AI and International Security
Pathways of Impact and Key Uncertainties

Jason Pruet, Anna Makanju, Jonathan Reiber, and Josh Achiam
OpenAI

February 6, 2026

Abstract

Artificial Intelligence will reshape international security and the balance of power among nations. The most significant changes are associated with the ways that AI enhances or threatens principal functions of statecraft. We examine these effects across three dimensions: deterrence and the projection of force; the resources supporting national power; and the ability of states to understand and adapt to a dynamic strategic environment. This analysis draws on interviews with senior national security leaders, which provided insights into the implications of AI and how the technology is viewed. A central theme of these discussions is deep technical uncertainty. Future characteristics of AI remain sufficiently unclear that they admit radically different, and often incompatible, visions of the security landscape. We identify the most important of these uncertainties. Efforts by the AI community to reduce these critical gaps in understanding will strengthen the ability of national leaders to shape policy for global stability. These efforts will also help technology companies better understand the broader implications of AI, improving their ability to make responsible decisions for the design and deployment of increasingly powerful systems.

1 Introduction

We wrote this paper amid three converging trends: a rapid growth in global investment to harness the power of advanced AI, a commensurate increase in attention to AI's potential impacts, and a generational change in the character of conflict [1–3]. Within the large AI companies, much of the focus has been on the challenges of alignment, safety, and broad access for beneficial applications [4, 5]. In public discourse focused on national security, risks associated with applications of AI – cyberspace operations, bioterrorism, and autonomous weapons – have received significant attention [6–10].

Here we consider the consequences of AI for relations between nations. Prosperity of society and the well-being of its people depends on stable peace. A growing body of research, and our discussions with national security leaders, suggest AI is already becoming capable enough to begin disrupting the foundations of international stability. In this sense AI is emerging as a distinct domain of geopolitical concern, alongside established areas such as cybersecurity and outer space. Developing the capacity to navigate these geopolitical impacts, including by leveraging AI for proactive solutions, will be at least as important for humanity's future as managing traditional cyberspace and biosecurity risks.

This paper focuses on AI capabilities that shape the core functions of statecraft. Among the most important of these are capabilities that accelerate scientific discovery; strengthen planning and coordination in complex strategic competition; and sharpen situational awareness (the ability to perceive, understand, and anticipate events). Our scope is frontier AI systems capable of broad, complex cognitive labor—systems that plan, reason, and coordinate – rather than narrow applications like image recognition or late-phase targeting. While specialized AI models are already changing warfare, the more general models may have the greatest potential for transforming the security landscape.

The boundary between the two types of capability is increasingly blurring as specialized functions yield to more general systems.

History shows that general-purpose technological advances and long-run shifts in economic strengths, rather than specific weapons systems, have often had the greatest influence on the global balance of power [11]. The marine chronometer, created for commercial navigation, enabled precise longitude determination to transform naval commerce and warfare [12]. Nations that had been safe from European navies soon no longer were. The electric telegraph enabled near-real-time command and control in the mid-nineteenth century [13]. Daniel Headrick's Tools of Empire shows how steamships and quinine extended European military reach [14]. Even the humble stirrup is credited with having reshaped social institutions and the approach of medieval states to warfare [15].

There is substantial uncertainty both in AI progress and in its geopolitical impact. The closest historical analog may be with the beginning years of the Cold War [16, 17]. Then a spectrum of powerful new technologies drove a basic reimagining of the security order. This required decades of work, the creation of new disciplines, and broad collaboration between political leadership and technical experts. Even so, there were harrowing close calls and a great degree of luck [18, 19]. Finding a way to wisely navigate and invest in an international order that lets us achieve the broad potential of AI for good, reduce the probability of conflict between peer competitors, and reduce reliance on luck, will require a coherent large-scale effort [3, 20, 21].

Our concern here is with one small part of the larger challenge. Inspired by previous studies and our discussions with national security experts, we start by examining characteristics of AI most likely to affect international security. These include capability thresholds that flip deterrence dynamics, resource dependencies that determine who can keep up, diffusion pathways that shape how fast advantages erode, and others. Another effect – the impact of AI on human cognition, perception, and trust – is also central to understanding implications of AI but is beyond the scope of this paper [22, 23].

To motivate our later discussion, consider a world in which by 2030 AI increases the rate of progress in undersea sensing technology by a factor of ten. That means that by 2040 we have capabilities we would not have expected until far beyond today's planning and acquisition cycles. Before the last of twelve planned Columbia class nuclear submarines has its first deployment, the state of science will have surpassed that at its planned end of life by fifty years. Because submarines provide the sea-based leg of the nuclear triad, and because their deterrent value rests on the difficulty of detecting them, their survivability is central to strategic stability. We have no structures, organizational culture, or experience that prepares us for a century of military science occurring every decade.

We identify the technical uncertainties that need to be addressed by AI developers to help policymakers prepare for the future. Although some illustrative scenarios are discussed, we avoid predictions about timelines or capability end-states. Expert disagreement is reasonable, and there can be no substitute for a cross-disciplinary program of research aimed at reducing substantive uncertainty.

2 Context

Recent advances in AI provide a frame of reference for understanding present-day debates about its influence on international security. Comprehensive accounts of this history are described in the references [24, 25]. Here we briefly summarize key evidence. This evidence has been a foundation for researchers attempting to develop a theory of security in light of the rapid arrival of increasingly capable chat, reasoning, and agentic AI systems.

2.1 Two Lessons from Recent History

Modern AI is based on deep learning – the training of deep artificial neural networks, or models that use many successive learned layers of computation to process information. As recently as 2006 it was considered very difficult to train neural networks with more than two hidden layers [26]. Efforts to train such networks routinely produced poor results. Six years later a landmark result from a team at the University of Toronto demonstrated a deep network with eight layers could significantly improve over the previous state-of-the-art in image recognition [27]. The authors made the following prescient observation:

> In the end, the network's size is limited mainly by the amount of memory available on current GPUs and by the amount of training time that we are willing to tolerate. … All of our experiments

> suggest that our results can be improved simply by waiting for faster GPUs and bigger datasets to become available.

Another significant development occurred about a year later. Mnih et al. trained deep neural networks with reinforcement learning (learning by trial and error) to play Atari games, in some cases surpassing the performance of human experts, based only on raw visual inputs without special preprocessing [28]. The learning procedure did not have any examples of human play to learn from.

These and other results catalyzed a flurry of activity that led to the development of highly-capable narrow AI systems. Development of AlphaGo in 2016 is regarded as a culminating point for this wave of research. By integrating many earlier advances, this system achieved superhuman play in Go far earlier than many knowledgeable people thought possible.

Out of this period of change came a first significant lesson: *For every challenge for which it is possible to measure success, deep learning will solve it if provided with enough computing power and data.* Here a basic requirement for a scalable deep learning method is that it is possible to specify which outputs are desirable with either a set of examples (as is the case with labelled images) or a reward function (such as winning at a video or board game).

The arrival of large language models based on the transformer architecture, coupled with the development of a quantitative science of their performance characteristics in the form of scaling laws and other insights, moved the field of AI from narrow applications such as visual recognition or game playing to general-purpose models [29, 30]. Today a single model can converse in natural language, perform useful tasks in coding and software engineering, solve certain kinds of advanced problems in mathematics, translate between languages, and reason strategically. In the last year AI has become capable of forecasting world events with accuracies comparable to those of experts [31, 32]. Some models are served in systems that have prototypes of long-term memory built in, enabling them to reference earlier conversations and experiences.

Advances like those held another significant and surprising lesson in the nine years following 2016: *A great many problems that humans can solve with thinking can be formulated so that a scalable deep learning method can be applied.* While many capabilities have not yet been demonstrated, there is optimism within the technical AI community that the pattern of rapid advances will continue. So far this has been biased towards problems with a clear notion of right and wrong, like those in math and science. With larger effort in developing evaluations, and techniques like synthetic data generation, the complexity and scope of problems that can be formulated this way grows each year.

These lessons led to quick change. Four years ago ChatGPT did not exist. Today roughly one person in ten in the world is using large AI models [33]. Three considerations point to still greater change:

- Empirical trends in model performance, rising investments in computing, and falling costs of computing power point to predictable improvements in AI capabilities and accessibility.
- Modern AI systems have arrived so quickly that there is still a great effort needed to integrate them into useful workflows.
- Advances in methods, or combinations of existing methods, may markedly improve capabilities in the near term.

The next section describes some of the most important empirical trends. The other two considerations are briefly touched on in later sections.

2.2 Empirical Trends

If general lessons from recent advances clarify the landscape of AI in broad brushstrokes, quantitative data on empirical trends fill in details needed for reasoning about the future. These trends provide a useful simplification that collapses complicated technical measurements into a concise summary salient for international security. Thanks to outstanding scholarship by groups around the world there is now high-quality analysis describing some of the most important measurable characteristics of AI.

The first of the major trends to have become clear described scaling laws for training. These relations show how models get better as more computing power is used to train them. This had been the subject of intense interest for decades [34]. However, it was only recently that the measurements for very large scales of computing were possible. The largest models available today were built with more than ten billion times the amount of computing used for the

largest models of 2010 [35]. That growth in computing was observed to provide a steady improvement performance of AI models [36].

In autumn of 2024 a breakthrough associated with a different type of scaling law was announced [37]. Instead of relying on more computing during training, this advance allowed models to spend more time reasoning (often called "test time compute" in the literature) before responding. There was an immediate jump in performance for many difficult problems. There are also implications for international security that are broader than those associated with improvements in model capabilities. For example, reasoning can make it more difficult to understand what models are capable of. Many tasks impossible with low reasoning effort are achievable with application of enough computing power to think through the problem. Another impact, related to the possibility that computing becomes an essential resource for national power, is discussed later.

One consequence of promising performance trends has been an enormous growth in computing available for both training and use of AI. A rough estimate is that in the United States there are now more than five million GPUs dedicated to AI. In terms of simple volume of output, these can produce in a day several times the sum of all books ever written by humanity when using the most sophisticated current reasoning models.

At the moment AI is either not capable enough, or we have not yet found the right methods, for using output from AI models at that scale to qualitatively change the foundations of power for a nation. It is possible that our ability to understand how to most effectively adopt AI has appreciably lagged its potential, much like the first paradigm for electricity used it as a direct substitute for centralized mechanical power in factories. It is also possible that a new generation of models will suddenly unlock a new span of applications.

3 Insights from the Community

Researchers have drawn a broad spectrum of conclusions about implications of AI for international security.[1] That range reflects different perspectives on two important dimensions of this complex topic. The first concerns the character and pace of advances in AI capabilities. Past trends are often not reliable predictors. The realm of reasonable expert views allows many competing estimates for when important breakthroughs will be made, or even if they are possible at all on the current path.

Different authors have also described a variety of pathways by which AI will affect security. A survey of some recent widely-cited works gives an impression of the diversity. The studies *Artificial General Intelligence Race and International Security* and *How Artificial Intelligence Could Reshape Four Essential Competitions in Future Warfare* by RAND explore multiple perspectives on how advances toward AGI could reshape international stability [38, 39]. Though not focused directly on security, Harari argues in *Nexus* that advanced AI will reshape how societies construct shared understanding, suggesting that changes in the means of producing stories about the world may become a source of instability [40]. The essays *Situational Awareness* and *AI2027* examined how rapid AI progress could give a nation with even small initial leads a decisive advantage over its competitors [41, 42]. *Superintelligence Strategy* goes further, arguing that the prospect of unilateral dominance, and the risk that a state loses control of its AI, would pose an unacceptable threat that motivates a new security regime [43].

Notwithstanding a growing body of thoughtful analysis, the national security community has not achieved a consensus on principal needs for preserving international security in a future with powerful AI. Building the path forward will depend on better information from the AI laboratories, just as understanding deterrence and non-proliferation depended on technical input from nuclear scientists in the Cold War era. To better understand the kinds of information that would be most useful we interviewed senior experts and leaders in national security. We also spoke with experts that had studied implications of AI for economics, a key aspect of national strength.

Our interview subjects come from diverse backgrounds and described a broad variety of potential impacts of AI on international security. In some cases, their concerns have been the subject of extensive study; in others, they provided insights that we did not find in the literature. Perhaps of most value was their first-person account of the challenges in adapting to a future with powerful AI. Nearly every individual we engaged has been on the front lines of this

[1]The term "international security" has been defined in many competing ways, and no single definition is universally accepted. In this paper we use it in a deliberately broad sense, referring to efforts and conditions aimed at maintaining peace and stability in relations among states.

challenge. They gave frank descriptions of the practical difficulties in moving organizations, of persuading national leadership, and of communicating effectively about a technology that is both strange and fast moving.

Many concerns related to AI and international security are with imminent near-term challenges. Adoption – whether militaries and institutions can integrate AI into doctrine, operations, and decision-making – was often highlighted as the decisive factor. "The impact of AI on the military is not predominantly dependent on the technology, but on the assimilation process," former Secretary of the Navy Richard Danzig said: "If I put a bounteous feast in front of you but your jaw is wired shut, you can't eat." There was also a sense that deeper direct experience with these technologies is essential for building an understanding. One of the experts we spoke with noted "We don't have an information gap, we have a visceral belief gap." With this view, effective communication is only part of the need. Creation of environments in which decision-makers and operators can safely engage, experiment, and train with the technology for their own needs is also essential.

Experts also underscored AI's many applications: it can accelerate logistics, intelligence, and productivity, but it can undermine defense with cyberattacks, biothreats, and escalation risks. The experts we consulted warned of the dangers in a future where AI compresses decision timelines, lowers political costs of war, and threatens to normalize "riskless warfare." The specter of authoritarian governments developing and exploiting AI without democratic accountability was also a common theme. The span of concerns raised included the use of AI to enable pervasive surveillance, suppress dissent, manipulate information at scale for the purpose of driving a population to war, concentrate decision-making power, and remove human judgment from critical political and security processes.

We are still early in building a collective understanding. No one we spoke with expressed confidence they can predict this future, or that the best course of action is known. As Lieutenant General (USAF, ret'd) Jack Shanahan noted, "The problem is massive uncertainty. Decision-makers are torn between claims that 'this will end the human race' and 'this can't add 4 digit numbers'." Along with uncertainty is the challenge of an uneven understanding of AI by civilian and military leadership. Several of the people we interviewed emphasized the need for collaborations with technology companies to quickly build a common understanding.

The most important uncertainties that the frontier AI labs can help with lie at the intersection of technology and pathways for changing international security. We turn to those in the next section.

4 Pathways and Key Uncertainties

We consider three pathways by which AI will shape relations between nations. These pathways concern changes in:

- The ability of national and subnational actors to project force and the resulting effects on deterrence.
- Resources essential for national power.
- The ability to understand the competitive environment.

AI's influence on these pathways depends on how it alters a small set of general-purpose functions of statecraft. The most important of these include scientific acceleration, situational awareness, planning, forecasting, and efficient resource use.

An account of how AI will influence the evolution of international security will take the effort of diverse communities working over the course of years. Our more modest goal is to identify the technical uncertainties whose resolution would most improve the ability of these communities to reason about the future of international security. A common understanding of these uncertainties provides a shared frame of reference for interpreting new information about AI capabilities as it emerges. Because much more is now known about key technical characteristics of AI than was understood even a year ago, some of the consequences for international security are already becoming clear.

Table 1 summarizes a set of technical uncertainties associated with the three pathways. To illustrate their importance and the kinds of analysis required to resolve them, we examine one key uncertainty for each pathway in detail.

Pathway: Changes in deterrence and the ability to project force	
Technical uncertainties	**Why this pathway matters**
• When and how will AI accelerate progress in science and technology for militarily significant applications? • To what extent will AI lower the resources needed to create and execute cyberspace, chemical, biological, and other threats? • What will enable AI to exceed human capabilities for devising strategy, meaningful courses of action, and negotiation tactics?	• Governs timing and magnitude of shifts in military effectiveness and strategic balance. • Determines efficacy of the defense plan of record. • Impacts proliferation of means to cause harm. • Sets importance of resilience, redundancy, and diversified pathways for projecting power. • Shapes offense-defense balances that may change deterrence (by denial and punishment) or coercion. • Creates possibilities for improved diplomacy. • Relates to the possibility of decisive advantage.
Pathway: Changes in the resources essential for national power	
Technical uncertainties	**Why this pathway matters**
• Will computing capability become a decisive resource for national power? • To what degree will the inference compute available to nations shape their economic growth and military advantage? • What scale of computing and research is needed to reach key capability thresholds in AI? • How will the cost of training and inferencing frontier AI change over time?	• Shapes distribution of advantage for national power. • Changes the resources required to generate and sustain military forces. • Influences race dynamics and preemption. • Determines pace and pathways of capability diffusion. • Informs our understanding of the adequacy of existing governance approaches for compute and AI capabilities. • Relates to the possibility of decisive advantage.
Pathway: Changes in the ability to understand the competitive environment	
Technical uncertainties	**Why this pathway matters**
• Will AI be able to discover protected information by inferring it from other information that is not protected? How robust is secrecy? • How good will AI be at creating actionable insights and superior options from multi-source data (including under uncertainty, deception, and tight time constraints)? • To what extent will AI compress decision cycles for complex operations? • Will AI enable significant scientific discovery in a datacenter with minimal observability? • How capable will AI be for behavioral forecasting and shaping?	• The fall of secrecy would undermine present security protections and require significant changes in posture and mindset. • Decision-making speed can collapse stability and enable escalation spirals. • Advantages in situational awareness translate to advantages in operations and policies. • Sets conditions for strategic surprise; erosion of traditional indicators and warnings and emergence of entirely new ones. • Shapes the efficacy of cognitive warfare. • May create heightened misperception, misattribution, and crisis instability.

Table 1: Pathways and key technical uncertainties.

4.1 Deterrence and Projection of Force

The concern for this category is with changes in the susceptibility of nations to violence. Possible consequences of AI include the gain or loss of asymmetric advantage, technological developments that weaken defense, and acquisition by sub-national actors of the means of causing large-scale harm [7, 44, 45]. The key uncertainty we focus on here is associated with one of the most important sources of major changes in military capabilities:

> How will AI markedly accelerate or otherwise disrupt science significant for deterrence and the projection of force?

Steady progress in science and technology, or just a few unexpected breakthroughs, has often been a path to upsetting balance in military relations [46]. The state of science provides the environment in which threats are created and executed. The scale of resources needed for major advances has influenced both the present structure of American society and the dominance of the U.S. military over the last eighty years [47]. In recent years the United States has often outspent the rest of the world in defense R&D. In 2016, for example, it invested about $78.1 billion in defense R&D, which was over seven times as much as all of the other OECD countries combined [48]. In 2025 DoD will spend about $140B in Research, Development, Testing and Evaluation [49].

> *The most important thing is to get there first so we can turn AI into real advances as fast as possible. The one that goes quiet so they can put fusion energy on the grid or develop the $500 electric car can own the economy of the world. The alarming scenario is where things are integrated right.*
>
> *— Pat Fitch, Deputy Director, Los Alamos*

A comparison of national defense R&D investments with those made by the private sector for AI is useful. Though public reporting does not typically separate spending on AI technology development from the costs of operating commercial AI applications, available information allows a rough estimate. Stanford's AI Index estimates U.S. private AI investment exceeded one hundred billion dollars in 2024. Leading U.S. hyperscalers are spending several times that amount in 2025 to build AI computing infrastructure. These infrastructure outlays are not fully captured in standard "private investment" tallies, and they still exclude costs for technical talent. It is plausible that private-sector spending aimed at developing and scaling advanced AI capabilities now exceeds the military research and development budgets of nearly every country in the world. This scale of civilian-led investment implies that governments will gain access to an AI-enabled accelerant for scientific and technological progress, along with other strategically significant capabilities, generated by an effort comparable in scale to their own defense research enterprises.

A number of authoritative studies describe types of advances in science and technology that would have significant implications for international security [50–52]. Table 2 gives just a few examples. These are all from areas that are the subject of active research and where at least some progress in the next few decades seems likely. Most recently, the war in Ukraine has illustrated how the confluence of digital operations and unmanned systems can undermine traditional assumptions about the effectiveness and survivability of large weapons platforms [53, 54]. From history we also know that many of the most important disruptions – such as the development of quantum algorithms for

Foundational area	Example of disruptive advance	Possible implications for deterrence / stability
Advanced materials and stealth	Radar-/IR-absorbing coatings	Harder-to-detect air/sea platforms complicate denial and warning; surveillance/counter-stealth races inject uncertainty into second-strike survivability.
Quantum and non-acoustic undersea sensing	Non-acoustic sensing concepts (e.g. quantum magnetometers) at scale	Erosion of strategic stability through nuclear ballistic missile submarine detectability.
Cryptanalysis and algorithmic breakthroughs	Threats to traditional and post-quantum encryption	Communication compromise raises risks of spoofing, misperception, and escalation; loss of secrecy protections.
Energy storage	Higher-density batteries for unmanned endurance.	More capable attritable assets strengthen denial but put pressure on deterrence forces.

Table 2: Examples of potential science and technology disruptions with implications for deterrence/stability.

breaking encryption, CRISPR (a powerful platform for targeted genome editing), and capable generative AI itself – came as a surprise. If AI is capable of novel discovery of those types, a still open question, the aperture of studying its effect on international security will need to be broader.

> *If AI is making scientific and engineering discoveries, then you will have new innovations bubbling up around the world and, all of a sudden, the technology landscape looks different than it did historically. Now, new places are producing new things. If that happens really fast, all over the world, that would cause serious disruption and change the balance of power.*
>
> *— Neil Thompson*

The example given in the introduction, of science advances that may affect submarines, was not purely speculative. Three NATO countries and Russia rely on the difficulty of detecting submarines carrying ballistic missiles as a central pillar of strategic nuclear deterrence. There is a long-standing effort to harness advances in AI for improving signal to noise discrimination in undersea magnetic anomaly detection, and for fusion of information from multiple signal types [55]. Generative AI is now beginning to demonstrate promise for advancing discovery of high temperature superconductors, although no breakthroughs have been confirmed [56, 57]. Because superconductors exhibit exceptional sensitivity to magnetic anomalies from metallic hulls, achieving such performance at higher temperatures would make practical field deployment more feasible. While success would bring many great benefits, it would also have implications for large-scale distributed platforms for submarine detection [58].

Our ability to adapt to the power unleashed by science has been tied to the need for humans as engines for progress. That need set timescales, constraints on capability, requirements for massive investments, and a measure of transparency useful for stability. There is a deep debate within the scientific community on how much AI will change these factors. Though some evidence for the impacts of frontier models on science has begun to accumulate, systematic research is still scarce [59].

Research by NATO's Science and Technology Organization on long-range trends gives valuable insights into present estimates of the timescales for advances in different areas. While that research emphasizes the possibility of disruptions to those timescales from AI, the technical uncertainties are so large that quantitative estimates are not possible at present. Arguments from thoughtful experts span a wide range. To give a sense of the spectrum of perspectives that national leadership is hearing, contrast the assessment in:

> [I]f AI allows us to make better predictions from incorrect theories, it might slow down scientific progress if this results in researchers using flawed theories for longer. In the extreme case, fields would be stuck in an intellectual rut even as they excel at improving predictive accuracy within existing paradigms. – Kapoor & Narayanan

with

> Our goal is to compress the next 250 years of chemistry and materials science progress into the next 25. – Microsoft CEO

Those holding the view that broadly capable AI models may not be very significant to progress in science point to evidence that overall productivity in science has become less efficient as the body of papers grows as well as evidence for basic weaknesses in model capabilities [60–63]. Those making the case that science will be accelerated emphasize the potential for non-linear gains as AI systems become capable of large-scale synthesis across previously disconnected fields, continuous verification and reanalysis of existing results, and increasingly automated integration of hypothesis generation, experimentation, and evaluation. They also note early successes in applying AI across the discovery lifecycle, the rapid adoption of increasingly larger models for scientific use, evidence of superior performance in predicting experimental outcomes, and targeted breakthroughs in domains that had previously resisted human effort [64–69].

As each day brings new results, the case that AI will substantially advance science gets stronger [70]. Still, there are very few measurements of its impact and no sound ability to project into the future. There is an urgent need for careful large-scale studies to quantify the impact of frontier reasoning models on science and technology. Current security plans were developed only with experience of historical rates of scientific progress. Great uncertainty about how scientific progress will be disrupted complicates assessment of the continued effectiveness of those plans. If AI

does markedly accelerate science, we face rising risks as we take refuge in the seemingly conservative side of that uncertainty.

4.2 Resources for National Power

Where the previous section examined how AI affects coercive force, this section considers how AI will reshape the basic resource structures that underpin national strength. AI may change dependencies on essential resources in many ways[2, 71]. Perhaps the most obvious is related to the possibility for shocks in employment, a well-studied source of disruption [72, 73]. Another relates to the ability of AI to advance science and technology in ways that reduce needs for existing resources. Ed Conway's *Material World* described how even the most advanced technologies depend on basic and rigid material foundations. History also shows that breakthroughs can rearrange those foundations. AI may do the same today. For example, there is now an active effort to use AI both to enhance the efficiency of rare earth element production and to accelerate the discovery of alternative materials [74, 75]. This type of technology-driven independence can strengthen security. It can also cause slow-burn destabilization of alliance and power structures, as happened with Germany's invention of the Haber-Bosch process that reduced reliance on natural saltpeter in explosives production [76].

> *It's not just that AI can affect strategic stability, it's that by itself it could become a central driver of strategic stability.*
>
> *— Gen. Shanahan (ret.)*

The uncertainty we consider here is associated with technology-driven dependence on new resources, particularly those needed for AI itself. Countries around the world have developed strategies reflecting the view that access to AI will be an important source of domestic vitality and national power [77–80]. The effect on international security of this new entrant in the set of vital resources will depend on technical details of the relationship between benefits that AI provides and the time, computing, and human talent needed for realizing these benefits.

To illustrate the significance of this relationship we will focus on the narrow question,

How do asymmetries in AI computing capacity across nations affect asymmetries in national power?

A simple definite form of this question examines whether a nation with ten times the AI computing resource of an adversary can make scientific discoveries at a decisively faster pace. The converse consideration is that a nation may make a mistake and invest too much in the wrong path for an AI-enabled future. That would lead to great opportunity costs and a weakening of their power relative to nations that had invested more wisely in AI or, in an extreme case, in different types of advances altogether.

Experience of the last few years provides initial insights. At present just a few countries in the world produce frontier models, although many groups are not far behind. The first to reach important thresholds have been those that are investing enormous computing resources, though progress from algorithmic improvements has also been very fast. There has also been a strong correlation between improvements in different capabilities of frontier models. Marked improvements in math, coding, writing, and other tasks tend to all be made at the same time. If that continues to hold, the group that first achieves one threshold in capability will be the first to achieve many.

A better understanding is important in part because computing needs set timescales. The pace at which computing is growing is bounded by technological trends and economics, and can at least be approximated [81, 82]. For illustration, consider the threshold associated with marked acceleration of GDP growth. According to EpochAI economic modeling, AI will double global world productivity in just three years if AGI only requires a training run that uses 2,500 times more computing than has been used for training the largest models to date (gains in economic productivity occur before all tasks are automated in this model [83]). As another example, consider capabilities for scientific discoveries that challenge present defense architectures. If such models require computing that is available next year, we could begin to see those disruptions decades before current acquisition programs have been completed.

The significance of asymmetries in scale also shapes the balance of power between nations. A need for much larger resources than have been invested so far would mean that only the United States and perhaps China could develop AI that reaches meaningful thresholds. For the case of a machine intelligence capable of accelerating science, this would mean that some nations risk falling ever further behind technologically. In particular, they could risk military irrelevancy unless some kinds of international agreements were made. On the other end of the spectrum, very low

barriers to achieving important thresholds could risk small groups having the potential to develop destabilizing military capabilities.

If a nation believes that it has time to develop an appropriate response to an AI-based disruption, that belief can be stabilizing as it will decrease the likelihood of the nation conducting a preemptive action or taking another step on an escalation ladder to defend itself. The uncertainty of imminent advantage by an opponent leads to preventive war as an attractive strategy [84].

There are two limiting cases:

Case 1. AI capabilities gently improve with increasing resources. This case is illustrated by discussions of AI as a "normal technology" [85].
Case 2. There are abrupt improvements in capability of AI with increasing resources. The possibility of this case is behind concerns that small leads in being the first to achieve superintelligence or other thresholds of AI capability will give decisive advantage [86].

A third case, sitting between these two, is of special interest to discussions of security. This describes "spiky AI" that has extraordinary capabilities in narrow areas. An example is the rapid improvement in AI models for areas related to cyber-security [10, 87].

To see why the difference between these cases is important, consider a simple scenario with two nations. In this scenario, one nation leads in AI with capabilities significant for strategic advantage. They can choose to race ahead or enter an international agreement to preserve rough parity. The other nation is permanently behind, and can choose to try to preemptively hobble the adversary's AI capability, or wait.

If the benefits of AI appear gradually, as in Case 1, the side that is behind does not have strong motivation to deny creation of AI to its adversary. They lose by attacking, and would not suffer unacceptable loss from not attacking. In Case 2, where AI quickly provides decisive advantage, the weaker side has stronger incentives to avoid a future where they would suffer more damage than would result from a preemptive attack. In either case, the nation with a lead in AI development has a motivation to safely preserve that lead because AI provides benefits even absent defense motivations. These qualitative considerations can be made formal with a game-theoretic treatment of the incentives for the side that is behind to attack [88].

An application that highlights the possibility of significant and abrupt improvements concerns the depth of machine-supported planning relative to an opponent. An effective planning horizon can be characterized as the depth (in sequential decisions) at which a system can reliably evaluate consequences under realistic time, compute, and noise constraints [89]. Many operations contain "trap" structures in which an initially attractive move becomes unfavorable only beyond some minimal depth, with logistics reconfigurations as a classic example. When the disparity between depth of planning crosses the trap depth, the feasibility of such designs changes. One side can set traps the other cannot avoid [90].

In this case the link between computing resources and their effects is discontinuous. We should expect jumps in operational options as planning depth passes typical trap depths. Deception risk becomes asymmetric. If one side expects the other's horizon to be short, traps with payoff asymmetries (large upside if sprung, moderate downside if declined) become rational to deploy and cheaper to proliferate with AI assistance, even absent intent to escalate.

Beyond better characterizing the significance of asymmetries in compute resources, it may be valuable to create and track simple metrics describing AI resources available to nations. One such quantity, National Inference Compute, would measure the sustained rate at which a country can deliver model tokens (or other measures of useful output) at a given quality. Another metric could describe the time it would take to focus computing resources in the event of a national emergency. This is useful because one nation may have a great advantage in decentralized computing that would not be of much help in the event of a crisis against an adversary that better integrates computing with military or intelligence functions. A Compute Mobilization Latency metric would account for this by estimating the time between a policy decision to harness computing for national purpose and the moment when the capacity is online.

4.3 The ability to understand the competitive environment

The final expression of AI's capabilities we consider relate to how states perceive and interpret the strategic environment. Changes in the ability of nations to understand this environment have been at the heart of many conflicts. Weak awareness can raise fears of threats in ways that invite preemption [91]. This was a consideration behind

transparency measures in arms control. Weak awareness can also increase complacency to erosion in the status quo and raise the risks of strategic surprise. Improved understanding of the environment can also pose risks. It can encourage adventurism by supporting confidence that victory can be achieved, or make a nation vulnerable if exploitations of defense become known [92].

There are several dimensions along which AI may alter the availability of critical information. One of the most important concerns the speed and quality of insights in periods of tension or conflict. AI can affect both crisis dynamics and relative advantage once fighting begins. As that has been the subject of many careful studies (e.g. [2, 93–95]), we focus here on a concern for which there may be a very sensitive dependence on capabilities of the most advanced AI models. The key uncertainty we consider is:

> Will AI shorten the time over which secrets can be preserved?

A closely related question is whether AI will make it much harder to understand our adversary's secrets. Though it may be counterintuitive, the answer to both questions can be yes. Advances made within the confines of a small datacenter may uncover information that an adversary wants to protect, and may also be essentially impossible for an adversary to see from outside. The question of when release of information about clandestine capabilities is stabilizing, and when it should be avoided, is a difficult one [96].

The importance of secrets for security is reflected in the great lengths that nations go to prevent their disclosure. The US defines top secret information as that whose release could cause "exceptionally grave" damage to national security, and secret information as that whose release could cause "serious" damage [97]. While it is possible that nobody knows the exact number, an enormous number of distinct statements that fall into one of these categories.[2].

It is useful to distinguish between "social secrets" and "technological secrets". By social secrets we mean those facts about the human world, such as details of cabinet-level deliberations on war plans, whose secrecy depends on the strength of security measures and the difficulty of predicting human behavior. A body of research has shown that statistical methods and AI can draw remarkably strong inferences for many aspects of human behavior and decision-making [98–100]. In an extreme limit, AI may do a good job of approximating what was said behind closed doors even if the deliberations are never spied on or leaked.

> *This is going to be a massive opportunity and massive risk in the hands of enemies finding us because we have patterns in the way we do operations.*
>
> *— Anne Neuberger*

Technological secrets are those that involve knowledge about the physical world. As an example, the British Government Communication Headquarters, an analog of the National Security Agency in the United States, discovered and kept private public key cryptography for years before it was described in open literature [101]. Things like algorithms or tricks in the applications of fundamental physics for weapons can in principle be found by a smart enough person, or a good enough artificial intelligence. This has the potential to suddenly expose information that protects the balance of international power without any theft of blueprints or spies eliciting information from scientists.

AI may also accelerate development of technologies for finding secrets. Shor's algorithm was published in 1994. Over the subsequent 20 years there was promising progress towards a quantum computer, and in 2015 the National Institute of Standards and Technology announced a competition to develop post-quantum cryptography algorithms. The significance of having enough lead time was emphasized by NIST:

> Some secrets remain valuable for many years. Even if an adversary can't crack that encryption that protects our secrets at the moment, it could still be beneficial to capture encrypted data and hold onto it, in the hopes that a quantum computer will break the encryption down the road. This idea is sometimes expressed as "harvest now, decrypt later", - and it's one of the reasons computers need to start encrypting data with post-quantum techniques as soon as possible [102].

[2]For example, as of fiscal year 2022, the Department of Energy alone maintained 128 separate classification guides. https://www.archives.gov/files/isoo/fcgr/fcgr-2022-final-report-doe-.pdf

Even had it been possible to develop trusted quantum-resistant algorithms quickly, it would have taken many years for human organizations to adapt to them. A particular concern, even with post quantum encryption techniques, is with the increasing ability of AI to make advances in mathematics and algorithms [66, 103].

Although it would take us too far afield to go into detail, it is worth emphasizing that AI's impact on secrecy has implications that are much broader than traditional state secrets. One example is for intellectual property protection. As Pat Fitch noted, as AI systems become capable of automated invention or rapid reverse engineering, the traditional logic of patents—public disclosure in exchange for years of exclusivity—may weaken. If AI makes "ordinary skill in the art" equivalent to what is available in open-source tools, reductions to practice and trade secrets may become far more valuable than patents with long timelines.

The US protects different types of information for different lengths of time. Some information is only kept secret a short while, some for 50 or seventy five years, and some like design details for weapons of mass destruction indefinitely. A clear sense of whether that will remain possible is needed. As partnerships between industry and government grow, there will be a better understanding of the evolution of the ability of frontier models to undermine secrecy. However, such a retrospective look at capabilities of models already developed is not very useful on its own. To provide useful warning we need some ability to look ahead.

5 Conclusions

Ensuring that AI benefits humanity will depend on carefully addressing its effects on relations between nations. This concern overlaps with worries about the use of AI in weapons and other direct means of harm. But it is also broader. AI has the potential to cause structural disruptions to the foundations of security that go far beyond weapons. As with the marine chronometer, the most significant effects will likely be indirect.

Uncertainties about basic characteristics of AI affecting international security are so significant that they lead to wildly divergent views of what the future will look like. This is unusual. Technologies like high performance computing, lasers, and hypersonic missiles are all understood to be important for the future balance of power between nations. They are also all to some degree unpredictable. But the span of uncertainty does not allow much debate about whether small groups will be able to undermine states, or whether the basic calculus of incentives for preemption can be preserved. With AI the present range of uncertainties is so large as to be unhelpful.

An ability to make decisions for navigating this future, and an ability to make these decisions in time, will depend on quickly reducing uncertainties about consequences of AI. The stakes for a better understanding are high. Those responsible for international security contend with the dissonance of the possibility that the future may not be much different on one hand, and the undermining of foundational assumptions for stability on the other. To the extent that great uncertainty is leading to inaction, it may have created a period of false stability. If the scientific community itself does not know if AI will be capable of quickly creating advances that disrupt present defense, or if the defense community cannot determine if AI will provide decisive advantage in strategic planning, costly actions to adapt to changes seem premature. The sooner we can resolve substantive uncertainties, the more time there is for measures to preserve stability.

It will take a coordinated, large-scale effort to build the needed understanding. The AI laboratories can make valuable contributions by reducing technical uncertainties that now lead to such wide variation in assessments for the implications of AI. While AI companies are not responsible for international security, they play an essential role in providing the technical foundations leaders will need to meet this challenge, and in shepherding technologies in ways that makes this possible.

A better understanding need not hinge on predicting the future of AI capabilities. The point was made many times in our interviews that the technology available today already has the potential to reshape international security, we just need deeper study in targeted areas. In fact, at this point it can be argued that clarification of timescales is not very important from the perspective of needs for action. Much of the argument about when we will see much more capable AI has now collapsed to the difference between two years from now and ten years from now. Both are very short compared to the pace of changes in institutions for international security.

The most significant near-term challenge is that we have no coherent programs or organizational structures to support such broad coordination. Encouragingly, the seeds of this work are beginning to appear across the AI ecosystem. Stanford has run workshops examining the intersection of AI and international security. The Special

Competitive Studies Project has written extensively on the topic and hosted several conferences that bring together AI companies and government leaders. Several other groups - including Georgetown's CSET and RAND - produce outstanding analysis, and Anthropic's fellows program incubates interdisciplinary expertise. At OpenAI, we have started a conference series aimed at convening a research community focused on anticipating and addressing the major challenges of the next century. But no single organization can meet the scale of what is needed. Progress will depend on a diverse coalition of researchers, practitioners, and institutions willing to collaborate on building the path forward.

Acknowledgements

In the course of drafting this paper, we engaged senior leaders in the fields of international security, economics, and science. We are grateful for their time and have tried to honor their expertise and lessons they have shared with us. Any errors in this paper are ours alone. For speaking with us, we would like thank Marv Adams, former head of Defense Programs in NNSA and Texas A&M professor; Ylber Bajraktari, Senior Advisor, Special Competitive Studies Project (SCSP); Sam Brannen, former Deputy Assistant Secretary of Defense for Plans and Posture; Aaron "Ronnie" Chatterji, Chief Economist, OpenAI and Mark Burgess and Lisa Benson-Burgess Distinguished Professor of Business and Public Policy at Duke University's Fuqua School of Business; Richard Danzig, Senior Fellow, Johns Hopkins Applied Physics Laboratory, and former Secretary of the Navy; Kevin Dixon, Senior Director at Sandia National Laboratories; Jon Finer, former Deputy National Security Advisor, The White House; Christine Fox, Senior Fellow, Johns Hopkins Applied Physics Laboratory, former Deputy Secretary of Defense; Pat Fitch, Deputy Director for Science and Global Security, Los Alamos National Laboratory; Michael Horowitz, Director, Perry World House, University of Pennsylvania, and former Deputy Assistant Secretary of Defense for Force Development and Emerging Capabilities; Colin Kahl, Senior Fellow at the Freeman Spogli Institute for International Studies, Stanford University, and former Under Secretary of Defense for Policy; Thom Mason, Director, Los Alamos National Laboratory; Jason Matheny, President and CEO, RAND Corporation, former Director of Intelligence Advanced Research Projects Activity (IARPA); General (ret) H. R. McMaster, Senior Fellow, Hoover Institution, Stanford University and former U.S. National Security Advisor; Anne Neuberger, Payne Distinguished Lecturer at Stanford University and former Deputy National Security Advisor for Cyber and Emerging Technology, The White House; General Jack Shanahan, former Director, U.S. Department of Defense Joint Artificial Intelligence Center (JAIC); Neil Thompson, Research Scientist, MIT Sloan School of Management and MIT Computer Science and AI Lab; and Bob Webster, Deputy Director, Los Alamos National Laboratory. We would also like to thank our colleagues Jackie Hehir, Richard Johnson, Alexis Bonnell, Katrina Mulligan, Caroline Zier, Connie LaRossa, Harrison Satcher, Kevin Weil, and Ludovic Peran of OpenAI, as well as Chip Usher at SCSP, for reading the manuscript and providing valuable feedback.

References

[1] Stanford University Institute for Human-Centered Artificial Intelligence. *Artificial Intelligence Index Report 2025*. Tech. rep. 2025. URL: https://hai.stanford.edu/ai-index.

[2] National Security Commission on Artificial Intelligence. *Final Report*. Mar. 2021. URL: https://reports.nscai.gov/.

[3] United Nations. *Governing AI for Humanity: Final Report*. Tech. rep. Final report of the UN Secretary-General's High-level Advisory Body on Artificial Intelligence. Accessed: 2026-01-22. United Nations, Sept. 2024.

[4] OpenAI. *Preparedness Framework*. Technical report. Version 2. Apr. 2025. URL: https://cdn.openai.com/pdf/18a02b5d-6b67-4cec-ab64-68cdfbddebcd/preparedness-framework-v2.pdf.

[5] Anthropic. *Responsible Scaling Policy*. Policy document. Version 2.2. May 2025. URL: https://www.anthropic.com/responsible-scaling-policy.

[6] A. J. Lohn. *Anticipating AI's Impact on the Cyber Offense–Defense Balance*. Washington, DC: Center for Security and Emerging Technology (CSET), Georgetown University, May 2025. URL: https://cset.georgetown.edu/publication/anticipating-ais-impact-on-the-cyber-offense-defense-balance/.

[7] M. C. Horowitz. "Artificial Intelligence, International Competition, and the Balance of Power". In: *Texas National Security Review* 1.3 (2018), pp. 36–57.

[8] P. Scharre. *Army of None: Autonomous Weapons and the Future of War*. New York: W. W. Norton & Company, 2018.

[9] M. C. Horowitz. "Battles of Precise Mass: Technology Is Remaking War—and America Must Adapt". In: *Foreign Affairs* (Oct. 2024).

[10] R. Danzig. *Artificial Intelligence, Cybersecurity, and National Security: The Fierce Urgency of Now*. PE-A4079-1. Santa Monica, CA: RAND Corporation, July 2025. DOI: 10.7249/PEA4079-1. URL: https://www.rand.org/pubs/perspectives/PEA4079-1.html.

[11] P. Kennedy. *The Rise and Fall of the Great Powers: Economic Change and Military Conflict from 1500 to 2000*. New York: Random House, 1987.

[12] R. T. Gould. *The Marine Chronometer: Its History and Development*. London: J. D. Potter, 1923.

[13] D. Showalter. "Soldiers into Postmasters? The Electric Telegraph as an Instrument of Command in the Prussian Army". In: *Military Affairs* 37.2 (Apr. 1973), pp. 48–52.

[14] D. R. Headrick. *The Tools of Empire: Technology and European Imperialism in the Nineteenth Century*. New York: Oxford University Press, 1981.

[15] J. White Lynn. *Medieval Technology and Social Change*. Oxford: Clarendon Press, 1962.

[16] B. Brodie. *Strategy in the Missile Age*. Undertaken by the RAND Corporation as part of its research program for the United States Air Force. Princeton, NJ: Princeton University Press, 1959. URL: https://www.rand.org/content/dam/rand/pubs/commercial_books/2007/RAND_CB137-1.pdf.

[17] F. Kaplan. *The Wizards of Armageddon*. New York: Simon and Schuster, 1983.

[18] D. E. Hoffman. *The Dead Hand: The Untold Story of the Cold War Arms Race and Its Dangerous Legacy*. New York: Doubleday, 2009.

[19] E. Schlosser. *Command and Control: Nuclear Weapons, the Damascus Accident, and the Illusion of Safety*. New York: Penguin Press, 2013.

[20] J. Black, M. Eken, J. Parakilas, S. Dee, C. Ellis, K. Suman-Chauhan, R. Bain, H. Fine, M. C. Aquilino, M. Lebret, and O. Palicki. *Strategic Competition in the Age of AI: Emerging Risks and Opportunities from Military Use of Artificial Intelligence*. Tech. rep. RAND Europe / RAND Corporation (commissioned by UK MOD and FCDO), Oct. 2024. URL: https://assets.publishing.service.gov.uk/media/6703f5ec080bdf716392ef44/Strategic_competition_in_the_age_of_AI.pdf.

[21] J. Shanahan. *Artificial Intelligence and Geopolitics: Hitching the Disruptive Technology Cart to the Geopolitics Horse*. essay. 2023.

[22] N. Wright, M. Miklaucic, and T. Veazie, eds. *Human, Machine, War: How the Mind-Tech Nexus Will Win Future Wars*. Maxwell AFB, AL: Air University Press, Apr. 2025. URL: https://www.airuniversity.af.edu/AUPress/Display/Article/4162241/human-machine-war-how-the-mind-tech-nexus-will-win-future-wars/.

[23] N. Wright. *Warhead: How the Brain Shapes War and War Shapes the Brain*. London and New York: Pan Macmillan and St. Martin's Press, Oct. 2025.

[24] