The Washington Quarterly

ISSN: 0163-660X (Print) 1530-9177 (Online) Journal homepage: www.tandfonline.com/journals/rwaq20

The Cassandra Trap: How Beijing Could Come to Doubt Its Own Digital Oracles

Philip B.K. Potter

To cite this article: Philip B.K. Potter (2026) The Cassandra Trap: How Beijing Could Come to Doubt Its Own Digital Oracles, The Washington Quarterly, 49:2, 135-147, DOI: 10.1080/0163660X.2026.2684818

To link to this article: https://doi.org/10.1080/0163660X.2026.2684818

Published online: 30 Jun 2026.

Philip B.K. Potter

The Cassandra Trap: How Beijing Could Come to Doubt Its Own Digital Oracles

The Chinese Communist Party (CCP) has fully embraced artificial intelligence (AI) as the foundation of its strategies for governance and social control. The plans for AI run the spectrum from mundane to Orwellian. At one end, AI eases traffic congestion and streamlines public services. In the middle, it helps enforce China's social credit system and monitors the performance of government bureaucrats. At the extreme, AI powers predictive policing programs which claim to anticipate social unrest and sustain the mass surveillance and detention regime in Xinjiang.

The prospect of AI-enhanced autocracy is a daunting one for China's democratic competitors. While Chinese investment in AI still pales in comparison to the United States, the People's Republic of China (PRC) is expending significant capital in its pursuit of AI dominance.[1] In summer 2025, the Chinese cities of Hangzhou, Shanghai, Shenzhen, and Chengdu pledged nearly $400 million to subsidize AI investment in their regions.[2] Chinese engineers have proven worthy competitors to Silicon Valley, as evidenced by the shockwaves set off by DeepSeek's advances in January 2025. DeepSeek, a Chinese AI model, demonstrated capabilities comparable to those of American frontier models, but at a fraction of the development time and cost. But for all the handwringing about Chinese

Philip B.K. Potter is the executive director of the National Security Data and Policy Institute (NSDPI) at the University of Virginia, where he is also a professor of public policy and the founding director of the Frank Batten School's National Security Policy Center. He has served as a university expert for the Intelligence Community and a senior advisor for the Department of Defense. He can be reached by email at pbp2s@virginia.edu.

The Washington Quarterly • 49:2 pp. 135–147
https://doi.org/10.1080/0163660X.2026.2684818

advances and investments, artificial intelligence also confronts the CCP with an underappreciated risk that I call the "Cassandra Trap."

China's leaders, including Xi Jinping, have a habit of framing their thorniest dilemmas as classical "traps"—from Graham Allison's Thucydides Trap, which warns of war between rising and established powers, to the Tacitus Trap, which holds that once a regime loses public trust even its honest actions invite suspicion.[3] The Cassandra Trap recalls Cassandra's tragic fate: she was cursed by Apollo so that her prophecies, though true, would never be believed. She foresaw Troy's fall but was powerless to prevent it. For the CCP, the danger lies in its reliance on AI—adopted out of distrust for its citizens and bureaucrats—colliding with the inherent opacity of the systems themselves. Because AI models can never be fully transparent, trust in their outputs is vulnerable to collapse. And once that collapse occurs, leaders have no fallback. A system built to provide omniscience instead risks producing doubt and paralysis. Therein lies the trap.

China's AI-enabled governance becomes brittle in a crisis when trust and adaptability matter most

What makes China's AI-enabled governance look formidable in normal times—its centralization, surveillance, and discipline—becomes brittle in a crisis when trust and adaptability matter most. Take, as a hypothetical, the PRC's integration of AI into real-time military decision-making. In a fast-moving crisis in the South China Sea where success depends on anticipating an adversary's fleet movements to ensure force readiness, loss of trust in these systems could lead to costly delay and potential defeat. Similarly, in a fast-moving social crisis, doubt in the instruments of state control could leave the CCP vulnerable to protest and political disruption.

The crux of the problem is that AI systems cannot be made fully explainable at scale or under crisis conditions, making confidence hard to restore once doubt is introduced. And uncertainty can creep in from many directions in a political system where information is often distorted and institutions rely on control rather than trust. Citizens and bureaucrats can withhold information and subvert reporting out of self-preservation. Adversaries can exacerbate the problem through cyber intrusions, real or even just feinted. The result is a system that projects power but can become unsettled, leaving leaders hesitant when they need their digital oracles most.

