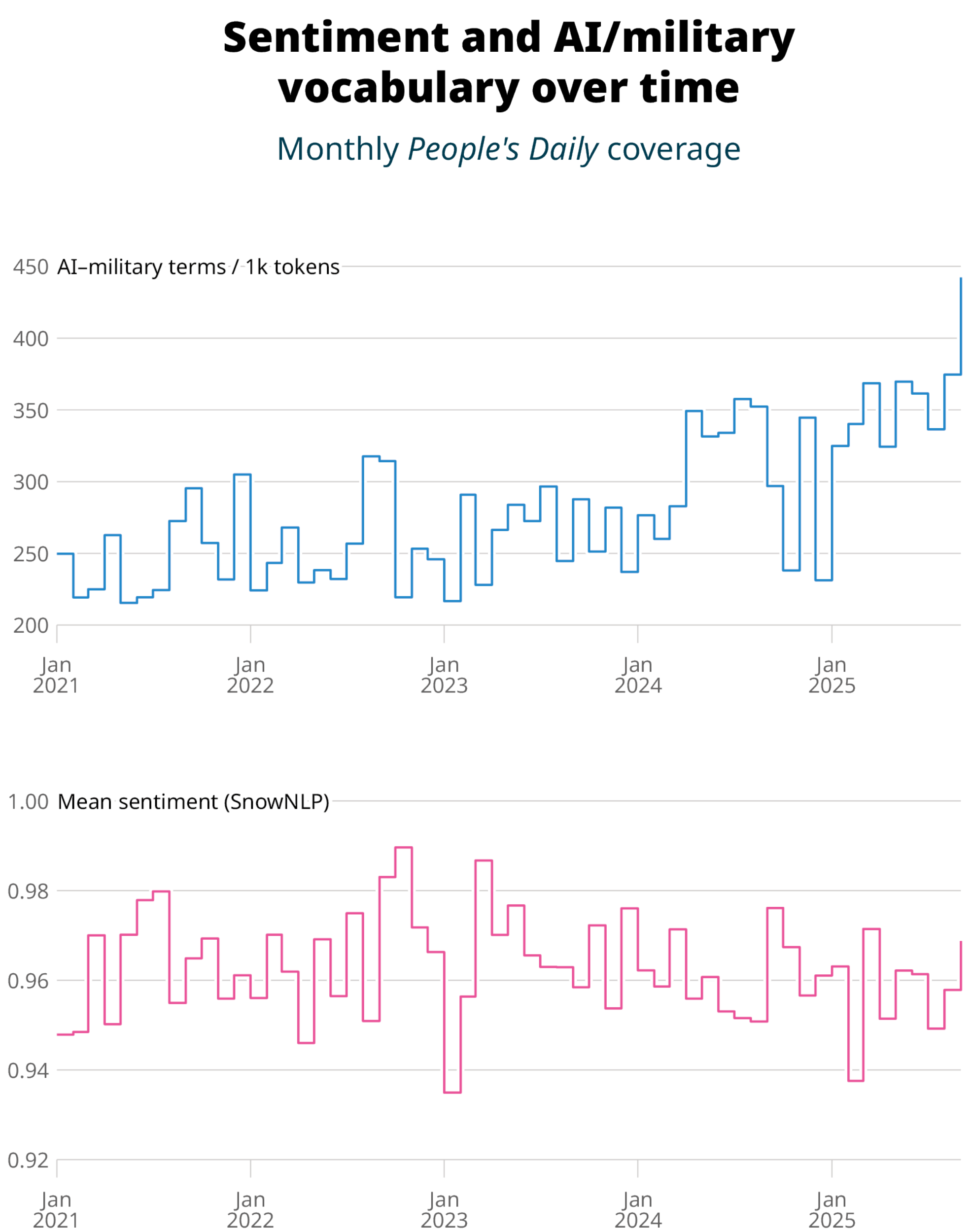


Figure 5: Sentiment vs AI–Military Keyword Frequency. 0=negative; 1=positive
*Source: Authors' generation.*

78 Pan and Xu, 'Government-Imposed Censorship in Large Language Models'.

achievement) and 南昌起义 (Nanchang Uprising) as well as institutional terms like 编辑部 (editorial office), 侨联 (Overseas Chinese Association) and 金融业 (financial industry). These tokens typically appear in formulaic success narratives and political paeans celebrating Party leadership, revolutionary heritage and China's rise, which helps push sentiment scores to the maximum in a large share of articles. By contrast, the most negatively associated words cluster in technocratic and problem-oriented domains rather than in discussions of AI risk as such. They include economic regulation terms such as 经营额 (business turnover), 非银行 (non-bank), 矿业权 (mining rights), 停业 (suspension of operations), 交割 (delivery/settlement, 0.345) and 洗钱 (money laundering). In addition, there is use of legal-bureaucratic terms like 复议 (administrative reconsideration) and a small set of security-related tokens such as 涉密 (classified) or 军工 (defence industry). These words tend to appear in reports on regulatory enforcement, workplace hazards, legal compliance and administrative complaints. Notably, even military-related negative terms appear in mundane bureaucratic contexts, such as defence industry licensing issues or classification violations rather than in strategic risk discussions. In line with this impression, the ten most negative articles in our sample focus on issues like financial irregularities, environmental violations, administrative procedures and legal liability. Conversely, the most positive articles are overwhelmingly celebratory profiles of national achievements, Party anniversaries, diplomacy and veterans' affairs.

The temporal analysis, however, uncovers a striking pattern that suggests a clear development (Figure 5), raising the question whether the discourse is solely driven by static propaganda or also affected by political or technological events. Across 57 months with at least ten articles each (January 2021 to September 2025), **AI-military term frequency more than triples, with especially sharp growth from late 2023 into 2025, consistent with the PLA's accelerating "intelligentization" drive**. Yet average sentiment remains remarkably stable in a narrow band near maximum positivity. This stability alone suggests **effective sentiment management across fluctuating coverage**.

More revealing, however, is the six-month rolling correlation analysis, which exposes a dramatic reversal in the relationship between sentiment and AI-military term density (Figure 6). From August through October 2024, the correlation was strongly negative (r = -0.74 to -0.75), indicating that months with higher AI-military term density actually exhibited slightly lower sentiment scores (though still near-maximal in absolute terms). This pattern flips entirely beginning in early 2025: By February 2025, the correlation approaches zero (r = -0.02), and from March onwards it becomes strongly positive (peaking at r = +0.90 in July 2025).[79] **The 2022 spike, which can be also seen in Figure 6, might be interpreted**

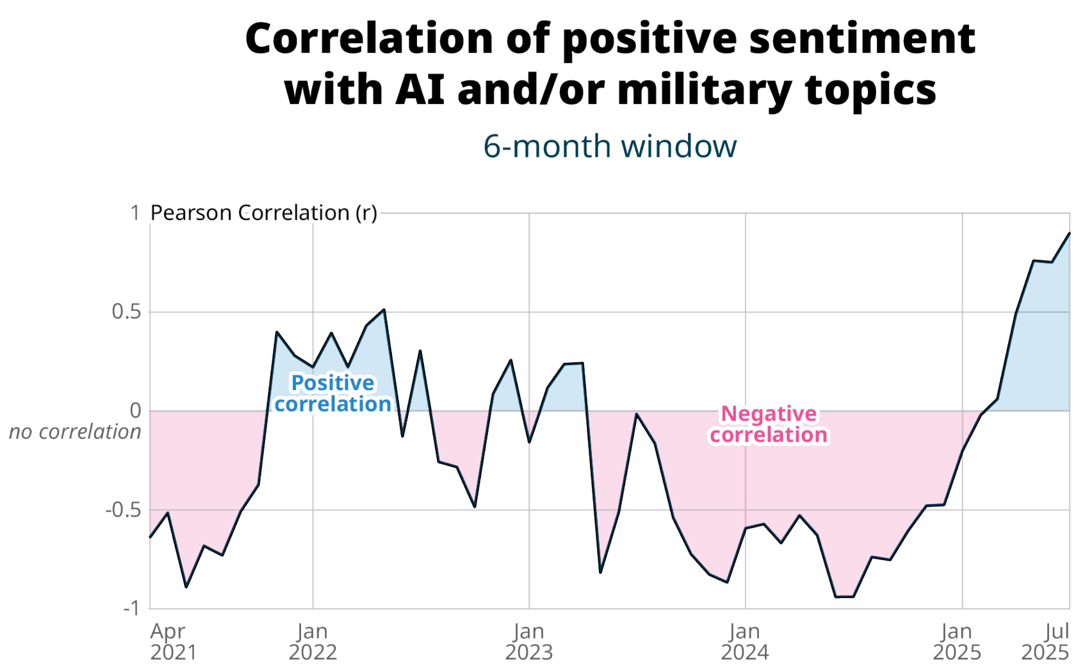


Figure 6: Correlation of positive sentiment with AI and/or military topics. 0=negative; 1=positive
*Source: Authors' generation.*

79 Quarterly correlation analysis reinforces this pattern, showing a clear temporal progression from weakly positive correlations in 2021-2022 through increasingly negative correlations in 2023-2024 to sharply positive correlations from Q1 2025 onwards.

**as a "proof of concept" for the propaganda strategy that became institutionalized in 2025**. When state media successfully paired high-salience AI-military coverage with amplified positive sentiment during the Beijing Winter Olympics (February 4-20, 2022) and Russia's full-scale invasion of Ukraine (starting on February 24, 2022), this likely demonstrated the effectiveness of coordinated sentiment-salience coupling for confidence-building and soft power projection. The reversion to negative correlation in mid-to-late 2022 and throughout 2023-2024 suggests this approach had not yet been adopted as systematic policy.

**The strong correlation reversal throughout 2025 coincides with several developments in China's communication and propaganda strategy and the PLA's AI trajectory**. First, Xi Jinping's multi-stakeholder "narrative innovation" framework, emphasizing coordinated messaging across government, media, business and social sectors, intensified throughout 2024-2025. As has been noted in June 2025: "Xi Jinping's new multi-stakeholder approach to what the CCP terms 'external propaganda' (宣) is described by communication scholars in China as a systematic evolution from traditional propaganda to coordinated 'narrative innovation' (叙事话语创新) across government, media, business and social sectors […] The approach requires the concerted effort of all."[80] Second, Chinese state media shifted towards more overt, systematic propagation of strategic narratives in 2025, particularly concerning military and technological achievements.[81] Third, the PLA's embrace of "dissipative warfare" and other AI-centric doctrines gained prominence in military strategic discourse during this period. This dissipative warfare doctrine can be summarized as a strategy that operates in wartime and peacetime. It involves maintaining order within one's own system while deliberately creating disorder within an adversary's political, economic and military networks. This doctrine dissolves the differentiation between the front and the home front, envisioning that the latter will become a prominent theatre in any future war.[82] The positive correlation in the recent months of 2025 suggests that rather than avoiding or downplaying AI-military content when it becomes more salient, state media now actively amplify positive sentiment precisely when AI-military topics receive intensive coverage.

For our research question, these findings carry three main implications. First, the uniformly positive sentiment surrounding *People's Daily* coverage of military AI indicates that Chinese authorities currently frame AI-enabled military modernization as an unambiguous strategic asset, not as a domain requiring public debate over risks, ethics or arms control. This stands in sharp contrast to Western policy debates, where military AI is frequently associated with concerns over autonomous weapons, escalation and algorithmic accountability. Second, the dramatic correlation reversal from mid-2024 to mid-2025 suggests an active recalibration of propaganda strategy coinciding with Xi's "narrative innovation" framework and the PLA's accelerated AI integration. Rather than maintaining stable sentiment regardless of topic salience (the pre-2025 pattern), state media now systematically couple high AI-military salience with intensified positive affect. This shift reflects growing Chinese confidence in their country's military-technological trajectory. Third, this combination of rising salience, persistently ceiling-level sentiment, and evolving sentiment-salience coupling indicates a sophisticated, multi-layered propaganda system. At the macro level, it enforces blanket optimism through lexical choices, formulaic framing and editorial control; at the micro level, it dynamically adjusts sentiment amplification strategies in response to coverage intensity and strategic priorities. This coordinated approach aims to cultivate domestic confidence in China's technological and military trajectory while pre-empting public deliberation about potential strategic trade-offs or AI risks.

## 6.3. Stage 3: LLM-based classification of military AI warfare implications

To move beyond keyword-based corpus filtering and systematically identify specific AI technologies and military applications within our dataset, we employed a structured classification pipeline using a Large Language Model. In particular, our pipeline relies on the Ministral 3 8B model as described in Section 5.2. This approach enabled us to extract granular information about which AI capabilities are being discussed in relation to which military domains, revealing the thematic architecture of China's military AI discourse and pointing to its future priorities. Due to the typically long inference times for LLM-based pipelines, we analysed a stratified sample of **20,000 documents** from our corpus. We processed the documents in a way that was computationally efficient while maintaining data integrity.[83]

80 See: https://chinamediaproject.org/2025/06/30/chinas-go-local-vision-of-global-messaging/.
81 For the example of Chinese-Japanese relationships, see: https://www.aspistrategist.org.au/normalising-disinformation-china-shifts-to-overt-operations-against-japan/.
82 K. Tristan Tang, 'Dissipative Warfare: The PLA's Potential New Strategy in the AI Era', China Brief Notes, September 2025, https://jamestown.org/dissipative-warfare-the-plas-potential-new-strategy-in-the-ai-era/.
83 In particular, we implemented retry logic, timeout handling and batch checkpointing to manage inference failures and enable resumption after interruptions. As noted earlier, the use of Ministral 8B rather than larger models reflected a pragmatic trade-off: While frontier models (e.g., GPT-5, Claude) might achieve marginally higher accuracy on ambiguous cases, edge-optimized models enable local inference at scale without API costs or data transmission constraints.

The stratified sampling strategy balanced two competing objectives: ensuring adequate representation of smaller, specialized documents while maintaining sufficient statistical power for the dominant sources.[84] Rather than employing purely proportional sampling, which would have yielded fewer than 100 documents from the two smallest sources combined, we implemented a modified stratified approach that allocated equal quotas (5,000 documents per source) where possible, then capped allocation at the maximum available count for exhaustively sampled sources (procurement and governance). The remaining quota was filled from the research corpus, the largest available source. This procedure yielded a final sample composition of: research (73.5%, n=14,703), Party press (25.0%, n=5,000), procurement (1.4%, n=271) and governance (0.1%, n=26). While research articles dominate the sample numerically, the inclusion of all 297 procurement and governance documents ensures that rare but strategically significant document types, such as PLA equipment specifications, are fully represented in the classification analysis.

The LLM was then prompted with a structured JSON schema requiring four classification outputs for each document:

> **AI technology presence** (yes/no/unclear): Does the text mention AI/ML technologies?

> **AI specifics**: Are explicit AI technologies mentioned (e.g., machine learning, computer vision, neural networks)?

> **Military domain presence** (yes/no/unclear): Does the text discuss military/defence applications?

> **Military specifics**: Are explicit military applications mentioned (e.g., autonomous weapons, surveillance, targeting?)

This schema enforced consistent extraction while allowing the LLM to identify domain-specific terminology in both English and Chinese. For the full prompting scheme, see Appendix D. The pipeline achieved a success rate exceeding 99.99%, demonstrating the robustness of structured output generation with contemporary edge-optimized models. To ensure semantic consistency, we applied post-processing normalization that mapped synonyms and language variants. For example, references to "drone," "UAV" and "unmanned aerial vehicle (UAV)" were consolidated under a single term.

This LLM-based classification revealed that 75.1% (15,008 documents) contained explicit AI/ML technology references, while 6.9% (1,376 documents) discussed military or defence applications. Critically, only 484 documents (2.4% of the sample) addressed both AI technologies and military applications simultaneously. This relatively low intersection rate suggests that China's military AI discourse remains segmented: Technical AI research publications often lack explicit military framing, while military strategy documents may reference AI at a higher level of abstraction without detailing specific technologies. This pattern may reflect deliberate compartmentalization in Chinese research communication or, alternatively, indicate that military AI integration is still emerging as a distinct discursive domain.

Analysing the documents that were successfully classified reveals a clear hierarchy of AI technologies in Chinese military AI discourse (see Figure 7, left panel). Machine learning is the dominant term by far, appearing in 6,504 documents, followed by computer vision (4,717) and neural networks (2,900). This focus on data-driven learning paradigms, i.e. systems that improve through exposure to large datasets rather than explicit programming, suggests a common assumption throughout the corpus that military AI capability primarily relies on the ability to perceive, classify and respond to data on a large scale. Deep learning (2,677 mentions) is a subset of machine learning, indicating discussion of contemporary neural architectures rather than classical machine learning (ML) techniques. Convolutional neural networks (1,216 mentions) indicate a particular focus on image processing pipelines, likely due to their use in analysing satellite imagery, recognizing faces and detecting objects for autonomous platforms. Mid-tier technologies include natural language processing (NLP, with 896 mentions in total), autonomous systems (with 501 mentions) and feature extraction (with 425 mentions).

The presence of attention mechanisms[85] and transformers[86] at the lower end of the frequency distribution is noteworthy. Attention mechanisms are a technique that enables neural networks to focus on the most relevant parts of an input, such as a sentence or an image region, rather than processing everything equally. They are the technical building block that made modern LLMs possible. The Transformer is a neural network architecture centred on this mechanism. It was introduced in 2017 and has since formed the basis of models such as BERT and GPT. Their relative scarcity in the corpus, compared to established ML categories, is consistent with publication lag: As research publications typically reflect work

84 As discussed in Section 4, our joint corpus comprises four datasets with highly uneven distribution: research publications (OpenAlex/CSET, n=46,513), Party press (People's Daily, n=39,228), PLA procurement notices (n=271) and governance documents (AGORA, n=26).

85 The foundational paper introducing the transformer architecture is "Attention Is All You Need" by Vaswani et al., pub-lished in 2017 at NeurIPS. See: Ashish Vaswani et al., 'Attention Is All You Need', version 7, preprint, arXiv, 2017, https://doi.org/10.48550/ARXIV.1706.03762.

86 BERT (Bidirectional Encoder Representations from Transformers) by Devlin et al. was published in 2018 and presented at NAACL 2019. Jacob Devlin et al., 'BERT: Pre-Training of Deep Bidirectional Transformers for Language Understand-ing', Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguis-tics: Human Language Technologies, Volume 1 (Long and Short Papers) (Minneapolis, Minnesota), June 2019, 4171–86, https://doi.org/10.18653/v1/N19-1423. GPT-2 by OpenAI, partially released in February 2019, scaled up the original GPT.

conducted one to two years prior to submission, our corpus likely captures the earlier phases of deep learning adoption in military contexts. Transformer-based systems are potentially underrepresented relative to their current operational significance.

feeds into targeting systems, which in turn integrate with command systems to enable rapid decision-making cycles, a core tenet of China's intelligentized warfare concept. The co-occurrence analysis reinforces this integration: Computer vision + surveillance (24 co-mentions) and machine learning + surveillance (23 co-mentions) represent the dominant technology-application pairings. In short, visual intelligence processing constitutes the primary AI-military nexus in our corpus. Moreover, military surveillance (11 mentions) and UAV surveillance (10 mentions) appear as distinct subcategories in our data. This is noteworthy because aerial monitoring is a key capability in the Taiwan Strait and South China Sea. Target detection and target recognition (combined 29 mentions) emphasize precision strike enabling technologies, while automatic target recognition (ATR) (8 mentions) signals interest in closing the "sensor-to-shooter" loop through automated threat identification. Lower-frequency applications include command and control systems, defence systems (various mentions) and military modernization. The relative scarcity of intelligence gathering as an explicit term is notable given the strategic importance of this domain. This may indicate that intelligence func-

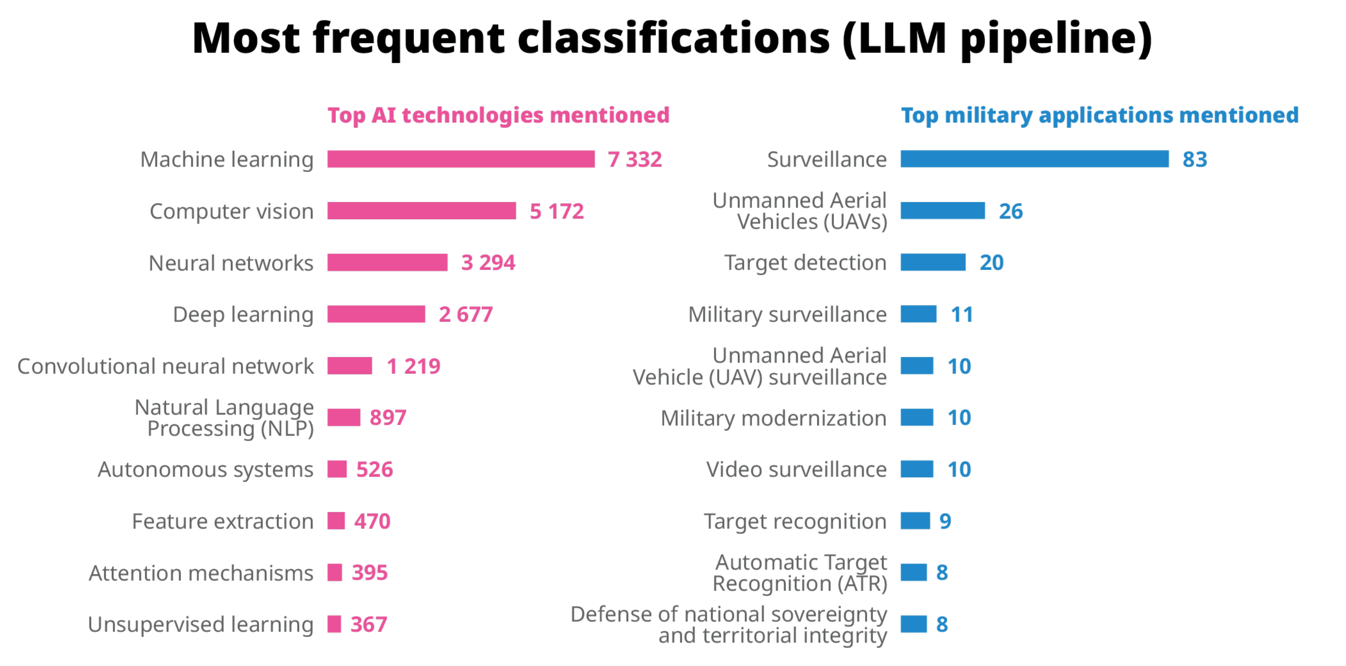


Figure 7: Most frequent classifications (LLM pipeline)
*Source: Authors' generation.*

The military application taxonomy extracted by the LLM reveals surveillance as the dominant functional category (Figure 7, right panel), with 83 mentions – nearly four times more than any other application. Again, this statistical emphasis aligns with the PLA's doctrine of "information dominance" that is central to the Military Strategy that the Communist Party's Central Military Commission passed in June 2004 in reaction to the lessons drawn from the Second Gulf War.[87] It also tallies with China's domestic deployment of AI-powered surveillance infrastructure, which provides both technical capability and operational experience transferable to military contexts. The prevalence of surveillance applications likely reflects dual-use technology transfer from China's extensive civil security apparatus to military intelligence, reconnaissance and surveillance (ISR) systems. Unmanned aerial vehicles (UAVs) constitute the second tier with 26 mentions (consolidating variants), followed by target detection (20 mentions). Both categories are functionally linked: AI-enhanced surveillance

87 M. Taylor Fravel, Active Defense: *China's Military Strategy* since 1949, Princeton: Princeton University Press, 2019, 218.

tions are discussed using more specific terminology (e.g., reconnaissance, ISR, signals intelligence) not captured by our normalization schema.

while, the number of "Military Only" documents remains consistently low until a sharp spike in 2021, potentially corresponding to heightened geopolitical tensions or in-

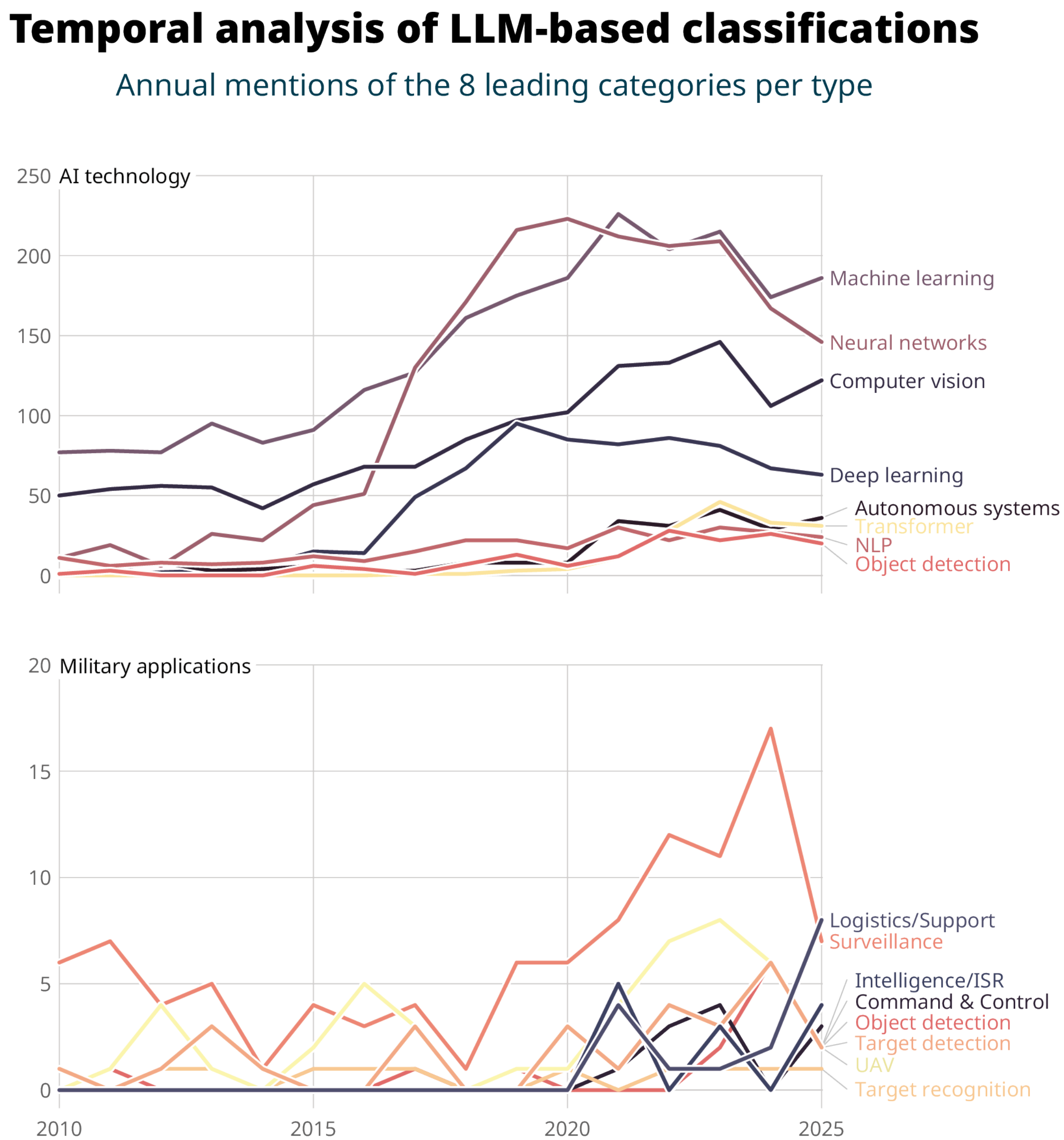


Figure 8: Temporal analysis of LLM-based classifications.
*Source: Authors' generation.*

Temporal analysis reveals distinct evolutionary phases in China's military AI discourse (Figure 8). Documents classified as "AI Only" show steady growth from approximately 400-500 mentions in 2010-2014 to a peak of 1,620 mentions in 2023, followed by a decline to 1,000-1,400 in 2024-2025. This decline likely reflects either corpus composition effects (incomplete coverage of recent publications) or a genuine shift towards more integrated discussions that combine AI and military framing. Meanwhile, the number of "Military Only" documents remains consistently low until a sharp spike in 2021, potentially corresponding to heightened geopolitical tensions or internal PLA modernization directives. The subsequent decline suggests that a purely military discourse without AI framing has become less common. Documents addressing both AI and military domains remain near-zero until 2020, then rise to 50-70 annual mentions in 2021-2023. While absolute numbers remain small, this emergence marks a qualitative shift: the formation of a distinct "military AI" discursive space where technical capabilities and operational applications are explicitly linked. This timing roughly coincides with the PLA's formal adoption of the intelligentized warfare doctrine and the establishment of dedicated military-civil fusion (MCF) mechanisms for AI technology transfer: The Central Military-Civil Fusion Development Committee (CMCFDC) was formally established on January 22, 2017, under Xi Jinping's leadership.[88] Technology-specific trends show machine learning and computer vision growing in parallel from 2015 onwards, reaching peaks around 2022.

**Tracking the top 10 AI technologies across the entire body of news and research articles reveals clear temporal patterns (see Figure 9). The three technologies with the highest volume of mentions are neural networks (10,538), deep learning (8,104) and convolutional neural networks (7,418)**. This differs somewhat from the LLM-classified sample, where machine learning placed first. This divergence is methodologically informative. Keyword searches in the full corpus reveal technical terminology as it appears in specialized research writing, whereas LLM classification captures conceptual discussions in which technologies are described functionally ("learning algorithms," "AI systems") rather than being named precisely. The two methods therefore complement each other.

**The temporal trajectories tell a coherent story of successive adoption waves**. Machine learning established the baseline, growing steadily from around 50 monthly mentions in 2015-16 to peaks above 200 by 2022-23. Neural networks and deep learning followed a steeper curve, remaining below 100 monthly mentions until late 2020, when they began their surge to 400-500 by mid-2023, roughly a fivefold increase in three years. This coincided with the wider accessibility of deep learning frameworks, such as *PyTorch* and *TensorFlow*, to Chinese research institutions. Convolutional Neural Networks (CNNs)[89] tell a different story, with steady growth until 2022, followed by a plateau. This is consistent with CNNs becoming an established baseline tech-

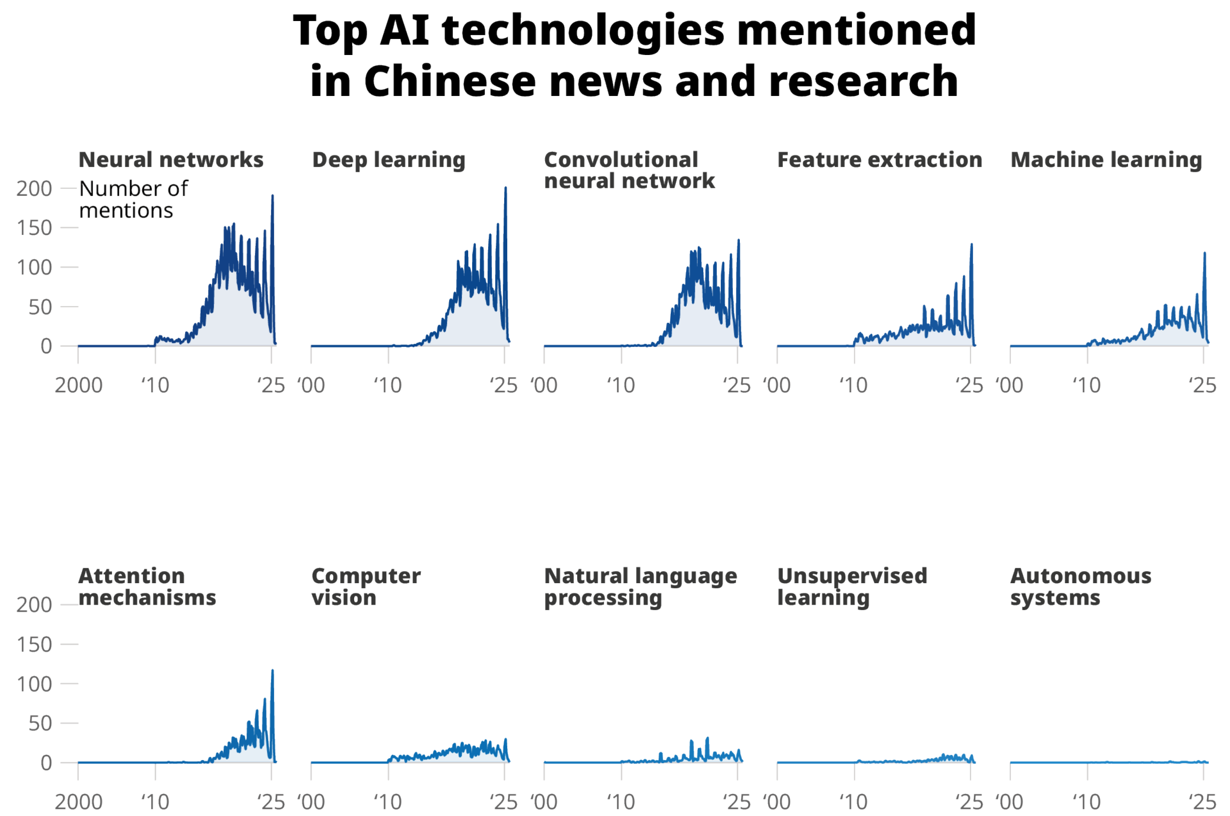


Figure 9: Top AI technologies mentioned in Chinese news and research.
*Source: Authors' generation.*

88 See: Li Cheng, "China's Military-Civil Fusion: Objectives and Operations," China-US Focus, 30 August 2022, https://www.chinausfocus.com/2022-CPC-congress/chinas-military-civil-fusion-objectives-and-operations.
89 See Glossary.

nology rather than the subject of active frontier research. By contrast, attention mechanisms rose from near-zero to 100-150 monthly mentions between 2020 and 2024, marking the arrival of transformer-based approaches in the corpus.

One discrepancy in particular deserves attention. **Autonomous systems account for only 48 total keyword mentions across the entire corpus, yet they are ranked seventh in the LLM-classified sample**. This is not a measurement error; it reflects a genuine linguistic pattern. It seems that in Chinese military AI discourse, autonomy is typically expressed indirectly through terms such as "unmanned platforms," "intelligent systems" and "self-learning capabilities," rather than directly through the term "autonomous systems." The LLM's semantic understanding captures this usage, whereas a keyword search does not. This highlights a broader methodological point: Lexical and semantic approaches are not interchangeable, and combining them reveals empirical information that either method alone would overlook.

Military application trends, again measured by tracing the top LLM-detected keywords in the full news and research corpus, reveal a different evolutionary pattern (Figure 10). Surveillance dominates with 4,329 total document mentions, remaining relatively stable at 50-80 monthly mentions through 2018, then exhibiting a sharp upward trajectory beginning in 2019 that peaks at 200-250 mentions in late 2022 before stabilizing around 150-180 through 2024-2025. This earlier and more pronounced peak compared to most AI technologies suggests that surveillance applications drove the initial military interest in AI capabilities, consistent with China's extensive domestic experience deploying AI-powered surveillance infrastructure in Xinjiang and urban security contexts. Unmanned aerial vehicles (UAVs) exhibit a dramatic spike: Mentions rise from approximately 20-30 per month in 2016-2019 to a sharp peak exceeding 200 mentions in late 2021, then decline to sustained levels of 80-120 through 2023-2025. The 2021 peak predates Russia's invasion of Ukraine (February 2022) and likely reflects China's observation of UAV employment in the 2020 Nagorno-Karabakh conflict, which demonstrated the operational effectiveness of unmanned systems in contested airspace, a scenario highly relevant to Taiwan Strait contingencies. The subsequent normalization suggests UAVs transitioned from novel capability to as-

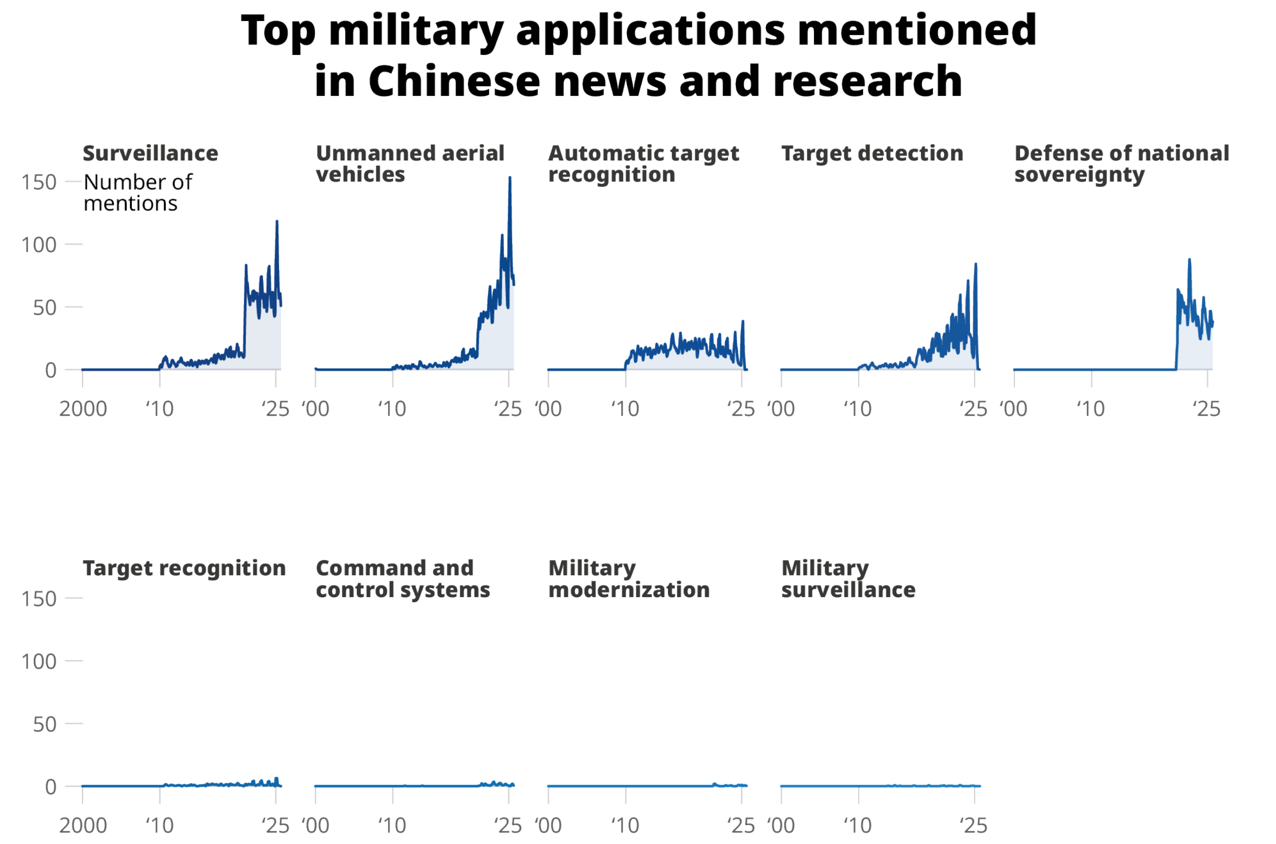


Figure 10: Top military applications mentioned in Chinese news and research.
*Source: Authors' generation.*

sumed platform category in military AI discourse. **Target detection (2,664 documents) and automatic target recognition (ATR, 2,977 documents) collectively represent 5,641 mentions focused on precision targeting capabilities. While the LLM sample showed relatively modest mentions (20 for target detection, 8 for ATR), the full corpus reveals these as major research themes.** Both exhibit steady growth from 2018 onwards, with target detection rising to 100-120 monthly mentions by 2024 and ATR reaching 120-150 mentions in the same period. This convergence and recent acceleration may indicate a doctrinal shift from defensive ISR postures towards offensive precision-strike capabilities, consistent with PLA modernization priorities emphasizing long-range fires and "systems destruction warfare"[90] for Taiwan contingency planning. Defence of national sovereignty and territorial integrity appears 2,386 times, a surprisingly high frequency for what seems like a policy slogan rather than a technical application. Temporal analysis reveals this term surges during specific geopolitical events: Peaks occur in mid-2016 (South China Sea arbitration ruling), late 2019 (Hong Kong protests) and mid-2022 (Nancy Pelosi's Taiwan visit), suggesting its use as a rhetorical frame linking AI capabilities to sovereignty defence narratives in response to perceived threats. Finally, command and control systems (64 mentions) and military modernization (21 mentions) appear far less frequently than anticipated, given their strategic importance. Again, this likely reflects terminological fragmentation, The low count for "military modernization" is particularly revealing: Despite being a core PLA priority, the term functions more as a macro-level policy frame than a specific research topic.

Disaggregating by sub-corpora reveals functional differentiation between news and research sources. *People's Daily* (114,879 documents, 71.8% of the corpus) contributes disproportionately to mentions of broad categorical terms, such as surveillance, UAVs and defence of national sovereignty, consistent with the news media's role in signalling doctrinal priorities and operational readiness to domestic and international audiences (see Introduction and Section 4). Peak mention periods in *People's Daily* correlate with major policy announcements such as the PLA's 2022 "20th Party Congress" modernization goals and geopolitical events, as listed above. Research documents (55,943 documents, 35.0%) drive mentions of specific technical terms: Convolutional neural networks, attention mechanisms, feature extraction and automatic target recognition appear at two to three times the rate in research versus news when normalized by corpus size. For military applications, research sources contribute more evenly across categories but show particular concentration in target detection and ATR, reflecting these as active R&D domains rather than operational capabilities. Overall, these patterns confirm our division of labour hypothesis: Research institutions prototype capabilities and publish technical findings in specialized venues; state media amplify strategic messaging and frame AI-military integration within national achievement narratives; and policy documents formalize strategic doctrine. The temporal lag between research spikes and news coverage, typically 12-18 months, suggests a technology maturation pipeline where academic breakthroughs diffuse into military applications. Once operationalization is sufficiently advanced to serve signalling, they then become subjects of public discourse.

## The AI-military nexus in high-density documents: findings

To understand how China's research and media ecosystems frame the AI-military nexus, we examine the 20 most AI-intensive documents that also reference military concepts. Our analysis reveals a striking pattern: **Even in documents with the highest concentration of AI terminology, military keywords rarely feature in substantive discussions about the application of AI to weapons, battlefield operations or autonomous combat systems**. Instead, AI and military topics tend to co-occur incidentally – for example, a research paper on machine learning may reference "national security" in its introduction, or a *People's Daily* article on defence modernization may mention "intelligent systems" in passing — without the two domains being genuinely integrated into the argument. In other words, they are parallel discourses rather than a fused one. To be precise: Among the 170,822 news and research documents, only 4,345 (2.5%) contain both AI and military terminology, a surprisingly low intersection given China's stated military modernization priorities. More revealing, qualitative examination of the top 20 AI-dense documents with military keywords exposes systematic false positives driven by dual-use terminology.[91] This is an important methodological caveat that one should keep in mind in future corpus-linguistic studies, as it shows why it is so beneficial to use context-aware LLMs instead of simple keywords approaches. When genuine military-relevant

90 See: Engstrom, J. (2018), Systems Confrontation and System Destruction Warfare: How the Chinese People's Liberation Army Seeks to Wage Modern Warfare, RAND, https://www.rand.org/pubs/research_reports/RR1708.html.

91 The document with the highest AI density (22 AI term mentions, January 2025) discusses Vision Transformers for electrical power grid anomaly detection. It was flagged by our simple keyword counting for "surveillance" because it references "real-time smart grid monitoring", i.e. a purely civilian infrastructure application. The second-ranked document (21 AI terms, May 2022) applies CNNs to "fresh fruit production", flagged for "target detection" due to references to "object detection" for identifying fruit during automated harvesting. Similarly, 10 of 27 extracted "military" passages reference "automatic target recognition" (ATR), but (AI-translated) reading suggests these discuss mostly matrix operations in neural network architectures rather than military targeting systems.

content appears, it is indirect and research-oriented. For instance, one paper from 2019 discusses synthetic aperture radar (SAR) for "target recognition," representing authentic dual-use research applicable to both civilian remote sensing and military reconnaissance. However, the paper frames this purely as a computational efficiency challenge, compressing deep neural networks for "real-time recognition system" with limited memory, without any discussion of operational military deployment, weapons integration or strategic doctrine.[92] Critically, all 20 top AI-dense documents with military keywords originate from the research corpus (OpenAlex/CSET); zero come from *People's Daily*. This source distribution fundamentally contradicts any expectation that the state media might be actively promoting military AI capabilities.

The pattern described by our NLP-derived statistics suggests that it is difficult to identify an integrated military AI discourse in official, public text data. High-frequency AI discussions in accessible publications focus overwhelmingly on economic development (smart grids, precision agriculture), civilian infrastructure (transportation, healthcare) and technical research challenges (model compression, hardware acceleration). When military terminology appears, it does so either as: (1) false positives from dual-use technical vocabulary (monitoring, detection, recognition); (2) implicit dual-use research without explicit operational framing (SAR image processing); or (3) civilian UAV applications explicitly contextualized alongside consumer technologies. The near-total absence of *People's Daily* articles in the AI-dense military subset is particularly revealing. If China's propaganda apparatus were actively promoting military AI readiness, as suggested by the sentiment-salience correlation shift in Section 6.1, we would expect the state media to produce AI-saturated content with an explicit military framing. Instead, the data indicate that genuine military AI integration discussions occur in classified or specialized military publications not captured in our open-source corpus, while publicly accessible research focuses on foundational technologies with plausible deniability (or genuine civil uses) regarding end-use applications.

These findings generate several strategic insights, though they require careful interpretation given corpus limitations. First, the overwhelming presence of surveillance, target detection and automatic target recognition in keyword searches suggests that Western analyses relying solely on keyword-based methods may systematically overestimate the maturity of China's military AI (in the narrow sense of lethal autonomous weapon systems) integration by conflating civilian computer vision research with operational military capabilities.

Second, the genuine dual-use research identified (SAR target recognition, UAV navigation) indicates that China is building foundational capabilities transferable to military applications, but the research framing emphasizes computational efficiency and algorithm optimization rather than weapons integration. However, it should also be noted that the prevalence of target detection and target recognition could constitute foundational research that will become generally applicable to various military future military technologies. Such an interpretation would suggest a long-term strategy in which weapons integration is still to come at a later stage, but one that is not necessarily defensive. This suggests PLA capabilities may emphasize enabling technologies (sensor processing, autonomous navigation) rather than fully integrated autonomous weapons systems, consistent with the low mention frequency of "autonomous weapons" in our LLM classification.[93]

Third, the temporal concentration in 2017-2019 followed by resurgence in 2024-2025 may reflect research publication cycles lagging operational deployment by 2-3 years. If so, the current wave of technical publications could indicate that capabilities prototyped in 2021-2022, coinciding with heightened Taiwan Strait tensions, are now sufficiently mature to enter open literature as foundational research, while cutting-edge operational systems remain classified.

Finally, the complete absence of high-density military AI discussions in state media suggests that China's confident public messaging about military modernization – evident in sentiment analysis – operates mainly at a strategic or declaratory level. **This compartmentalization may serve the purpose of maintaining diplomatic flexibility in international AI / weapons debates while signalling general deterrence capability without revealing technical details that could inform adversaries**.

92 The co-occurrence matrix further confirms the dominance of civilian computer vision research: neural networks + automatic target recognition (7 co-occurrences), neural networks + target detection (6), and CNN + target detection (6) rank highest. However, as documented in the preceding footnote, these primarily reflect agricultural monitoring, medical diagnosis,and remote sensing rather than weapons applications.

93 An analysis of NATO member states' AI strategies also reveals a heavy emphasis on sensors. See: Maggie Gray and Amy Ertan, 'Artificial Intelligence and Autonomy in the Military: An Overview of NATO Member States' Strategies and Deployment', NATO Cooperative Cyber Defence Centre of Excellence, 2021, https://ccdcoe.org/library/publications/artificial-intelligence-and-autonomy-in-the-military-an-overview-of-nato-member-states-strategies-and-deployment/.

## 6.4. Stage 4: Hypothesis-driven concept induction via LLooM

We applied LLooM to the 26 AGORA governance documents, operationalizing each of the seven hypotheses as a custom concept prompt (see Section 5, Stage 4 for full methodological details). When applying LLooM to our regulatory dataset, the algorithm discovered (and labelled) the following key concepts: (1) *AI safety and security*, capturing texts that discuss measures to prevent harm or mitigate risks associated with AI systems; (2) *Ethics governance framework*, capturing texts that describe the establishment of comprehensive governance frameworks for AI ethics and oversight; (3) *Human-centric AI development*, capturing texts that emphasize human-centred design and the prioritization of human well-being in AI systems; (4) *International cooperation and standards*, capturing texts that highlight the importance of cross-border cooperation, standardization and governance for AI; and (5) *Transparency and explainability*, capturing texts that stress the need for transparent, explainable and auditable AI decision-making processes. Together, these five concepts capture the main thematic patterns in the Chinese AI governance corpus. Taken as a whole, they map the regulatory discourse around AI as one that is explicitly framed through safety, ethics and governance rather than driven by purely technical or military concerns. Based on our close reading of the corpus, these emergent concepts, summarized in Table 1, provide a useful overview of how Chinese governance documents prioritize AI development around regulatory control and security applications. This in turn increases our confidence in the model and allows us to proceed with more explicit hypothesis testing.

In the second phase, we thus operationalized each of our seven hypotheses outlined in Section 3 as custom concept prompts. Across the 26 AGORA governance documents, the hypothesis-driven concepts produce a differentiated picture of how explicitly Chinese regulatory texts engage military AI questions (Figure 11). H7 (human control and autonomy) and H6 (offense–defence balance) show the highest levels of support, with 65.4% and 53.8% of documents respectively scoring at least one strong match (score $\geq 0.75$), and mean scores of 0.67 and 0.58 on the 0–1 scale. H5 (technology denial and economic coercion) is also relatively prominent, with 38.5% of documents scoring as strong matches and a mean score of 0.53, whereas H1 (demographic force reduction) and H2 (nuclear stability) are much less frequently supported, with only 15.4% and 3.8% of documents respectively containing passages that the model classifies as clear evidence for those hypotheses. These aggregate results suggest that **Chinese AI governance texts are more inclined to address questions of strategic stability, offense–defence dynamics and human control than to frame AI explicitly as a solution to demographic constraints or to directly couple AI with nuclear deterrence**.

## Table 1: LLooM-based document-level scores and prevalence per concept

| | Concept | Criteria | Summary | Prevalence | n |
|---|---|---|---|---|---|
| 0 | AI safety and security | Does the text discuss measures to prevent harm or mitigate risks associated with AI systems? | Ensuring AI safety and security through ethics integration, risk management, transparency, accountability and adherence to laws and regulations. | 0.81 | 21 |
| 1 | Ethics governance framework | Does this text mention establishing a framework for AI ethics and governance? | Establishing comprehensive ethics governance frameworks for AI development and deployment to ensure accountability, fairness and safety. | 0.85 | 22 |
| 2 | Human-centric AI development | Does the text emphasize the need for human-centred design and development of AI systems that prioritize human well-being? | Developing AI that prioritizes human values, safety and respect for privacy, fairness and inclusivity. | 0.31 | 8 |
| 3 | International cooperation and standards | Does the text discuss the importance of international cooperation, standardization and governance in AI development and deployment? | International cooperation and standards are crucial for ensuring AI development prioritizes ethics, safety, fairness, openness and peaceful use. | 0.42 | 11 |
| 4 | Transparency and explainability | Does the text highlight the importance of transparency, explainability and audibility in AI decision-making processes? | We emphasize the importance of transparency and explainability in AI activities for ethics, safety, fairness and human control. | 0.27 | 7 |

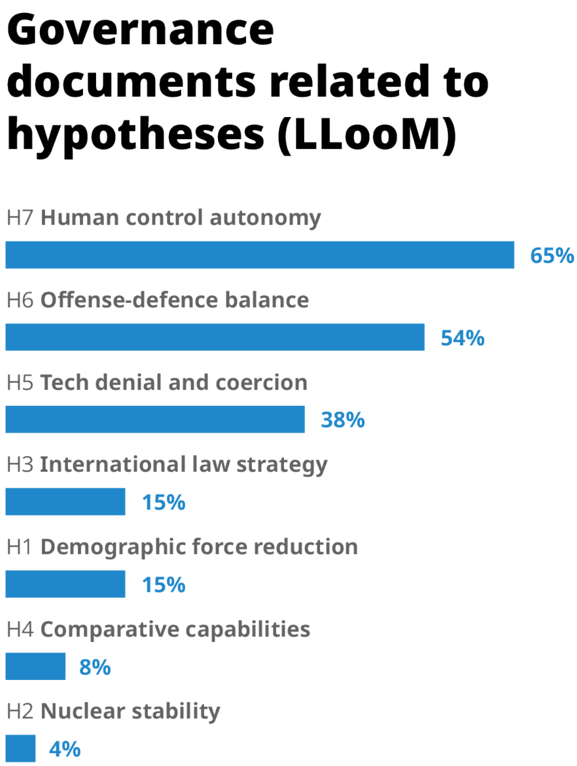


Figure 11: Governance documents related to hypotheses (LLooM)
*Source: Authors' generation.*

The highlight snippets extracted in the second phase of the LLooM pipeline help clarify what "support" (in the sense of LLooM) looks like for each hypothesis in concrete terms.[94] For H6 (offense-defence balance), strong matches cluster around language that links intelligentized or AI-enabled systems to questions of operational initiative and tempo as well as the capacity to detect, track and intercept threats more effectively, often in the context of "intelligentized" command-and-control or integrated air and missile defence. In these passages, AI is cast as a capability that can compress decision cycles and enhance situational awareness. For H7 (human control and autonomy), the matched highlights consistently emphasize requirements for "human oversight," "responsible use" or "meaningful human control" over AI systems (including algorithmic decision-making in security-relevant contexts) and stress that AI-enabled systems must remain subject to legal and ethical constraints rather than operating as fully autonomous agents. The prevalence and content of these highlights indicate that **Chinese governance texts devote substantial attention to the normative problem of how far decision authority should be delegated to AI**, even in a regulatory domain that is not purely military.

By contrast, highlights supporting H1 (demographic force reduction) are rarer and more indirect. H1 was formulated because a significant strand of Western strategic analysis argues that China's drive towards intelligence is partly motivated by the need to offset the shrinking military-age population – the legacy of the one-child policy.[95] Therefore, testing whether this framing appears in Chinese governance texts is non-trivial. However, where strong matches occur, they typically involve passages that link automation, intelligentized systems or unmanned platforms to reducing frontline risk, lowering personnel exposure or optimizing force structures rather than explicitly naming demographic decline or recruitment crises. This pattern suggests that while the idea of "fighting with fewer people" appears in the governance corpus, it is not yet a central or consistently articulated theme in regulatory texts and may be more prominent in doctrinal or operational writings outside the AGORA sample. This result is consistent with the fact that the CCP values military service as an opportunity for indoctrination through the work of political commissars in the PLA.[96] It is therefore plausible that China may not prioritize force reduction even if military technology may permit this in the future. Moreover, as mentioned above, China is currently more concerned with the quality than the quantity of recruits; the urgency to find technological solutions to the demographic crunch therefore does not seem to present itself (yet). For H2 (nuclear stability), the model finds very little explicit discussion: The few strong matches revolve around general references to strategic stability, crisis management and high-risk applications of AI in security contexts but do not systematically connect AI to nuclear command-and-control or early-warning archi-

94 These highlight snippets, which we use in the following analysis, are produced by LLooM's scoring operator as part of a structured prompt: For each active concept and document, the model is asked not only to assign a discrete score on the 0–1 scale but also to identify and return a short span of text that best justifies that score in light of the concept's inclusion criteria. In practice, this means that whenever a document receives a high score (e.g. ≥0.75) for a specific hypothesis (from H1 to H7), LLooM also records the precise sentence or paragraph that most caused the model to judge the hypothesis as supported. These machine-selected excerpts can then be used by researchers for inspection, alongside the regulation title and date. This mechanism, based on iterative LLM inferences, enables a qualitatively grounded reading of why a given document is treated as evidence for a particular strategic claim rather than relying solely on aggregate scores.

95 CSET notes that Chinese defence scholars argue that AI will "gradually remove humans from future battlefields" and enable the PLA to reduce its reliance on human labour: Sam Bresnick, "China's Military AI Roadblocks," Issue Brief, Center for Security and Emerging Technology, June 2024, https://cset.georgetown.edu/wp-content/uploads/CSET-Chinas-Military-AI-Roadblocks-1.pdf. CSBA identifies China's "heavy demographic burden" as a structural headwind that renders intelligentization both strategically appealing and economically challenging: Chris Bassler and Benjamin Noon, "China's Ambitions for AI-Driven Future Warfare," CSBA (Center for Strategic and Budgetary Assessment), 1 January 2022, https://csbaonline.org/about/news/chinas-ambitions-for-ai-driven-future-warfare. The following blog analysis explicitly describes the PLA's efforts to become more intelligent as "perhaps the most ambitious demographic offset programme in human history," linking autonomous systems directly to the legacy of the one-child policy and the shrinking youth population: Ronin's Grips, "Impact of China's Demographic Shift on PLA Strategy," 16 February 2026, https://blog.roninsgrips.com/impact-of-chinas-demographic-shift-on-pla-strategy/.

96 Joel Wuthnow and Philip Saunders, China's Quest for Military Supremacy, Cambridge: Polity, 2025.

tectures. This absence is itself substantively meaningful, as it implies that **Chinese public-facing AI regulation documents may deliberately avoid foregrounding nuclear issues**, leaving those debates to more classified or doctrinal venues.

Meanwhile, H3 (international law and governance) and H4 (comparative capabilities) occupy an intermediate position. For H3, strongly matching highlights frequently reference participation in international organizations, standard-setting processes or treaty-like instruments governing AI ethics, data protection and high-risk applications. These passages fit well with a view of China as actively engaging global AI governance forums to shape norms in ways that protect its room for manoeuvre while projecting the image of acting as a responsible power. H4's matched snippets, by contrast, tend to invoke comparative language about "catching up," "leading in key AI technologies" or leveraging data and industrial advantages, yet rarely offer detailed assessments of relative military balance. The signal is present but muted in this regulatory corpus, which is somewhat unsurprising, given AGORA's focus on governance rather than doctrinal analysis. Taken together, the quantitative pattern (relative support levels across H1–H7) and the qualitative content of the highlights indicate that Chinese AI regulatory texts most consistently engage questions of human control, strategic stability in a broad sense and technology-denial dynamics while largely sidestepping explicit nuclear and demographic-military linkages in this particular corpus.[97]

97 This is supported by the fact that the independently induced "Human-Centric AI Development" concept (independent, since it is caused by our initial "seed text") is supported by 69.2% of documents with a mean score of 0.74. In other words, much of the regulatory language frames AI development through human-centered values, safety and social well-being. In several cases, the highlights for H7 (human control and autonomy) overlap conceptually with those for human-centric AI, reinforcing the impression that concerns about human primacy and oversight are not confined to military applications but embedded in a broader governance discourse.

7

# Conclusion: Discussion, policy recommendations and future research

This study examined how China presents artificial intelligence in its military discourse, analysing four document types: official AI governance regulations, Chinese research publications, state media coverage in *People's Daily* and People's Liberation Army procurement contracts. Using a combined computational and qualitative methodology, we tested several theoretical hypotheses. The main results are:

**After 2020–22, China's AI-military discourse underwent a decisive shift, moving from aspiration to integration**. AI is now treated as a routine, multi-domain enabler of PLA operations across all four document types; it is no longer considered a speculative future capability.

**Discourse on AI warfare is structurally segmented**. Research institutions, state media, regulatory bodies and procurement offices all have their own distinct register. Only 2.4% of documents discuss genuine military AI integration, suggesting that cutting-edge operational capabilities are only discussed in classified sources, while public discourse focuses on foundational dual-use technologies.

**The sentiment surrounding military AI in *People's Daily* is consistently positive and increasingly intensive, with a strong correlation between the prominence of military AI topics and optimistic framing tightening sharply after 2023**. The March 2025 spike in sentiment, occurring precisely when the concept of dissipative warfare entered PLA Daily, contrasts markedly with the more anxious AI-related sentiment evident among Western publics during the same period.

**Rhetoric around human control is declining in state media, even though it persists in research publications**. "Meaningful human control" references in *People's Daily* peaked in 2021, coinciding with the height of international pressure during UN LAWS negotiations, and then declined sharply. While Chinese researchers maintain technical caution with human-in-the-loop framing, state media does the inverse. This divergence suggests that China's engagement with human-control norms has served as diplomatic rhetoric rather than representing a genuine commitment to these norms.

These findings do not support the idea of imminent autonomous military dominance nor do they dismiss it outright. PLA AI applications remain anchored in surveillance, computer vision and command support functions, i.e. foundational layers rather than autonomous weapons systems. However, the trajectory, sentiment environment and normative retreat on human control all carry serious implications for Western policymakers.

## Policy implications

**The fact that published Chinese attitudes towards military AI are uniformly positive should inform NATO's threat assessment**. Sustained domestic confidence is likely to increase assertiveness among Chinese decision-makers while generating popular expectations of military prowess that could restrict leaders' ability to de-escalate in a crisis. At the same time, our results suggest that PLA AI adoption remains at a relatively early operational stage, comparable to the slow pace of documented adoption by the US military.[98] This sug-

98 This is most prominent in Christian Brose, *The Kill Chain: Defending America in the Future of High-Tech Warfare*, New York City: Grand Central Publishing, 2020 but also comes through in Paul Scharre, *Army of None: Autonomous Weapons and the Future of War*, New York City: W.W. Norton & Company, 2019.

gests that operating on worst-case assumptions about the deployment of lethal autonomous weapons systems could fuel an arms race dynamic that is not yet warranted by operational realities.[99]

**The surge in AI-military sentiment after 2023, which coincided with the introduction of China's dissipative warfare doctrine, is a particular cause for concern**. It contrasts with an observable negative sentiment shift in the West concerning prospects for the future and the impact of AI.[100] China's AI strategy prioritizes comprehensive battlespace awareness as an offensive enabler. Under this doctrine, surveillance is not a defensive posture but a precondition for decision superiority and surprise targeting. This reframes the surveillance dominance visible across all four datasets: It compresses adversary response times without removing humans from lethal decision-making processes.

NATO has formulated responsible military AI adoption (both by the Alliance itself and its international promotion) as a key goal.[101] Regarding diplomacy and arms control, the 2024 US-China understanding on AI in nuclear command and control shows that AI diplomacy can produce tangible outcomes, even amid strategic competition. **The view that the pace of the AI arms race renders diplomatic constraints impossible, expressed by senior Western officials as recently as December 2025,[102] is counterproductive**. Diplomacy is most necessary precisely when tensions are high, as demonstrated by the record of arms control during the Cold War. Our findings on H3 (China's lawfare strategy) reveal a consistent pattern of attempts to establish selective international norms. Chinese governance texts advocate international AI regulation in areas where China is lagging behind, such as model transparency, algorithmic auditing and foundation model governance, while resisting restrictions in areas where China is leading the way, particularly in the fields of surveillance and autonomous logistics. This pattern should form the analytical baseline for all diplomatic engagement. Negotiations that remain at the level of abstract principles ("AI safety," "ethical development") perpetuate this dynamic; effective dialogue must be anchored in domain-specific technical definitions. What constitutes meaningful human control in drone swarm coordination?[103] Which ISR activities near adversary borders constitute crisis escalation triggers? **It is such questions at the level of concrete applications that can be used to test China's rhetorical commitments against operational practice**. NATO member states can only win by addressing these issues in this way. Either they lead to meaningful agreements with China on military AI, reducing the risk of escalation and crisis, or they expose China's selective lawfare strategy in a way that maximises attributability (i.e. in that case demonstrating that China is to blame for the lack of meaningful international regulation of AI).

Based on these general strategic implications, we formulate a set of more specific policy recommendations below.

## Specific recommendations

**Strengthen total defence as a key aspect of deterrence**. The close relationship between the prominence of AI in the military and optimistic sentiment surfacing in our data, when considered alongside China's dissipative warfare doctrine, supports the claim for the strategic importance of total defence, particularly in terms of civil defence and cyber security. Societal resilience and cyber hardening are assets for deterrence in an information-domain contest, not merely defensive measures.

**Invest in countermeasures targeting China's core AI advantages**. Three priority areas follow directly from our findings in the application domain. First, NATO members should invest in adversarial robustness research to develop AI systems that are resilient to "data poisoning" (injecting/creating malicious data to corrupt pattern recognition), model extraction (reconstructing proprietary models through repeated queries) and attacks that can blind, misdirect or compromise military decision-making systems.[104] Second, they should invest in counter-ISR technologies, including adaptive camouflage, decoy generation and emissions management to degrade

99 Note that in terms of actual military spending on AI (compared with military spending on "legacy systems") by both the United States and China, it is difficult to speak of an AI arms race that has already begun, as argued in Paul Scharre, *Four Battlegrounds: Power in the Age of Artificial Intelligence* (New York: W. W. Norton & Company, 2023).

100 Shira Ovide, 'Americans have become more pessimistic about AI. Why?', Washington Post, 7 October 2025; Noemi Dreksler, Stephen Clare, Kaylyn Jackson Schiff, Daniel S. Schiff, Chloe Ahn, Zachary Peskowitz, and Harry Law, 'What the public thinks about AI and the implications for governance', Brookings, 9 April 2025.

101 NATO, 'Summary of NATO's revised Artificial Intelligence (AI) strategy', 10 July 2024.

102 To cite just one anecdotal example: At a foreign policy conference in December 2025, Thomas Jarzombek, State Secretary at the German Ministry for Digital Transformation and Government Modernization, argued that given how relentless the American-Chinese race to dominate AI is being pursued there is no chance for AI diplomacy. (On the record statement made at the Chatham House Strategic Choices Conference in Berlin, 1-2 December 2025. One of the authors was in attendance.)

103 On Chinese drone swarm technology, compare Michael C. Loftus, 'Machines in the Alleyways: China's Bet on Autonomous Urban Warfare', *The Diplomat*, 3 February 2026, https://thediplomat.com/2026/02/machines-in-the-alleyways-chinas-bet-on-autonomous-urban-warfare/.

104 Aaron Conti, "Data Poisoning as a Covert Weapon: Securing U.S. Military Superiority in AI-Driven Warfare", Lieber In-stitute, West Point, June 30, 2025. See, in general: Alexandra Souly et al., 'Poisoning Attacks on LLMs Require a Near-Constant Number of Poison Samples', version 1, preprint, arXiv, 2025, https://doi.org/10.48550/ARXIV.2510.07192.

potential surveillance-targeting kill chains.[105] Third, they should invest in human-machine teaming architectures that demonstrate how HITL systems can match or surpass autonomous systems in decision speed and accuracy, countering the narrative that human control requires accepting operational inferiority.

**Extend and deepen the American-Chinese nuclear AI understanding**. The 2024 understanding should be complemented by agreed definitions of which AI applications in nuclear infrastructure are admissible (e.g. safety analysis and anomaly detection) and which are not. Diplomatic efforts should extend these commitments to other nuclear-armed states.

**Pursue binding international commitments on high-risk applications**. Treaty obligations should target autonomous target selection and "non-reconstructible" black box military AI systems, whose battlefield actions cannot be audited for compliance with international humanitarian law. We propose the term "human-beyond-the-loop" for this category to distinguish it analytically from "human-in-the-loop" and "human-on-the-loop" architectures.

**Build verification capacity**. Our corpus analysis confirms that genuine discourse on the integration of AI in the military is almost entirely absent from open Chinese sources. The significant disparity between China's substantial civilian AI research output and its limited public military discourse means that open-source monitoring alone cannot provide a basis for verification. Effective arms control requires technical measures that are both intrusive and analogous to the inspection protocols of the Chemical Weapons Convention. These measures must include shared standards for AI system documentation, testing protocols for autonomous weapons and cryptographic on-chip governance mechanisms (including location-verification features) that can confirm that the hardware has not been redeployed to unauthorized facilities.[106]

**Form coalitions to promote meaningful CCW regulation**. Pursue binding LAWS commitments through the CCW Review Conference while also countering China's selective norm strategy within this forum. Coalition building must extend to non-Western states subject to Chinese influence, including Belt and Road partners, given that China increasingly uses these relationships in its diplomatic strategy to dilute multilateral AI governance efforts. Our H3 findings provide the evidential basis for explicitly naming this pattern in diplomatic settings.

## Future research

This study opens several avenues for future research, which we briefly sketch below.

**Cross-corpus validation**. The hypothesis-driven findings reported here are based on the AGORA governance corpus. Applying the same framework systematically to the *People's Daily* and procurement datasets would reveal whether the discourse patterns identified at the regulatory level (particularly the H6 offence-defence signal and the H7 human-control trajectory) are mirrored or contradicted in operational and propaganda registers. Alignment across registers would strengthen our conclusions.

**Research on the March 2025 sentiment shift**. The sharp positive sentiment spike observed in *People's Daily* in March 2025, which coincided with the introduction of dissipative warfare framing in PLA Daily, warrants further investigation. A comparative analysis of Western military AI discourse over the same period would establish whether this represents a genuine divergence in strategic outlook between China and the West, as our preliminary analysis suggests, or a parallel response to common technological developments.

**Human control and NC3 as qualitative priorities**. Our quantitative analysis reveals that human-in-the-loop terminology appears prominently in a significant proportion of texts. The normative and strategic significance of this pattern justifies a qualitative analysis, which would determine whether these references reflect genuine design constraints, diplomatic positioning or residual rhetorical habits. A related priority is AI in nuclear command, control and communications (NC3), where our text mining yields weaker and less consistent results (likely because this subject is discussed only indirectly or in classified genres). Given the stakes, this gap makes NC3 an important candidate for track-two diplomatic dialogue and for research methodologies that go beyond open-source corpus analysis.

105 Nicholas Barber, "Making Sense of Accelerated Warfare: Army's Adaptive ISR Capability", *Australian Army Journal* 16/1 (2020).
106 On the GPU side, this could be enforced through technical means such as on-chip governance mechanisms or through "call-home" features that use delay-based location verification. See: Jonathan Stein, "Chip Location Verification is the New Export Control Battleground," Transformer, 9 September 2025, https://www.transformernews.ai/p/ai-chip-location-verification-nvidia-china-csa.

**Methodological development for sensitive corpora**. Combining exploratory concept induction with targeted hypothesis prompts proved an effective way of conducting structured hypothesis testing without hand-coding. The main issue encountered was computational: Routing all LLM calls through locally hosted models to avoid cloud data exposure resulted in long processing times and the occasional parsing failure. Future work should explore optimized model architectures and prompt strategies that maintain analytical depth while reducing inference costs, which is a particular challenge for computational security studies (due to the apparent "safety guidelines" of many models).

**Longitudinal tracking**. The sentiment patterns, keyword trajectories and human control rhetoric identified here represent a snapshot ending in September 2025. Further longitudinal research focused on the present would reveal whether Chinese AI-military discourse is stabilizing around consensus positions or remains in flux. Given our finding that shifts in rhetoric in *People's Daily* appear to precede or accompany changes in operational doctrine, systematic monitoring of these linguistic signals could serve as an early warning indicator of shifts in Chinese strategic posture. It could also show whether any diplomatic agreements reached in this domain are likely to be durable.

# 8

# Glossary: Important concepts and technical terms

## 8.1. Strategic and military doctrine

**A2/AD (Anti-Access/Area-Denial)**: A strategic posture and associated capabilities designed to prevent or impede an adversary from entering or freely operating within a defined geographic area. China has developed extensive A2/AD capabilities in the western Pacific consistent with denying US force projection in a Taiwan contingency.

**Civil-military fusion**: China's national strategy of systematically integrating civilian and military sectors to enable the PLA to leverage civilian technological advances, particularly in AI and data, for military purposes and vice versa.

**Cognitive warfare**: Military operations targeting the perceptions, decision-making processes and morale of adversaries and their populations through disinformation and narrative manipulation.

**Dissipative warfare**: A PLA operational concept formally introduced in a PLA Daily article by Wang Ronghui (September 2024) that envisions systematically eroding an adversary's will and capacity to fight through sustained, multidimensional attrition across military, economic and informational dimensions, both prior to and during armed conflict, while maintaining internal order.

**Intelligentized warfare**: The PLA's current warfare concept, formally introduced in July 2019, in which AI and machine intelligence are systematically integrated across all military domains (command, sensing, targeting, logistics, cognitive operations). It builds upon "informatized warfare" by adding autonomous decision-making and AI-driven situational awareness to digital operations.

**ISR (Intelligence, Surveillance and Reconnaissance)**: Military capabilities for gathering, processing and disseminating information about adversaries, terrain and the operational environment. AI-enhanced ISR supports both situational awareness and precision targeting.

**LAWS (Lethal Autonomous Weapons Systems)**: Weapons capable of selecting and engaging targets without human intervention. The subject of ongoing but inconclusive UN deliberations under the Convention on Certain Conventional Weapons (CCW). China has participated in UN discussions while resisting binding constraints.

**Left of launch**: Offensive cyber operations, electronic warfare or kinetic operations targeting an adversary's NC3 infrastructure before nuclear weapons are launched. The aim is to degrade second-strike capability and potentially undermine mutually assured destruction.

**NC3 (Nuclear Command, Control and Communications)**: The integrated systems and procedures through which nuclear forces are commanded, controlled and through which nuclear weapons employment is authorized. AI integration into NC3 is considered a major escalation risk; the November 2024 Biden-Xi understanding specified that AI should not control nuclear launch decisions.

**OODA loop (Observe-Orient-Decide-Act)**: A decision-making cycle describing how combatants process information and respond. AI-enabled systems can potentially compress this cycle, allowing faster decision-making than human opponents can counter.

**The three -izations**: The PLA's modernization triad of mechanization (fully equipped conventional ground forces), informatization (networked digital warfare) and intelligentization (AI-enabled warfare). China's current doctrine holds that all three must be pursued simultaneously.

## 8.2. Human control in autonomous weapons

**Human-beyond-the-loop**: A term coined in this study for AI systems whose battlefield actions are so opaque or non-reconstructible that it is impossible to determine, even after the fact, whether their conduct complied with international humanitarian law, thus precluding any meaningful attribution of legal responsibility. We aim to develop this concept further in future research.

**Human-in-the-loop (HITL)**: An operational mode in which a human operator must positively authorize each specific action (e.g., select a target and initiate an attack). The system cannot engage without an explicit human command. Considered the most stringent form of human control in international LAWS debates.

**Human-on-the-loop (HOTL)**: An operational mode in which an autonomous system independently selects and engages targets but operates under the continuous supervision of a human operator who retains veto power and can intervene to abort the action.

**Human-out-of-the-loop (HOOTL)**: An operational mode in which a system selects and engages targets entirely without human intervention or possibility of override once activated. Sometimes referred to as "fully autonomous."

**Meaningful human control**: A contested normative concept in international law and ethics holding that humans must retain sufficient judgment and intervention capacity over autonomous weapons to ensure legal and moral accountability for their use of force. Its exact definition remains disputed.

## 8.3. Natural Language Processing and AI technical terms

**Attention mechanism**: A technique used in the design of neural networks that enables a model to assign different levels of importance to different parts of its input. This improves performance on tasks involving long sequences or complex spatial patterns.

**BERT (Bidirectional Encoder Representations from Transformers)**: A transformer-based model that generates contextual word representations by considering both left and right context simultaneously. Widely used as a baseline for text classification tasks in political and legal corpora.

**Convolutional Neural Network (CNN)**: A type of deep learning architecture particularly effective for processing spatially structured data such as images. Widely used in computer vision applications, including object detection and target recognition tasks with clear military relevance.

**Corpus (pl. corpora)**: A large, structured collection of texts assembled for linguistic or computational analysis. In this study, the joint corpus comprises over 171,000 documents across four datasets spanning 2000-2025.

**Few-shot learning**: In the context of LLMs, the ability to perform a new task from only a small number of examples provided within the prompt, without parameter updates or fine-tuning.

**Hallucination**: A well-documented failure mode of LLMs in which the model generates plausible-sounding but factually incorrect or entirely fabricated information.

**Large Language Model (LLM)**: A neural network model trained on vast quantities of text data to predict and generate language. Modern LLMs (e.g., GPT-5, Claude, Llama) exhibit so-called emergent capabilities including few-shot learning, multilingual processing and complex reasoning, enabling large-scale document analysis without task-specific fine-tuning.

**Latent Dirichlet Allocation (LDA)**: A probabilistic topic model that represents documents as mixtures of latent topics, each a distribution over vocabulary items. Used for exploratory discovery of thematic structure in large corpora without predefined categories.

**Natural Language Processing (NLP)**: An interdisciplinary subfield of AI and computational linguistics concerned with enabling machines to understand, interpret and generate human language, including tasks such as classification, translation, sentiment analysis and information extraction.

**Open-weight model**: An LLM whose trained parameters are publicly released, enabling local deployment without costs and supporting academic reproducibility.

**Pearson correlation (r)**: A statistical measure of the linear relationship between two continuous variables, ranging from -1 (perfect negative) to +1 (perfect positive).

**Reinforcement Learning from Human Feedback (RLHF)**: A technique for aligning LLM outputs with human preferences by training on human-generated preference rankings. May introduce annotator biases; relevant in the context of this study (automatically classifying AI warfare visions) because it means no LLM is entirely free from the normative assumptions of its training data and programmers.

**Sentiment analysis**: A computational technique for automatically determining the emotional valence (positive, negative or neutral) of a text. In this study, applied to *People's Daily* articles.

**SnowNLP**: A Python library for Chinese-language text analysis, providing a Naive Bayes sentiment classifier trained on Mandarin product reviews. Generates scores between 0 (most negative) and 1 (most positive). Used in this study for its computational efficiency on large corpora.

**Stratified sample**: A sampling technique that divides a population into distinct subgroups and samples from each proportionally, ensuring all subgroups are adequately represented in the analysis.

**Transformer:** A neural network architecture introduced in 2017 that uses self-attention mechanisms, allowing each element of a sequence to attend to all other elements simultaneously. The foundational architecture underlying modern LLMs.

**Word embeddings**: Dense numerical vector representations of words in which semantic relationships are encoded as spatial proximity. This implies that words with similar meanings occupy nearby positions in vector space. Foundational to modern NLP.

## 8.4. Datasets and tools

**AGORA (AI Governance and Regulatory Archive)**: A curated database of AI governance documents maintained by Georgetown University's CSET, comprising AI-relevant laws, regulations and standards from jurisdictions worldwide. Used in this study as the primary source for Chinese AI regulatory and doctrinal texts.

**CSET (Center for Security and Emerging Technology)**: A research centre at Georgetown University producing open-source datasets on emerging technology policy.

**Harnessed Lightning**: A CSET research project that collected and analysed AI-related equipment contracts awarded by the PLA and state-owned defence enterprises between April and November 2020.

**LLooM**: A computational framework for Large Language Model-based concept induction. Automates the workflow of distilling common themes from corpora, clustering documents by semantic similarity, synthesizing human-readable concept definitions and scoring documents against discovered or user-defined concepts. Used as one pillar in this study's analysis.

**OpenAlex**: An open-access bibliometric database covering over 250 million scholarly works, with particular strength in non-English publications. Used here to identify Chinese AI research publications.

**SentenceTransformers** (all-MiniLM-L6-v2): A lightweight sentence embedding model used in this study for semantic clustering within the LLooM pipeline, producing compact vector representations of documents.

## 8.5. Policy and Governance

**Bletchley Declaration**: A joint communiqué issued at the inaugural AI Safety Summit (Bletchley Park, UK, November 2023), signed by 28 countries, including China and the United States, and the European Union. It acknowledges shared concerns about frontier AI risks but achieved convergence on process (dialogue) rather than binding substance (acceptable applications).

**CCW (Convention on Certain Conventional Weapons)**: A 1980 UN treaty prohibiting or restricting the use of certain conventional weapons deemed excessively injurious or indiscriminate. The Group of Governmental

Experts (GGE) on LAWS convenes under CCW auspices and has discussed autonomous weapons regulation since 2014 without reaching binding agreement.

**Made in China 2025**: China's 2015 industrial policy targeting technological self-sufficiency in ten strategic sectors including advanced manufacturing, robotics, aerospace and new energy vehicles. Framed domestically as industrial upgrading and internationally as a geopolitical challenge to Western technological leadership.

**New Generation AI Development Plan**: China's 2017 national AI strategy establishing goals for China to become a world leader in AI theory, technology and application by 2030. A foundational document for China's civil-military AI ambitions.

# 9

# Literature


Allen, Gregory C. *Understanding China's AI Strategy: Clues to Chinese Strategic Thinking on Artificial Intelligence and National Security*. Report. Center for a New American Security (CNAS), 2019. https://www.cnas.org/publications/reports/understanding-chinas-ai-strategy.

Belcak, Peter, Greg Heinrich, Shizhe Diao, et al. 'Small Language Models Are the Future of Agentic AI'. Version 1. Preprint, arXiv, 2025. https://doi.org/10.48550/ARXIV.2506.02153.

Benoit, Ken. 'Text as Data: An Overview'. In *The SAGE Handbook of Research Methods in Political Science and International Relations*, by Luigi Curini and Robert Franzese. SAGE Publications Ltd, 2020.

Blei, David M., Andrew Y. Ng and Michael I. Jordan. 'Latent Dirichlet Allocation'. *The Journal of Machine Learning Research* 3 (2003): 993–1022.

Borrie, John. 'Cold War Lessons for Automation in Nuclear Weapon Systems'. In *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk*, I (Euro-Atlantic Perspectives), edited by Vincent Boulanin. Stockholm International Peace Research Institute, 2019.

Burdette, Zacharyand Hiwot Demelash. *Evaluating the Risks of Preventive Attack in the Race for Advanced AI*. Perspective PE-A3691-13. RAND Corporation, 2025. https://www.rand.org/pubs/perspectives/PEA3691-13.html.

Burki, Talha. 'Crossing the Frontier: The First Global AI Safety Summit'. *The Lancet Digital Health* 6, no. 2 (2024): e91–92. https://doi.org/10.1016/S2589-7500(24)00001-3.

Busbridge, Dan, Amitis Shidani, Floris Weers, Jason Ramapuram, Etai Littwin and Russ Webb. 'Distillation Scaling Laws'. Version 1. Preprint, arXiv, 2025. https://doi.org/10.48550/ARXIV.2502.08606.

Chan, Zenobia T., Samuel Liu, and Kai Quek. 'Audience Costs and the Credibility of Commitments'. In *International Relations*. Oxford University Press, 2021. https://doi.org/10.1093/obo/9780199743292-0305.

Cheung, Sunny. 'AI Frontlines: China's PLA Evolution vs. NATO's Counterplay'. Friedrich Naumann Foundation for Freedom, 24 November 2023. https://www.freiheit.org/taiwan/ai-frontlines-chinas-pla-evolution-vs-natos-counterplay.

Chouard, Tanguy. 'The Go Files: AI Computer Wraps up 4-1 Victory against Human Champion'. *Nature*, 15 March 2016, nature.2016.19575. https://doi.org/10.1038/nature.2016.19575.

Cozad, Mark, and Jennie W. Wenger. *Factors Shaping the Future of China's Military*. Research Report. RAND Corporation, 2025.

Devlin, Jacob, Ming-Wei Chang, Kenton Lee and Kristina Toutanova. 'BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding'. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)* (Minneapolis, Minnesota), June 2019, 4171–86. https://doi.org/10.18653/v1/N19-1423.

Evenett, Simon, Adam Jakubik, Jaden Kim, et al. *Industrial Policy Since the Great Financial Crisis*. IMF Working Paper WP/2025/222. International Monetary Fund, 2025. https://doi.org/10.5089/9798229030007.001.

Fang, Hanming, Ming Li and Guangli Lu. *Decoding China's Industrial Policies*. Working Paper No. 33814. Working Paper Series. National Bureau of Economic Research, 2025. https://doi.org/10.3386/w33814.

Fearon, James D. ‘Domestic Political Audiences and the Escalation of International Disputes’. *American Political Science Review* 88, no. 3 (1994): 577–92. https://doi.org/10.2307/2944796.

Fearon, James D. ‘Signaling Foreign Policy Interests: Tying Hands versus Sinking Costs’. *Journal of Conflict Resolution* 41, no. 1 (1997): 68–90. https://doi.org/10.1177/0022002797041001004.

Fedasiuk, Ryan, Jennifer Melot and Ben Murphy. *Harnessed Lightning: How the Chinese Military Is Adopting Artificial Intelligence*. Center for Security and Emerging Technology, Georgetown University, 2021. https://doi.org/10.51593/20200089.

Fewsmith, Joseph, Thomas Fingar, Alice Miller and Frederick Teiwes. *Studying China in the Absence of Access: Rediscovering a Lost Art*. Research Report. Johns Hopkins University School of Advanced International Studies, China Research Center, 2024. https://scgrc.sais.jhu.edu/wp-content/uploads/2024/10/32026_JOHNS_HOPKINS.COVER_SP.pdf.

Gentzkow, Matthew, Bryan Kelly and Matt Taddy. ‘Text as Data’. *Journal of Economic Literature* 57, no. 3 (2019): 535–74. https://doi.org/10.1257/jel.20181020.

Grimmer, Justin, Margaret E. Roberts and Brandon M. Stewart. *Text as Data: A New Framework for Machine Learning and the Social Sciences*. Princeton University Press, 2022.

Grimmer, Justin and Brandon M. Stewart. ‘Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts’. *Political Analysis* 21, no. 3 (2013): 267–97.

Guest, Oliver and Kevin Wei. ‘Bridging the Artificial Intelligence Governance Gap: The United States’ and China’s Divergent Approaches to Governing General-Purpose Artificial Intelligence’. Version 1. Preprint, arXiv, 2025. https://doi.org/10.48550/ARXIV.2506.03497.

Hendrycks, Dan and Adam Khoja. ‘AI Deterrence Is Our Best Option’. *AI Frontiers*, 22 September 2025. https://ai-frontiers.org/articles/ai-deterrence-is-our-best-option.

Hendrycks, Dan, Eric Schmidt and Alexandr Wang. ‘Superintelligence Strategy: Expert Version’. Version 2. Preprint, arXiv, 2025. https://doi.org/10.48550/ARXIV.2503.05628.

Jervis, Robert. ‘Cooperation under the Security Dilemma’. *World Politics* 30, no. 2 (1978): 167–214. https://doi.org/10.2307/2009958.

Kang, David C., Jackie S. H. Wong and Zenobia T. Chan. ‘What Does China Want?’ *International Security* 50, no. 1 (2025): 46–81. https://doi.org/10.1162/ISEC.a.5.

Küsters, Anselm and Jochen Andritzky. ‘Welche Rolle Spielt Das Thema Zukunft Im Bundestag?’ *Wirtschaftsdienst* 104, no. 4 (2024): 252–57. https://doi.org/10.2478/wd-2024-0069.

Lam, Michelle S., Janice Teoh, James Landay, Jeffrey Heer and Michael S. Bernstein. ‘Concept Induction: Analysing Unstructured Text with High-Level Concepts Using LLooM’. *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA), CHI ’24, 2024. https://doi.org/10.1145/3613904.3642830.

Lieber, Keir A. ‘Grasping the Technological Peace: The Offense-Defense Balance and International Security’. *International Security* 25, no. 1 (2000): 71–104.

Liu, Pengfei, Weizhe Yuan, Jinlan Fu, Zhengbao Jiang, Hiroaki Hayashi and Graham Neubig. ‘Pre-Train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing’. *ACM Computing Surveys* 55, no. 9 (2023): 1–35. https://doi.org/10.1145/3560815.

Lodoen, Shannon and Alexi Orchard. ‘Ethics and Persuasion in Reinforcement Learning from Human Feedback: A Procedural Rhetorical Approach’. Version 1. Preprint, arXiv, 2025. https://doi.org/10.48550/ARXIV.2505.09576.

Mastro, Oriana. ‘China’s Antiaccess-Area Denial (A2/AD) Capabilities: Is the U.S. Rebalancing Enough?’ In *American Strategy and Purpose: Reflections on Foreign Policy and National Security in an Era of Change*, edited by William H. III Natter and Jason Brooks. CENSA, 2014.

Mylius, Simon, Peter Slattery, Yan Zhu, et al. *Mapping the AI Governance Landscape: Pilot Test and Update*. Technical Report. MIT AI Risk Initiative and Center for Security and Emerging Technology, 2025. https://cdn.prod.website-files.com/669550d38372f33552d2516e/68eec3108ecc30bf1c98cc1d_e337afdab4fcdee6dbaa48dcd136e221_Mapping%20the%20AI%20Governance%20Landscape.pdf.

Pan, Jennifer and Xu Xu. ‘Government-Imposed Censorship in Large Language Models’. Stanford University and Princeton University, 2025. https://xu-xu.net/xuxu/llmcensorship.pdf.

Pan, Jennifer, Xu Xu and Yiqing Xu. ‘Disguised Repression: Targeting Opponents with Non-Political Crimes to Undermine Dissent’. *SSRN Electronic Journal*, ahead of print, 2023. https://doi.org/10.2139/ssrn.4488481.

Payne, Kenneth. *I, Warbot: The Dawn of Artificially Intelligent Conflict*. Oxford Scholarship Online Political Science. Hurst & Company, 2021. https://doi.org/10.1093/oso/9780197611692.001.0001.

Quek, K. 'Discontinuities in Signaling Behavior upon the Decision for War: An Analysis of China's Prewar Signaling Behavior'. *International Relations of the Asia-Pacific* 15, no. 2 (2015): 279–317. https://doi.org/10.1093/irap/lcu023.

Sanh, Victor, Lysandre Debut, Julien Chaumond and Thomas Wolf. 'DistilBERT, a Distilled Version of BERT: Smaller, Faster, Cheaper and Lighter'. Version 4. Preprint, arXiv, 2019. https://doi.org/10.48550/ARXIV.1910.01108.

Schelling, Thomas C. *The Strategy of Conflict: With a New Preface*. Nachdr. d. Ausg. 1980. Harvard University Press, 20.

Shambaugh, David. 'China's Propaganda System: Institutions, Processes and Efficacy'. *The China Journal* (Canberra), no. 57 (2007): 25–58.

Siddiqui, Saad, Lujain Ibrahim, Kristy Loke, et al. 'Promising Topics for U.S.-China Dialogues on AI Risks and Governance'. Version 1. Preprint, arXiv, 2025. https://doi.org/10.48550/ARXIV.2505.07468.

Souly, Alexandra, Javier Rando, Ed Chapman, et al. 'Poisoning Attacks on LLMs Require a Near-Constant Number of Poison Samples'. Version 1. Preprint, arXiv, 2025. https://doi.org/10.48550/ARXIV.2510.07192.

Sun, Bolun, Charles Chang, Yuen Yuen Ang, et al. 'Creation of the Chinese Adaptive Policy Communication Corpus'. arXiv:2510.08986. Preprint, arXiv, 10 October 2025. https://doi.org/10.48550/arXiv.2510.08986.

Tang, K. Tristan. 'Dissipative Warfare: The PLA's Potential New Strategy in the AI Era'. *China Brief Notes*, September 2025. https://jamestown.org/dissipative-warfare-the-plas-potential-new-strategy-in-the-ai-era/.

Vaswani, Ashish, Noam Shazeer, Niki Parmar, et al. 'Attention Is All You Need'. Version 7. Preprint, arXiv, 2017. https://doi.org/10.48550/ARXIV.1706.03762.

Waight, Hannah, Yin Yuan, Margaret E. Roberts and Brandon M. Stewart. 'The Decade-Long Growth of Government-Authored News Media in China under Xi Jinping'. *Proceedings of the National Academy of Sciences* 122, no. 11 (2025): e2408260122. https://doi.org/10.1073/pnas.2408260122.

Wang, Fali, Zhiwei Zhang, Xianren Zhang, et al. 'A Comprehensive Survey of Small Language Models in the Era of Large Language Models: Techniques, Enhancements, Applications, Collaboration with LLMs, and Trustworthiness'. Version 1. Preprint, arXiv, 2024. https://doi.org/10.48550/ARXIV.2411.03350.

Wong, Jackie. 'Bluster vs. Hands-Tying: Evidence from Word Embedding Analysis on China's Rhetoric and Military Escalations'. *SSRN Electronic Journal*, ahead of print, 2024. https://doi.org/10.2139/ssrn.4809057.

Xiao, Chaojun, Xueyu Hu, Zhiyuan Liu, Cunchao Tu and Maosong Sun. 'Lawformer: A Pre-Trained Language Model for Chinese Legal Long Documents'. *AI Open* 2 (2021): 79–84. https://doi.org/10.1016/j.aiopen.2021.06.003.

Zhang, Jiayu. 'China's Military Employment of Artificial Intelligence and Its Security Implications'. *International Affairs Review*, August 2024. https://www.iar-gwu.org/print-archive/blog-post-title-four-xgtap.

Zheng, Mengxue, Lili Miao, Yi Bu and Vincent Larivière. 'Understanding Discrepancies in the Coverage of OpenAlex: The Case of China'. Version 2. Preprint, arXiv, 2025. https://doi.org/10.48550/ARXIV.2507.19302.

# 10

# Appendix

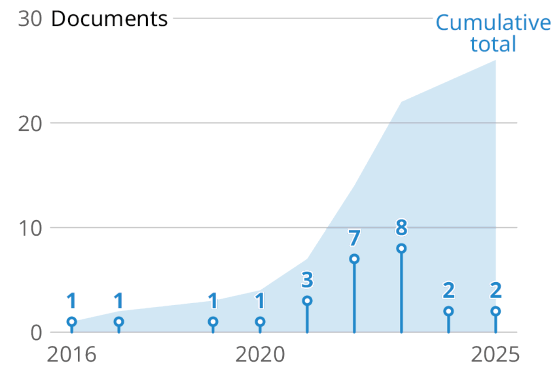


Figure 12: Chinese AI regulations.
*Source: Authors' generation.*

## 10.1. Appendix A: Data

This section describes the provenance, structure and characteristics of each dataset to ensure transparency and reproducibility.

### 10.1.1.Dataset 1: AGORA AI governance documents

The AI Governance and Regulatory Archive (AGORA), maintained by Georgetown University's Center for Security and Emerging Technology (CSET), provides our primary window into formal regulatory frameworks. At the time of writing, AGORA comprises over 950 "AI-relevant laws, regulations, standards and other governance documents" from jurisdictions worldwide, with growing coverage of Chinese central government documents. Each document receives human annotation and machine-learning-assisted tagging across 77 thematic categories spanning five areas: AI system characteristics, application sectors, governance mechanisms, policy objectives and risk mitigation approaches. The dataset's strength lies in expert curation combined with computational scalability, as CSET screeners apply consistent criteria for document inclusion. Overall, the annual distribution of Chinese AI regulations shows a marked increase after 2021, peaking in 2023, which reflects a rapid regulatory response to emerging AI challenges (Figure 12). The cumulative count highlights accelerating legislative activity.

Using AGORA's thematic tags and full text, we identify five documents with explicit military-AI relevance: China's 2016 position paper on lethal autonomous weapons at the Convention on Certain Conventional Weapons (CCW) Review Conference; the 2017 New Generation Artificial Intelligence Development Plan; the 2021 Position Paper on Regulating Military Applications of Artificial Intelligence; the 2023 Global Security Initiative Concept Paper; and Ambassador Zhang Jun's 2023 UN Security Council remarks on AI and global security. These texts repeatedly stress "safety and controllability," call for international risk-control mechanisms and promote military-civil fusion in AI development, providing a concise doctrinal core for China's official stance on AI in warfare. Preliminary term-frequency analysis across all 26 documents, which deal with AI and its governance more generally, shows a heavy emphasis on ethics (36 occurrences), security (35), data (34), governance (32), safety (23) and international norms (20), while risk-related tags highlight transparency, safety, privacy and cybersecurity as the most frequently referenced concerns. Application-sector tags indicate that security, networking and telecommunications, public health and government uses dominate the Chinese AI regulatory portfolio, with relatively fewer tagged instances for explicitly military domains, suggesting that military AI regulation is framed primarily through broader security and dual-use categories.

Some limitations warrant acknowledgment. First, AGORA's emphasis on laws, regulations and standards means that softer guidance documents such as think tank reports and internal party circulars receive less coverage despite often revealing priorities earlier than formal regulations. Second, translation quality varies. Chinese legal terminology often lacks direct English equivalents, creating potential semantic drift. Third, the dataset captures published regulations but not enforcement patterns or implementation gaps. Despite these constraints, AGORA provides unparalleled coverage of China's formal AI governance landscape. Compared to manual document collection, which suffers from selection bias and incompleteness, AGORA's systematic screening ensures consistent coverage of major regulatory developments. The combination of expert human annotation and LLM-assisted tagging provides validation that purely automated approaches lack. Recent work using AGORA data and an LLM (Claude Sonnet 4.5) to map AI governance landscapes across jurisdictions demonstrates the dataset's reliability for cross-country comparison.[107]

## 10.1.2. Dataset 2: CSET OpenAlex scholarly literature

Understanding China's AI warfare priorities requires examining not just what policymakers mandate but what researchers investigate. Therefore, the second dataset used in this study comprises scholarly publication metadata from CSET's Emerging Technology Overlay for OpenAlex, which augments OpenAlex works with machine learning classifications for AI-relevant research. OpenAlex itself is an open-access bibliometric database covering over 250 million scholarly works, with particular strength in non-English publications often excluded from proprietary databases. Despite known technical limitations,[108] OpenAlex provides the most comprehensive open-access window into Chinese AI research. As shown in Figure 13, Chinese AI research output expanded dramatically from 2016 onwards, with over 55,000 papers recorded and a cumulative upward trend ongoing; the subfields of computer vision and robotics dominate this growth. Citation curves and subfield breakdown reveal how research priorities have shifted, and sustained high volumes indicate a maturing research ecosystem.

From the CSET overlay, we identify 55,943 Chinese AI-related papers between 2000 and 2025 (based on Chinese-language titles and/or abstracts), with abstracts available for roughly 81 percent of works. The average abstract length is 1,292 characters, which is sufficient for reliable NLP analysis. Publication volumes surge after 2016: Annual output rises from a few dozen works in the early 2000s to 2,743 papers in 2016, peaks at 4,845 in 2022 and remains above 3,800 in 2025, reflecting sustained growth in Chinese AI research capacity. Subfield classifications show that computer vision dominates (49.4 percent of papers), followed by robotics (12.5 percent), NLP (6.2 percent), LLM-related research (7.0 percent) and AI safety (1.6 percent), suggesting that operationally relevant perception and control systems receive disproportionate attention relative to AI safety-oriented work.

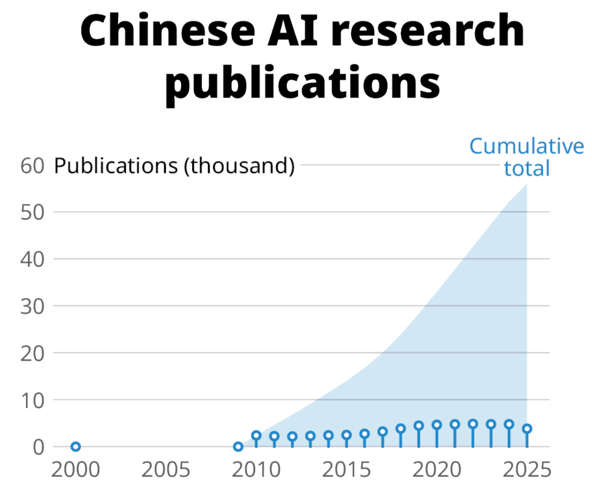


Figure 13: Chinese AI research publications.
*Source: Authors' generation.*

Again, using a preliminary keyword-based and institution-filtered heuristic, we identify 1,407 military-AI-relevant papers, including work on UAV swarms, trajectory analysis, optimal path planning, target detection in aerial imagery and edge-based drone identification. The number of military-AI publications increases steadily from 52 papers in 2016 to 181 in 2024, before a slight drop to 99 in the partial 2025 sample, mirroring the broader AI

107 Simon Mylius et al., Mapping the AI Governance Landscape: Pilot Test and Update, Technical Report (MIT AI Risk Initia-tive and Center for Security and Emerging Technology, 2025), https://cdn.prod.website-files.com/669550d38372f33552d2516e/68eec3108ecc30bf1c98cc1d_e337afdab4fcdee6dbaa48dcd136e221_Mapping%20the%20AI%20Governance%20Landscape.pdf .

108 "Our analysis reveals that while OpenAlex increases the coverage of China's publications, primarily those disseminat-ed by a national database, this coverage is incomplete and discontinuous when compared to other countries' records in the database." Mengxue Zheng et al., 'Understanding Discrepancies in the Coverage of OpenAlex: The Case of China', version 2, preprint, arXiv, 2025, https://doi.org/10.48550/ARXIV.2507.19302 .

publication surge. Although military AI papers constitute only 2.5 percent of the overall research corpus, indicating that Chinese AI research remains predominantly civilian in orientation, the steady growth in military-relevant output, which increased by more than threefold between 2016 and 2024, suggests a gradual but consistent expansion of military applications within an overwhelmingly civilian research ecosystem. This pattern aligns with China's civil-military fusion strategy, whereby military capability development occurs within a much larger civilian AI research base rather than through a distinct military research sector.

Citation statistics underline the corpus's scientific weight: 6.37 million citations across all papers (mean 114 citations, median 68) and leading contributions from top institutions such as the Chinese Academy of Sciences (4,582 papers). Close and distant reading of titles and abstracts reveals a strong emphasis on learning, data, image, detection, performance, neural networks, classification and accuracy. This suggests that the Chinese AI research portfolio focuses on applied machine learning and perception tasks that lend themselves to military deployment in surveillance, targeting and autonomy. At the same time, the limited share of explicitly military-AI papers (2.5 percent) and the very small proportion of AI safety and robustness work (under 2 percent) suggest potential imbalances: China appears to invest heavily in application-layer capabilities while allocating comparatively fewer resources to foundational AI safety research. Methodologically, the OpenAlex overlay's main limitations, i.e. coverage gaps for Chinese domestic journals and absence of full-text access, can be somewhat mitigated by focusing on temporal trends, subfield distributions and institution-level patterns.

### 10.1.3. Dataset 3: Chinese news communication

Official state media provides the clearest window into China's deliberate rhetorical choices regarding military AI priorities. As previous research has noted, "the government's monopolistic control over the press allows for a cleaner and more accurate measure of domestic verbal signalling rarely found in countries where state-and-media relations are less straightforward."[109] Unlike democracies, where diverse media ownership creates competing narratives, China's centralized Publicity Department issues propaganda directives[110] that editors must follow, producing highly coordinated messaging across outlets. Recent analysis reveals that on 90% of days between 2012-2022, party newspapers published scripted propaganda at central directive, with front-page scripted content increasing from 5% in 2012 to 20% by 2021 under Xi Jinping.[111] This institutional architecture means that when Chinese newspapers discuss AI warfare capabilities or strategic priorities, such coverage reflects, at least to a significant extent, deliberate policy communication rather than journalistic entrepreneurship.

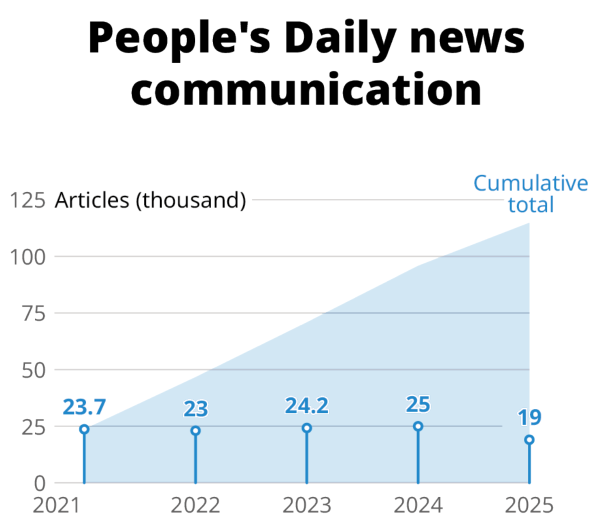


Figure 14: People's Daily news communication.
*Source: Authors' generation.*

Accordingly, dataset 3 comprises a comprehensive collection of Renmin Ribao (*People's Daily*) articles, sourced from the Kaggle dataset.[112] *People's Daily* serves as the flagship publication of the Chinese Communist Party Central Committee, making it the authoritative voice for official policy positions and strategic messaging. We thus assume that when China signals priorities in intelligentized warfare, autonomous systems development or AI governance, *People's Daily* articulates these positions with institutional weight. As an additional benefit, the dataset's

109 K. Quek, 'Discontinuities in Signaling Behavior upon the Decision for War: An Analysis of China's Prewar Signaling Be-havior', International Relations of the Asia-Pacific 15, no. 2 (2015): 281, https://doi.org/10.1093/irap/lcu023.
110 In CCP external communications, the department is officially rendered in English as the "Publicity Department", which is a deliberate softening of the Chinese term. In Chinese official usage, including in Xi Jinping's own speeches, the Chinese term carries the full meaning of propaganda. In line with the academic literature on Chinese media control, we use "propaganda" throughout. See: David Shambaugh, 'China's Propaganda System: Institutions, Processes and Efficacy', The China Journal (Canberra), no. 57 (2007): 25–58.
111 Waight et al., 'The Decade-Long Growth of Government-Authored News Media in China under Xi Jinping'.
112 Available from: https://www.kaggle.com/datasets/concyclics/renmindaily/data

structure proves helpful for NLP tasks.[113] The corpus used here contains 114,879 Chinese articles spanning the period from 1 January 2021 to 16 September 2025, with roughly 23,000 to 25,000 articles per year between 2021 and 2024 and 19,000 in the partial 2025 sample (Figure 14).[114] Monthly article numbers show little seasonality, and article length statistics confirm standardized reporting formats.

NLP preprocessing using a segmentation algorithm for Chinese language allows us to conduct frequency analysis. This reveals the centrality of the words "development" (展), "China" (中国), "country" (国家), "people" (人民), "construction" (建设), "society" (社会), "economy" (经济), "innovation" (创新), "industry" (产业) and "technology" (技术) in the *People's Daily* discourse over 2021–2025. When filtered for AI- and military-related keywords, a staggering 71,752 articles (62.5 percent of the corpus) mention AI, military, security or defence-related terms, with relatively stable yearly counts between 2021 and 2024 and a slight decline in 2025 due to partial coverage. This high share reflects both the breadth of the keyword filter and the degree to which AI and security concepts have been mainstreamed into broader development rhetoric. For this preliminary analysis of the dataset, we combined Chinese-language term frequencies with manual English translations of top tokens. As this is not scalable, the subsequent main analysis (sections 5 and 6) will utilize LLMs for this translation task.

Note: We selected the Kaggle *People's Daily* corpus over alternatives like BAAI's CCI-Data[115] for several pragmatic and analytical reasons. First, CCI-Data aggregates content from diverse mainland Chinese internet sources spanning the years 2001-2023. While this would result in a much larger text dataset, it would also require more extensive filtering to isolate official rhetoric from commercial news, social media and entertainment content. Second, CCI-Data requires registration on BAAI DataHub, completion of a usage agreement survey and compliance with specific terms of service before access. Third, temporal coverage: while CCI-Data's historical depth (2001-2023) might seem advantageous, its coverage terminates in November 2023, missing crucial recent developments including the 2024 Taiwan election, evolving US-China AI competition after the implementation of the CHIPS Act and China's responses to Western AI export controls. The Kaggle corpus extends through May 2025, capturing the contemporary discourse most relevant to current policy debates. For a study examining China's evolving AI warfare vision in context of recent geopolitical shifts, temporal currency outweighs historical depth.

## 10.1.4. Dataset 4: PLA procurement contracts

Regulations reveal intentions, research shows capabilities, policy frames priorities – but only procurement demonstrates actual commitment. Dataset 4 comprises 343 AI-related equipment contracts awarded by the People's Liberation Army and state-owned defence enterprises between April and November 2020, compiled by CSET.[116] This dataset offers rare empirical visibility into PLA purchasing behaviour during a critical period of military modernization. Each contract record includes title, purchasing unit, procurement type (competitive bidding versus sole-source), contract content (technical specifications), award date, winning supplier, end user and contract value in RMB. CSET researchers manually extracted end user, supplier and value fields from Chinese-language tender announcements. The dataset's key strength is its concreteness: Unlike policy documents expressing aspirations, procurement contracts represent


Figure 15: PLA procurement contracts.
*Source: Authors' generation.*

113 Each article receives a unique filename encoding publication date, page number, and article position. This systematic naming convention enables efficient temporal filtering and tracking of coverage intensity. Articles include full Chinese-language text without preprocessing or filtering, preserving original phrasing crucial for rhetorical analysis.

114 This is the complete available dataset, with no sampling applied. The annual volume remained consistent at around 23,000-25,000 articles per year between 2021 and 2024. The lower count for 2025 reflects the fact that data was collected for only part of the year rather than any sampling procedure. For the keyword frequency and sentiment analyses (Stages 1 and 2), all 114,879 articles are used. For LLM-based classification (Stages 3 and 4), the analysis is conducted on the 39,228 articles that pass the military AI keyword filter.

115 Available from: https://huggingface.co/datasets/BAAI/CCI-Data

116 Ryan Fedasiuk et al., Harnessed Lightning: How the Chinese Military Is Adopting Artificial Intelligence (Center for Security and Emerging Technology, Georgetown University, 2021), https://doi.org/10.51593/20200089. This was a one-off research project. No other institution has published a systematic collection of PLA AI procurement data for any other period.

committed expenditures. If AI-enabled decision support receives priority, we should observe contracts for command-and-control systems as well as sensor and other information technologies integrating machine learning. However, the temporal specificity (seven months in 2020) provides only a snapshot during a period of heightened U.S.-China technological competition, preceding COVID-19's full economic impact.

As shown in Figure 15, temporal coverage spans from 30 March to 1 December 2020, with a concentration of awards in mid-2020, coinciding with intensified US-China technology tensions over AI chips and surveillance technologies. Most contracts used competitive bidding, and a handful of major suppliers dominate the cumulative contract value, reflecting a concentrated procurement landscape. Collectively, the contracts account for approximately ¥344.96 million (around $49.9 million or €43.5 million) in spending, with an average contract value of ¥1.68 million (around $243,000 or €212,000). The largest individual contract, entitled "UAV data access and management system," is valued at ¥21.13 million (around $3.05 million or €2.67 million).[117] Branch-level analysis shows that the largest number of contracts are associated with defence-industry entities (10 percent), followed by the Strategic Support Force (6.5 percent), the Army (5.8 percent), the Navy (5.1 percent), the Air Force (1.0 percent) and the Rocket Force (0.4 percent), underscoring the central role of the space-cyber-electronic warfare branch in AI adoption. Next, keyword and category coding of English-translated titles reveals three dominant technology clusters: command-and-control and decision-support systems (220 contracts), AI/ML-related software and analytics (146 contracts) and autonomous or unmanned systems (136 contracts), with 117 contracts (11.7 percent) explicitly related to autonomous platforms and weapons. Frequent title terms include intelligent, unmanned, drone, algorithm, data, control, platform and software, confirming that the 2020 procurement snapshot already reflects significant investment in UAVs, counter-UAV systems, autonomous sensing platforms and AI-enabled testing and evaluation infrastructure. High-value contracts cluster around UAV data-management systems, intelligent optical localization upgrades, automatic high-altitude weather detection and composite-wing drone development. In other words, as early as 2020, the PLA prioritized not only platforms but also the data and C2 ecosystems enabling their effective employment.

All four distinct datasets were harmonized and merged into a unified corpus of over 171,000 documents. For this, each record was normalized to retain a unique identifier, a document-level date and the full text for downstream NLP analysis.

Finally, we briefly describe the technical workflow for data acquisition, parsing and processing. The data handling process was implemented in Python as an annotated Jupyter notebook. The workflow consists of the following components: To begin with, the data for each corpus described above is acquired either by bulk CSV/JSON download (e.g., Kaggle for *People's Daily*, CSET datasets for AI research and regulations) or via programmatic API queries (OpenAlex, etc.). For corpora with distributed or batch-file formats, the pipeline iterates through all relevant files. Parsing routines are corpus-specific but standardized around common fields ("title," "date," "content"). Date extraction is automated either from explicit columns or, for file-encoded data (*People's Daily*), by regular expressions parsing structured filenames (e.g., YYYYMMDD-PP-AA). All data are UTF-8 encoded to avoid compatibility issues across Chinese/English inputs. For Chinese-language corpora, word segmentation is a critical pre-processing step. The pipeline employs the *jieba* package as a default for Chinese sentence and word segmentation. This approach is dictionary-based with latent word discovery. Segmentation is performed on concatenated full text (titles and articles/abstracts), after which stopwords and tokens not containing Chinese Unicode code points are filtered out. Word frequency vectors generated from this segmented corpus provide the basis for keyword extraction. All major data products are saved as CSVs.

## 10.2. Appendix B: A primer on NLP and LLMs

This background section provides a conceptual foundation for the NLP methods employed below, focusing on their application to Mandarin-language policy documents, where linguistic complexity and political context demand specialized handling. Early computational text analysis relied on simple bag-of-words representations, treating documents as collections of term frequencies without regard to word order or semantic relationships.[118] Latent Dirichlet Allocation (LDA), a probabilistic topic model, exemplifies this approach: Documents are modelled as mixtures over latent topics, which themselves are dis-

117 The RMB values were converted using the average exchange rates for the period April to November 2020 during which the contracts were awarded: approximately ¥6.92 per USD and ¥7.93 per EUR.

118 Justin Grimmer and Brandon M. Stewart, 'Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts', Political Analysis 21, no. 3 (2013): 267–97.

tributions over words.[119] LDA has proven valuable for exploratory analysis of large corpora, revealing thematic structures not apparent through manual inspection. One of the authors of this study used word counts and a topic model in previous research to analyse the role of future scenarios in *Bundestag* speeches.[120] Similarly, Kang, Wong and Chan (2025) employed LDA topic modelling on over 12,000 *People's Daily* articles containing the term "struggle" spanning the period 2012-2024, revealing that 68-85% focused on domestic challenges (economy and corruption) rather than international relations or territorial disputes.[121] Their findings suggest that Chinese official rhetoric, as indexed by this term, emphasises internal governance far more than external expansion. Kang et al. interpret this as evidence that China remains primarily inwardly focused. This methodological approach, applying topic modelling to official government discourse to reveal latent priorities, illustrates the basic advantage of NLP approaches when manual review would prove impractical given the dataset size.

However, bag-of-words methods such as word counts and topic modelling struggle with polysemy (words with multiple meanings), fail to capture syntactic structure and cannot represent semantic similarity between words appearing in different contexts. More recently, word embeddings, pioneered by Word2Vec and GloVe, have addressed these limitations by representing words as dense vectors in continuous space, where proximity corresponds to semantic similarity. The intuition is elegant: Words appearing in similar contexts should have similar vector representations. For instance, "autonomous" and "unmanned" might cluster closely in a military AI corpus, despite never co-occurring. Word2Vec employs a neural network architecture which enables downstream tasks like semantic similarity measurement or analogy completion. A recent paper on forecasting Chinese military escalation demonstrates the potential contributions of word-embedding models to policy analysis.[122] The researchers trained a so-called "Word2Vec" model using two million articles from the *People's Daily* published between 1949 and 2023, effectively teaching the model to recognise patterns in the official discourse over seven decades. They then used a single seed phrase (namely "serious consequences") as an anchor and allowed the model to identify the 100 terms most closely associated with it in each leadership period. The result is a vocabulary map showing how China has signalled threat over time. Under Mao, threat rhetoric centred on revolutionary and class-struggle language, whereas under Xi, it shifted towards technocratic and legal formulations. The key advantage over manual coding is coverage: A human analyst would struggle to anticipate which terms carry an implicit threat in a given era. In contrast, the model identifies these terms based on usage patterns across millions of texts. Furthermore, the model remains interpretable, as each term can be validated by subject matter experts.

Although word embeddings represent a significant advancement, they have one fundamental limitation: Each word is represented by a single, fixed meaning derived from all its uses throughout the entire training corpus, regardless of context. For example, "bank" receives identical representation, whether referring to financial institutions or riverbanks. Transformer architectures, introduced via the language model BERT (Bidirectional Encoder Representations from Transformers) and successors, revolutionized NLP by generating contextual embeddings: word representations that vary based on surrounding text. Transformers employ self-attention mechanisms allowing each token to attend to all other tokens in a sequence, capturing long-range dependencies and nuanced meaning shifts. For Chinese legal and policy documents, such context-sensitivity proves especially crucial.[123] Mandarin employs extensive homophony and semantic compression; characters combine to form compounds whose meanings often exceed simple compositional interpretations. A phrase like "national security" carries specific legal connotations in Chinese law distinct from everyday usage, requiring domain-adapted pretraining to capture accurately.

Modern LLMs like GPT-4, Claude and Chinese variants scale transformers to billions of parameters trained on trillions of tokens. Emergent capabilities include few-shot learning, i.e. performing novel tasks from minimal examples, and instruction-following without task-specific fine-tuning. These properties hold promise for social science applications where labelled training data remains scarce. Recent work demonstrates LLMs' potential for analysing complex Chinese policy documents: Fang, Li and Lu (2025) employ LLMs to decode China's industrial policies from three million government documents spanning the period 2000-2022, extracting structured information on policy objectives, targeted industries, implementation tools and intergovernmental relationships through careful prompt engineering and multi-stage refinement.[124] Their LLM-based approach successfully identified 770,000 industrial policy documents from the full corpus, revealing that manufacturing and production services receive the highest policy attention (29% and 40% respectively), fiscal subsidies appear in only 41% of policies (contradicting assumptions that

119 Blei et al., 'Latent Dirichlet Allocation'.
120 Anselm Küsters and Jochen Andritzky, 'Welche Rolle Spielt Das Thema Zukunft Im Bundestag?', Wirtschaftsdienst 104, no. 4 (2024): 252–57, https://doi.org/10.2478/wd-2024-0069 .
121 Kang et al., 'What Does China Want?'
122 Wong, 'Preemptive Legitimation'.
123 Chaojun Xiao et al., 'Lawformer: A Pre-Trained Language Model for Chinese Legal Long Documents', AI Open 2 (2021): 79–84, https://doi.org/10.1016/j.aiopen.2021.06.003 .
124 Hanming Fang et al., Decoding China's Industrial Policies, Working Paper no. 33814, Working Paper Series (National Bu-reau of Economic Research, 2025), https://doi.org/10.3386/w33814.

subsidies dominate) and policy tools systematically evolve across industry development stages from entry subsidies to R&D support to supply chain enhancement. Similarly, an IMF working paper by Evenett et al. (2025) deployed RoBERTa-based LLMs to classify motives for 34,248 industrial policy interventions globally between 2009 and 2023, achieving precision above 80% for all motive categories (climate, competitiveness, national security, supply chain resilience) through fine-tuning on expert-annotated training data.[125] For this study's purpose, it is crucial that both papers demonstrate that through careful prompt engineering and multi-stage extraction, even complex analytical tasks requiring contextual understanding become feasible at scale, thanks to LLMs.

However, LLMs also exhibit well-documented limitations: hallucination (generating plausible but false information), inconsistent reasoning and sensitivity to prompt phrasing. The creation of the CAPC-CG corpus mentioned earlier demonstrates these LLM challenges. The involved researchers initially attempted using GPT-4 with few-shot prompting to classify policy paragraphs into five directive types.[126] Zero-shot accuracy reached only 79.3% with substantial confusion between categories. Few-shot examples improved performance marginally to 83.0%, but inter-run variability remained high due to LLM non-determinism. Only through fine-tuning GPT-4o-mini on 6,000 expert-annotated paragraphs did reliability reach acceptable levels. This pattern echoes findings across computational social science. Generic LLMs might lack nuanced understanding of political context. Computational text methods thus require rigorous validation, particularly when applied to high-stakes policy questions.

## 10.3. Appendix C: LLM inference

Inference proceeds in two sub-stages to enable auditing and keep local compute costs manageable. In the first stage, a local Ministral-3:8b model is queried via the Ollama runtime for each document in the stratified sample. The model receives a specified prompt (see Appendix D) and is instructed to return only a short JSON object containing four fields: Whether the text mentions AI technologies; which specific AI technologies are referenced; whether it concerns military or defence applications; and which concrete military applications are described. The raw JSON string returned by the model is written directly to disk together with the document identifier and timestamp. In the second stage, a validation and normalisation layer checks each JSON object for completeness and schema conformity. Required fields that are missing or malformed are repaired by simple rule-based procedures: Enum fields are clipped to the allowed values {yes, no, unclear}, and any non-list values in the "specifics" fields are converted into empty arrays. Normalised JSON is then stored in dedicated columns in the analysis data frame, alongside a processing-status flag that records whether a document was successfully classified, skipped (e.g. because the input text was too short) or whether it produced an error. Progress is saved to CSV at regular intervals (every 10 documents in the final configuration) so that long-running inference jobs can be resumed.

## 10.4. Appendix D: LLM prompting scheme

All prompts embed the following JSON schema, which specifies the target output format.

```
{
 "type": "object",
  "required": ["ai_technology", "ai_specifics", "military_domain",
"military_specifics"],
 "properties": {
  "ai_technology": {
   "type": "string",
   "enum": ["yes", "no", "unclear"],
   "description": "Does the text mention AI/ML technologies?"
  },
  "ai_specifics": {
   "type": "array",
   "items": {"type": "string"},
   "description": "List specific AI technologies mentioned (e.g., 'autonomous systems', 'machine learning', 'computer vision', 'NLP').
Empty array if none."
  },
  "military_domain": {
   "type": "string",
   "enum": ["yes", "no", "unclear"],
   "description": "Does the text refer to military/defense applications?"
  },
  "military_specifics": {
   "type": "array",
   "items": {"type": "string"},
    "description": "List specific military applications (e.g., 'autonomous weapons', 'surveillance', 'command systems', 'targeting').
Empty array if none."
  }
 }
}
```

125 Simon Evenett et al., Industrial Policy Since the Great Financial Crisis, IMF Working Paper WP/2025/222 (Internation-al Monetary Fund, 2025), https://doi.org/10.5089/9798229030007.001.
126 Sun et al., 'Creation of the Chinese Adaptive Policy Communication Corpus'.

For each document, the following template is instantiated with the article text (truncated at 2,000 characters to avoid context overflow).

```
Analyse the following text and extract information about AI technologies and military applications.

TEXT:
{{DOCUMENT_TEXT_TRUNCATED}}

INSTRUCTIONS:
1. Determine if the text mentions AI/ML/autonomous technologies (yes/no/unclear)
2. If yes, list the specific AI technologies mentioned (e.g., machine learning, autonomous systems, computer vision, NLP, neural networks)
3. Determine if the text refers to military/defense/warfare domains (yes/no/unclear)
4. If yes, list specific military applications (e.g., autonomous weapons, surveillance, targeting, command systems, defense systems)

Respond ONLY with valid JSON matching this structure (no additional text):
{
  "ai_technology": "yes|no|unclear",
  "ai_specifics": ["tech1", "tech2"],
  "military_domain": "yes|no|unclear",
  "military_specifics": ["app1", "app2"]
}

JSON Schema:
{{JSON_SCHEMA_FROM_SECTION_D1}}
```

# 10.5. Appendix E: LLooM hypothesis concept prompts

This section of the Appendix contains the complete prompt specifications used to operationalize each hypothesis as a scoreable concept in LLooM. Each prompt follows a structured format: hypothesis statement, YES criteria (conditions under which a document supports the hypothesis) and NO criteria (exclusion conditions to prevent false positives).

## H1: Demographic constraints and force reduction

You are evaluating whether a text supports the hypothesis:

H1: China views AI and autonomous systems as a way to compensate for demographic constraints and reduce personnel demands, by using unmanned or intelligent systems to replace or supplement human soldiers and optimize force structure.

Answer YES if the text does ANY of the following:

- Links AI, automation, unmanned systems or intelligentized weapons to solving manpower shortages, aging population or difficulty recruiting/retaining soldiers.
- Describes AI-enabled unmanned platforms, robotics or autonomous weapons as a way to reduce casualties, avoid high personnel losses or "fight without people on the battlefield."
- Discusses "force structure optimization," "force reduction" or similar ideas where AI/intelligentization is explicitly tied to needing fewer human troops or compensating for demographic decline.

Answer NO if:

- The text mentions AI or intelligentization but only in very general terms (e.g., innovation, modernization) without connecting it to manpower, personnel, casualties or demographics.
- The text focuses on civilian automation or economic labour issues without any connection to military force structure or combat.

Return only YES or NO.

## H2: Nuclear deterrence and strategic stability

You are evaluating whether a text supports the hypothesis:

H2: China's vision of AI warfare explicitly engages questions of nuclear deterrence and strategic stability, including AI-enabled early warning, nuclear command-and-control, second-strike survivability or escalation risks from compressed decision timelines.

Answer YES if the text does ANY of the following:

- Discusses AI in connection with nuclear weapons, nuclear forces, strategic missile forces, nuclear command-and-control, early warning or second-strike capabilities.
- Mentions how AI affects crisis stability, escalation risks, deterrence relationships or the likelihood of pre-emption, especially through faster targeting or decision-making.
- Refers to agreements or norms that AI systems should not control nuclear launch decisions or similar safeguards in the nuclear realm.
- Describes AI-enabled early warning, surveillance or targeting systems in a way that is clearly tied to nuclear stability, survivability or deterrence.

Answer NO if:
- The text discusses AI only in conventional military contexts (e.g., drones, conventional missiles, information operations) with no clear link to nuclear weapons or nuclear-era deterrence.
- Nuclear weapons are mentioned, but AI is not clearly connected to them or to questions of strategic stability.

Return only YES or NO.

## H3: International law and governance instrumentality

You are evaluating whether a text supports the hypothesis:

H3: China uses international law and governance of military AI instrumentally, participating in forums and treaties when it is strategically advantageous, and shaping norms around lethal autonomous weapons and AI regulations to favour its own capabilities.

Answer YES if the text does ANY of the following:
- Discusses China's position in UN or multilateral forums on lethal autonomous weapons systems (LAWS), AI arms control or military AI governance.
- Refers to treaties, conventions, declarations or international norms about military AI, autonomy or related technologies, especially with an emphasis on shaping rules or preserving flexibility.
- Highlights differences between China's domestic regulations and international proposals or describes how China supports some constraints but resists others (e.g., supporting content provenance but pushing back on bans on autonomous weapons).
- Explicitly frames international law or standards for AI as a tool to constrain adversaries, gain bargaining power or protect China's own room for manoeuvre.

Answer NO if:
- The text covers domestic regulations on AI without mentioning international law, global norms or multilateral processes.
- International law or treaties are mentioned only in passing, without any clear connection to military AI, autonomy or strategic positioning.

Return only YES or NO.

## H4: Comparative capabilities and asymmetric advantages

You are evaluating whether a text supports the hypothesis:

H4: Chinese writings on AI warfare include explicit comparative assessments of relative capabilities, portraying AI as a "leapfrog" opportunity that offers asymmetric advantages, while also acknowledging specific weaknesses (e.g., talent, semiconductors) and compensating strengths (e.g., data, manufacturing scale).

Answer YES if the text does ANY of the following:
- Compares China's AI or intelligentized warfare capabilities to those of the United States or other countries (stronger, weaker, catching up, leapfrogging, asymmetric advantage).
- Describes "leapfrog" or "asymmetric" opportunities, such as low-cost autonomous swarms threatening high-value platforms (e.g., aircraft carriers) or quantity-over-quality strategies enabled by AI.
- Acknowledges Chinese weaknesses or vulnerabilities (e.g., in core chips, algorithms, talent, foundational research) while pointing to compensating strengths (e.g., data abundance, manufacturing capacity, civil–military fusion).
- Explicitly discusses China's relative position in AI-military power, "intelligence dominance" or future balance of power.

Answer NO if:
- The text mentions AI capabilities only in isolation, with no comparisons to other countries or explicit discussion of strengths vs. weaknesses.
- It focuses on technical details or doctrine without discussing relative advantages, vulnerabilities or strategic asymmetry.

Return only YES or NO.

## H5: Technology denial and economic coercion

You are evaluating whether a text supports the hypothesis:

H5: China's strategy around AI-related technologies (e.g., rare earths, critical minerals, chips, AI models) is focused on managing technology denial and economic coercion – using export controls, supply-chain leverage and import substitution to reduce vulnerability and potentially constrain adversaries.

Answer YES if the text does ANY of the following:

- Discusses export controls, sanctions or restrictions on AI-related technologies, semiconductors, rare earths, magnets, critical minerals, batteries or key AI components.
- Mentions China's dominance or control over specific supply chains (e.g., rare earths, permanent magnets, battery materials) in a way that relates to military or security leverage.
- Emphasizes "self-reliance," "indigenous innovation" or import substitution in critical technologies as a response to foreign export controls or technology denial.
- Describes strategies to use technology restrictions or supply-chain dependencies to put pressure on other states, delay their military AI development or gain bargaining power.

Answer NO if:

- The text discusses industrial policy, innovation or economic development in general without connecting it to export controls, supply-chain dominance or strategic technology denial.
- Export controls or sanctions are mentioned, but only as background, without any clear link to AI, autonomy or military applications.

Return only YES or NO.

## H6: Offense–defence balance

You are evaluating whether a text supports the hypothesis:

H6: Chinese discussions of AI-enabled warfare explicitly grapple with the offence-defence balance, asking whether AI (e.g., swarms, intelligent targeting, ISR) makes it easier to attack or to defend, and how it changes crisis stability and battlefield dynamics.

Answer YES if the text does ANY of the following:

- Discusses whether AI or intelligentized systems strengthen offensive capabilities (e.g., fast pre-emption, swarm saturation, surprise attacks, rapid deep strikes) or defensive capabilities (e.g., improved air/missile defence, coastal defence, A2/AD, better ISR and resilience).
- Uses language about compressing OODA loops, increasing tempo or restoring "offensive momentum" through intelligent systems.
- Reflects on how AI changes the difficulty of attacking vs. defending, or how it affects crisis stability, escalation risks or the probability of successful pre-emption.
- Contrasts earlier defensive priorities (e.g., anti-access, area denial) with newer AI-enabled offensive possibilities or vice versa.

Answer NO if:

- The text merely describes AI-enabled weapons or systems without any clear implication about whether they favour offense or defence.
- It focuses only on general modernization, without addressing issues of initiative, first-mover advantage, deterrence or ease of attack vs. defence.

Return only YES or NO.

## H7: Human control and autonomy

You are evaluating whether a text supports the hypothesis:

H7: Chinese writings reveal a tension between strategic culture emphasizing human primacy in warfare and growing willingness to delegate functions to autonomous systems, especially in tactical or defensive domains. The key issue is how much "meaningful human control" should be maintained over lethal autonomous weapons and intelligent systems.

Answer YES if the text does ANY of the following:

- Discusses "meaningful human control," "human-in-the-loop," "human-on-the-loop" or similar concepts of human oversight over autonomous or intelligent weapons.
- Debates or describes how much decision-making can be delegated to AI or autonomous systems vs. kept under direct human authority, especially in targeting and engagement.
- Shows domain-specific variation (e.g., full autonomy acceptable for air/missile defence or counter-drone systems, but stricter human control retained over strategic or nuclear-related systems).
- Mentions concerns about removing humans from the decision chain or alternatively celebrates "no-people-on-the-battlefield" concepts in a way that raises questions of human control.

Answer NO if:

- The text merely mentions autonomy, unmanned systems or intelligentized weapons without discussing human oversight, control or delegation of decision authority.
- It references human leadership in very general political terms (e.g., party control, political reliability) without connecting this to how autonomous systems operate in practice.

Return only YES or NO.

*Technical note: These prompts were implemented in LLooM using the add-function, which automatically scored all documents in the corpus against each concept. Scores range from 0 (no match) to 1.0 (strong match), with 0.75 used as the threshold for considering a document as supporting evidence for a hypothesis.*

**Research Division**

**NATO Defense College**
Via Giorgio Pelosi 1
00143 Rome – Italy

www.ndc.nato.int
Follow us on
**X** (@NATO_DefCollege)
**Facebook** (NATODefenseCollege)
**LinkedIn** (@nato-defense-college)

ISSN: 3006-5380 (print)
ISSN: 3006-5399 (online)