This fragility is both caused by and exacerbates the information problems inherent to authoritarian regimes.[4] Because dictators cannot easily tolerate transparency or dissent, the data they receive is often limited and skewed. For the

authoritarian, the promise of AI lies in the potential for it to overcome these informational weaknesses by collapsing disparate signals into centralized decision systems. For example, Beijing has turned to AI-enabled monitoring of satellite imagery and emissions sensors to identify pollution violations that local officials have incentives to conceal.[5] These systems are designed to give Party leaders direct visibility into conditions on the ground without relying on potentially complicit local cadres. Yet, the very attempt to automate trust can deepen the problem. Once data integrity is in doubt—whether through adversarial manipulation, bureaucratic distortion, or simple technical opacity—the system lacks a capacity for redundant, credible information confirmation.

Understanding how this trap forms, and how it might be sprung, requires tracing its logic from first principles to strategic implications. The first portion of this article identifies and defines the three features of the Cassandra Trap: elusive explainability, the unreliability of data in autocratic systems, and the vulnerability to adversarial interference. With those ingredients established, I turn to the implications for US-China competition, examining how the structural differences between the two political systems shape each country's relationship with AI and its failure modes. I then develop four concrete strategies that US policymakers could employ to exploit the vulnerabilities baked into the CCP's AI-dependent governance. The article concludes by arguing that Beijing's most likely approach—treating AI as an omniscient oracle rather than a fallible tool whose outputs demand scrutiny—may ultimately prove more consequential than any gap in hardware or algorithms.

> The CCP's reliance on AI amplifies a risk inherent to autocracy

The Features of the Trap

Unlike other technologies at the heart of US-China competition—solar panels, rare-earth magnets, lithium batteries—artificial intelligence, particularly when based on large language models, is intrinsically social. It is trained on human inputs and relies on human interpretation. To trust a complex model's outputs, decisionmakers must have confidence in the process by which the model reached its conclusion. This is commonly referred to as "AI explainability."[6]

"Explainable," however, is not the same as "explained," and the distinction matters a great deal for trust in systems deployed to monitor at governmental and societal scales. While technical progress is being made on the explainability problem, it remains vexing in the short term and is likely not fully solvable at

scale in critical moments. Even advanced explainability tools are like MRI scans—invaluable for diagnosis but impractical for continuous use. Just as doctors use MRIs to probe invisible ailments layer by layer, engineers rely on tools like attribution maps and counterfactual traces to peer inside a model's reasoning.[7] These methods can provide a targeted diagnosis of a suspicious output but are far too slow and resource-intensive to assess every decision in real time. Moreover, the accuracy and usefulness of outputs depend heavily on the quality of the queries themselves, which require contextual details that users may not possess, may hesitate to record digitally in authoritarian contexts, or may lack the time and incentive to provide.

The only real way to mitigate these problems is through mechanisms that build redundancy and cross-checking into the system. Effective incorporation of AI into governance requires not only confidence in the models but also overlapping channels of verification that allow inconsistencies to be addressed before they cascade. Whether AI enhances governance therefore hinges on institutional context: it can either function as one component of a system of systems which augments human capacity, or as a monitoring tool intended to centralize, monitor, discipline, and control those systems.

Democracies can rely on independent institutions, bureaucratic expertise, free media, and pluralistic debate to perform this function. In China, however, where dissent is often punished and political incentives reward conformity, these safeguards are weaker. The brittleness of China's centralized approach is fertile ground for a Cassandra Trap. Bureaucrats and functionaries routinely withhold concerns, dissenting views, or inconvenient truths out of fear or futility. Silence is reinforced by a political culture which can reward subordinates for echoing leadership's expectations rather than challenging them.[8] In such an environment, even when AI systems produce accurate warnings or insights, those signals may be ignored, downplayed, or filtered before they reach decision-makers. The CCP's reliance on AI thus amplifies a risk inherent to autocracy—without trusted redundancy, any crack in model validity can trigger cascades of distrust which no parallel institution is equipped to correct.

The problems, however, go even deeper than explainability. Uncertainty is magnified when the training data and model designs are themselves potentially poisoned by internal political incentives or outside actors. The same political incentives pushing Beijing to adopt AI so aggressively raise questions about the quality and reliability of the models themselves. CCP leadership doggedly pursues AI systems because it distrusts its own bureaucratic cadres, to say nothing of its own citizens, ethnic minorities, and political dissidents. But the distrust embedded in these relationships can poison the very data that powers AI models. For decades, lower-level officials have falsified performance metrics to appear compliant, while citizens have self-censored and obscured their true

behaviors and preferences—meaning that the models are trained on inputs that are saturated with systematic errors, bias, and misinformation. Academics have found evidence of such "distorted information flows" within the Chinese bureaucracy, where fiscal data has been systematically inflated to meet certain performance metrics.[9] The prompts and data that officials and citizens put into the models that monitor them will presumably have similar biases and strategic omissions.

This pattern of systemic distortion is not hypothetical.[10] In the early stages of the COVID-19 outbreak, Wuhan officials withheld case data from Beijing, fearing political retribution—a decision that cost critical response time.[11] Similar bureaucratic delays occurred during the 1998 Yangtze River floods and the 2008 Sichuan earthquake, when local leaders suppressed warnings to avoid blame.[12] Under pressure, Chinese bureaucrats often prioritize self-preservation over accuracy. If such behavior persists in the era of AI-driven governance (and there is no reason to think that it will not) the CCP risks training and refining its systems based on falsified inputs and receiving correspondingly distorted outputs. This could deepen the very crises the Party wants to predict and control.

Beyond the challenges presented by explainability and biased data, Beijing faces the very real risk that adversaries might deliberately seek to compromise China's AI systems through cyber intrusions. Empirical research in adversarial machine learning underscores how even robust models can be derailed with minor but carefully crafted inputs. Imperceptible changes to text, images, or code can induce incorrect outputs.[13] The risk of data poisoning, the use of compromised training data to subtly bias model behavior, has been demonstrated at scale.[14] Notably, because the Cassandra Trap is about perception and trust rather than reality, breaches need not be successful or even real to erode confidence. Mere suspicion of the integrity of the system can be enough to sow doubt.

Consider a scenario in which an AI-enabled internal security system flags an emerging pattern of coordinated labor unrest across multiple inland transport hubs, based on anomalous mobile phone co-location data and shifts in freight movement. Even if the underlying signal is accurate, a single report, regardless of whether it is true, that the data or model has been corrupted would immediately complicate decision-making. Rather than acting on the alert, senior officials would likely require cross-validation from provincial security bureaus, manual review of raw feeds, and supplemental human reporting before authorizing a response. Each additional layer of verification introduces delay, and in a system conditioned to avoid political blame, delay and obfuscation are the rational responses. The result is not paralysis because the warning is false, but because leaders can no longer be confident that it is true. In a crisis, that loss of confidence can be decisive.

These three elements—elusive explainability, the unreliability of data in autocratic systems, and the risk of adversarial interference—form the necessary ingredients of a Cassandra Trap for the CCP. If the trap springs, Party decisionmakers could swing abruptly from steady-state reliance on AI systems to outright loss of confidence in their outputs. With no trusted fallback, hesitation and paralysis may follow, rendering Beijing disadvantaged in the face of domestic unrest or deterred at a decisive moment, such as a confrontation in the Taiwan Strait.

Implications for US-China Competition

AI derives its power and peril as much from how it is woven into social and governance structures as from the technology itself. Algorithms must be trained on human behavior, institutional records, and political priorities; in turn, their outputs reshape incentives, norms, and authority. This feedback loop means that two nations can develop and deploy the same technology and arrive at radically divergent outcomes. For CCP leaders in particular—who have plainly stated their intent to integrate AI across industry, government, and society—this outcome could be unexpectedly adverse.[15] The structural foundations of the US and Chinese governance systems create fundamentally different relationships with information, authority, and trust—differences which are especially consequential in the context of AI integration and crisis decision-making.

At its best, the US political system is characterized by distributed authority, institutional transparency, and a culture of public accountability. Power is intentionally fragmented across branches as well as federal, state, and local levels. While sometimes inefficient, this fosters resilience and adaptability. In the military domain, principles such as mission command empower subordinate decision-making and decentralized execution by focusing on a commander's intent rather than detailed orders, reinforcing resilience through a culture of initiative and decentralized problem-solving.[16] This is mirrored in civic life, where trust tends to flow upward from local institutions, and where corruption is popularly perceived as more prevalent at the top than at the bottom, imbuing a degree of grassroots legitimacy and responsiveness.

Vocal opposition by some segments of the US populace to broad scale AI adoption and data centers, an increasingly polarized information ecosystem, and declining trust in institutions overall might lead some to believe the US possesses ingredients of a domestic Cassandra Trap. But while these are real vulnerabilities in the US system, and it will require work to address them, they pale in comparison to the issues dogging the CCP. The US political system invites disagreement and public accountability whereas the CCP, by design, does not. The CCP has engineered a system anchored in centralized control, enforced political

loyalty, and pervasive surveillance. Authority is expected to flow vertically, and deviations from central directives, while common, are seen as pathologies rather than adaptations. Local fragmentation and provincial autonomy are a longstanding feature of governance, but they are cast as errors to be corrected, not strengths to be cultivated.[17] The People's Liberation Army reflects the same logic, privileging more rigid hierarchies and political reliability over initiative, while civic culture is shaped by conformity, compliance, and the avoidance of dissent.

While there are real vulnerabilities in the US system, they pale in comparison to issues dogging China

These distinctions are reflected in both nations' approach to AI governance. The United States and China are both home to competitive, commercially driven AI sectors with numerous private actors racing to develop frontier capabilities. The meaningful difference lies not in market structure but in how each system regulates the relationship between AI and the institutions of governance and social control. In China, that interface is tightly managed: private AI actors operate in a permissive commercial environment but within a framework that mandates cooperation with state surveillance objectives, requires data sharing with Party institutions, and orients deployment toward social monitoring and political loyalty.[18] In the United States, the equivalent interface is fragmented, contested, and subject to public scrutiny, a messier arrangement, but one that preserves flexibility, invites correction, and limits the degree to which AI outputs can be insulated from accountability.[19]

This contrast underscores a long-recognized paradox in complex systems called the "irony of automation."[20] As machines take over routine functions, the role of the human operator does not disappear, but rather becomes more critical and more demanding. Instead of handling a mix of easy and hard tasks, human operators are left with only the most difficult and time-sensitive problems, often without the benefit of practice on simpler cases. They must also supervise the automated system itself, catching errors or anomalies that the machine cannot see, which introduces an additional layer of vulnerability. Highly automated systems therefore require more attentive, skilled, and trusted human operators, not less.[21] Yet this human-in-the-loop role is prone to errors, including automation bias (over-trusting machine outputs), loss of situational awareness, and "skill fade" as operators lose proficiency through lack of use.[22] For the CCP, embedding AI deeply into governance and military decision-making risks demanding more judgment and initiative from human actors even as the political system discourages both.

Chinese leadership's push for AI integration into governance suggests they believe AI will *not* fall victim to the "irony of automation," but there is little evidence to support this confidence. For the foreseeable future, effective AI will require significant human-machine collaboration, especially in high-risk environments. Such successful collaboration, though, requires the human operator to question dubious outputs and take responsibility when the machine makes errors or produces something ambiguous. It is unlikely the need for human judgement will completely disappear in high-risk decision-making, so it would be prudent for CCP leadership to take this into account when developing their AI strategy. Structural biases in the PRC system, however, suggest that they have not and will not.

The two countries also have divergent approaches to AI governance. In the United States, efforts are fragmented but adaptive—driven by private-sector innovation and partially constrained by public scrutiny. Initiatives like the National Science Foundation's National AI Research Resource and National Institute of Standards and Technology's AI Risk Management Framework emphasize openness, interoperability, and responsible use.[23] In contrast, China's 2017 "New Generation AI Development Plan" and 2021 guidelines from the Cyberspace Administration of China explicitly call for the use of AI to achieve state-centric goals such as ideological conformity, data sovereignty, and societal control.[24] These mandates accelerate deployment but also bind technical systems to political incentives, making Beijing's AI outputs a reflection of what the regime wishes to see rather than what may objectively be true.

A final potential amplifier of mistrust lies in the CCP's hunger for predictive analytics. AI systems are getting better at interpretation and forecasting, but the inevitable mistakes that arise from rapid adoption can leave leaders wary. On the other side of the equation, repeated errors or overzealous actions which punish citizens unfairly also risk sparking backlash. For instance, the social credit system has already generated public anger when individuals were blacklisted for trivial infractions such as incorrectly separating garbage or eating on the subway, prompting authorities to scale back or quietly suspend certain enforcement measures.[25] These episodes show how missteps can erode legitimacy and trust: once leaders and citizens learn that the system can be wrong, it can be challenging to restore confidence.

Contemplating the Cassandra Trap highlights an important and more general asymmetry in US-China competition. The United States' messier, decentralized system may at times bog down, but it also provides resilience in the form of multiple channels for information, pluralistic scrutiny of data, and operators trained to question rather than blindly accept automated outputs. Beijing, by contrast, is betting on AI as a substitute for social trust and bureaucratic initiative, and to otherwise address challenges which stem from deep inside their autocratic

model. Under these conditions, sustained pressure may be sufficient to expose the internal frictions and self-defeating dynamics built into China's AI-dependent system.

US planners can pursue four strategies to exploit China's Cassandra Trap

Opportunities for Contestation

While the CCP's AI strategy is engineered to tighten control, the very design choices that make it powerful in peacetime expose a series of exploitable pressure points in crisis. American planners can lean into the Cassandra Trap by pursuing four strategies to use the system's weakness to their advantage.

First, target model integrity and data reliability. Injecting calibrated ambiguity—through diplomatic messaging, cyber-enabled influence, or the selective release of contradictory open-source evidence—can force Beijing to question the provenance and trustworthiness of the data feeding its AI systems. The objective need not be to deceive decisively or to overwhelm China's information architecture, but simply to introduce enough uncertainty to disrupt confidence in automated assessments. Because senior officials already distrust many data sources and struggle to interpret opaque, rapidly updated models, even modest amounts of synthetic noise can trigger demands for cross-validation, manual review, and additional reporting. In a highly centralized system, which punishes false positives and political misjudgment more severely than delay, these verification cycles consume time, diffuse responsibility, and blunt initiative. The result is not confusion for its own sake, but a predictable shift from speed to caution that consumes time, compromises initiative, and provides strategic advantage.

Second, exploit human incentives inside the bureaucracy. Routine anti-corruption drives and the threat of scapegoating encourage officials to prioritize self-preservation over candor. By highlighting inconsistencies between central directives and local realities, competitors can take advantage of this culture of caution, motivating regional leadership to pause for further verification rather than make a mistake. The same logic applies to the "lie-flat" ethos spreading among mid-level Chinese technocrats and young professionals; the more attention is drawn to the risks of acting, the more likely those officials are to freeze in place.[26]

Third, illuminate and expand the Chinese system's blind spots. Ordinary citizens, hackers, and even local police units have developed ad-hoc workarounds to avoid triggering the government's automated surveillance. For example, in the face of Beijing's use of spyware on citizens' phones, one coder created an Android application that hides sensitive information like location, messages,

and contacts.[27] Showcasing and normalizing these behaviors (but without necessarily directing them) erodes the perceived omniscience of the monitoring architecture. As doubts accumulate, CCP leaders must assume that their dashboards may be poisoned or simply incomplete. That credible fear alone can slow decision cycles, buying the precious hours or days needed to de-escalate a showdown.

Fourth, exploit frictions between Beijing and local officials. Provincial and municipal leaders frequently drag their feet on directives that shift costs or liability onto them, especially when compliance threatens local budgets, political standing, or personal security. Empirical research on China's anti-corruption campaign shows that heightened central scrutiny systematically alters local officials' behavior: when inspection intensity rises and punishment becomes more visible, officials respond by reducing discretionary activity, delaying approvals, and avoiding initiatives that could expose them to future blame.[28] Rather than improving compliance, intensified oversight often produces caution and inertia, as cadres prioritize political survival over policy execution. By underscoring the gap between central mandates and local capacity—policies that impose financial burdens, require improvisation, or entail visible downside risk—outside actors can amplify these existing incentives for delay and dilution. The result is not open resistance, but quiet non-implementation, slowing policy rollout and weakening the appearance of seamless central control from within.

Some might reasonably question whether setting a Cassandra Trap is too passive a tool given the stakes of AI competition, but its strength lies in its subtlety. It is far easier—and often more effective—to sow uncertainty than to deliver consistent, timely disruption. Undermining trust in the system's outputs, especially during moments of crisis, can be just as powerful a deterrent as degrading the system itself.

Resilience Over Omniscience

The United States and its democratic allies would do well to lean into their decentralization, adaptability, and trust in distributed decision-making. Rather than racing to match China's intrusive use of AI, the United States should simply take its cue from Apollo (the Greek God, not the space program), sowing doubt and creating conditions under which Beijing's omniscient centralized systems become its undoing. Thus, for China's democratic competitors, the long-term strategy must not be simply technical, but cultural. Resilience in the age of AI will hinge on the capacity of individuals and institutions to question, reinterpret, and when necessary, override automated outputs. That requires an educational system designed not merely to produce technical proficiency but to foster independent judgment, critical thinking, and institutional initiative.

Students must learn how to work with AI tools without surrendering to them—understanding both their potential and their fallibility. In effect, cultivating the habits of autonomy and distributed decision-making in the classroom prepares future leaders to preserve those same qualities in government, the military, and civil society. Such investments ensure that democratic societies retain the adaptive advantage in an era when authoritarian rivals may be paralyzed by rigid, over-centralized reliance on their digital oracles.

> For democratic competitors, the long-term strategy is not simply technical, but cultural

In the long run, democracies will retain an advantage not just by building better AI, but by avoiding the brittleness that comes with over-centralized reliance on it. Beijing's gamble is that AI can substitute for trust; the Cassandra Trap suggests that gamble may one day prove its undoing.

Notes

1. Shanshank Bhatt, "US vs China Tech Race 2025: Who Leads in AI, Semiconductors & Robotics," Outlook Business, December 29, 2025, https://www.outlookbusiness.com/explainers/us-vs-china-tech-race-2025-who-leads-in-ai-semiconductors-robotics.
2. Michael Schuman, "The Race for Global Domination in AI," *The Atlantic*, January 3, 2026, https://www.theatlantic.com/international/2026/01/china-ai-competition-differences/685389/.
3. Graham Allison, "Thucydides's Trap," The Belfer Center for Science and International Affairs, Harvard Kennedy School, accessed January 7, 2026, https://www.belfercenter.org/programs/thucydidess-trap.
4. Eddie Yang and Margaret E. Roberts, "The Authoritarian Problem," *Journal of Democracy* 34, no. 4 (2023): 141–150.
5. Xinhua, "China to Establish Modern Eco-Environment Monitoring System," Gov.cn, March 16, 2024, https://english.www.gov.cn/news/202403/16/content_WS65f5a1e4c6d0868f4e8e527b.html.; "How New Monitoring Systems Shaped China's War on Air Pollution," VoxDev, March 12, 2024, https://voxdev.org/topic/energy-environment/how-new-monitoring-systems-shaped-chinas-war-air-pollution.
6. "What Is Explainable AI (XAI)?" IBM, March 29, 2023, https://www.ibm.com/think/topics/explainable-ai.
7. Reduan Achtibat et al., "From Attribution Maps to Human-Understandable Explanations through Concept Relevance Propagation," *Nature Machine Intelligence* 5, no. 9 (2023): 1006–19, https://doi.org/10.1038/s42256-023-00711-8.
8. Steve Tsang, "Party-State Realism: A Framework for Understanding China's Approach to Foreign Policy," in *China's Big Power Ambition under Xi Jinping* (London: Routledge, 2021): 103–117; Susan L. Shirk, *China: Fragile Superpower*, vol. 165 (Oxford: Oxford University Press, 2008).
9. Jeremy L. Wallace, "Juking the Stats? Authoritarian Information Problems in China," *British Journal of Political Science* 46, no. 1 (2016): 11–29.; Heyuan Han, "Illusion of

Growth: Data Inflation among China's Local Governments," *ThinkChina*, accessed January 8, 2026, https://www.thinkchina.sg/economy/illusion-growth-data-inflation-among-chinas-local-governments.
10. Jiayuan Li, Xing Ni, and Rui Wang, "Blame Avoidance in China's Cadre Responsibility System," *The China Quarterly*, no. 247 (2021): 681–702.
11. Zolan Kanno-Youngs, Edward Wong, and Julian E. Barnes, "Coronavirus: Local Officials in China Hid Dangers from Beijing, US Report Finds," *The Irish Times*, August 20, 2020, https://www.irishtimes.com/news/world/asia-pacific/coronavirus-local-officials-in-china-hid-dangers-from-beijing-us-report-finds-1.4334827.
12. "Authorities Urged Not to Limit Coverage of Sichuan Earthquake," Reporters Without Borders, April 22, 2013, https://rsf.org/en/authorities-urged-not-limit-coverage-sichuan-earthquake.
13. Battista Biggio and Fabio Roli, "Wild Patterns: Ten Years after the Rise of Adversarial Machine Learning," *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*, 2018.
14. Matthew Jagielski, Alina Oprea, Bo Li, Chang Liu, and Raluca Ada Popa, "Manipulating Machine Learning: Poisoning Attacks and Countermeasures for Regression Learning," in *2018 IEEE Symposium on Security and Privacy*, 2018.
15. "Opinions of the State Council on Deepening the Implementation of the 'Artificial Intelligence+' Initiative," Center for Security and Emerging Technology, n.d., accessed January 7, 2026, https://cset.georgetown.edu/publication/china-ai-plus-opinions-2025/.
16. US Army, "Understanding Mission Command," April 30, 2015, https://www.army.mil/article/106872/understanding_mission_command.
17. Shui-Yan Tang, "Rethinking Local and Regional Governance in China: An Institutional Design and Development Perspective," *Urban Governance* 1, no. 1 (2021): 51–58, https://doi.org/10.1016/j.ugj.2021.11.002.
18. "The Chinese Communist Party's Layered Artificial Intelligence Strategy," National Security Data & Policy Institute, University of Virginia, 2025.
19. Michael Schuman, "The Race for Global Domination in AI," *The Atlantic*, January 3, 2026, https://www.theatlantic.com/international/2026/01/china-ai-competition-differences/685389/.
20. Lisanne Bainbridge, "Ironies of Automation," *Automatica* 19, no. 6 (1983): 775–79, https://doi.org/10.1016/0005-1098(83)90046-8.
21. Jonas Pöhler et al., "Keeping the Human in the Loop: Are Autonomous Decisions Inevitable?" *I-Com* 24, no. 1 (2025): 9–25, https://doi.org/10.1515/icom-2024-0068.
22. Lauren Kahn et al., "AI Safety and Automation Bias," Center for Security and Emerging Technology, November 2024, https://cset.georgetown.edu/publication/ai-safety-and-automation-bias/.
23. "National Artificial Intelligence Research Resource Pilot, NSF - U.S. National Science Foundation," accessed January 8, 2026, https://www.nsf.gov/focus-areas/ai/nairr.; National Institute of Standards and Technology, "AI Risk Management Framework," July 12, 2021, https://www.nist.gov/itl/ai-risk-management-framework.
24. "Full Translation: China's 'New Generation Artificial Intelligence Development Plan' (2017)," *DigiChina*, n.d., accessed January 8, 2026, https://digichina.stanford.edu/work/full-translation-chinas-new-generation-artificial-intelligence-development-plan-2017/.; "Cyberspace Administration of China (2021). Provisions on the Administration of Algorithmic Recommendations for Internet Information Services. 9th Decree," Scientific

Research Publishing, accessed January 8, 2026, https://www.scirp.org/reference/referencespapers?referenceid=3962757.

25. "Information Control and Public Support for China's Social Credit System," Stanford University et al., accessed January 8, 2026, https://sccei.fsi.stanford.edu/china-briefs/information-control-and-public-support-chinas-social-credit-system.
26. Lina Sun, "Beyond Resistance and Resignation: A Narrative Inquiry into the 'Lying Flat' Movement and the Reimagining of Work, Success, and Autonomy among Chinese Urban Youth," *Asian Journal of Social Science* 53, no. 2 (2025): 100198, https://doi.org/10.1016/j.ajss.2025.100198.
27. Jane Li, "How People in China are Trying to Evade Beijing's Digital Surveillance," Quartz, July 20, 2022, https://qz.com/1659328/chinese-people-are-pushing-back-on-beijings-digital-surveillance.
28. Haoyuan Ding, Hanming Fang, Shu Lin, and Kang Shi, *Equilibrium Consequences of Corruption on Firms: Evidence from China's Anti-Corruption Campaign*, NBER Working Paper No. 26656 (Cambridge: National Bureau of Economic Research, January 2020): https://doi.org/10.3386/w26656.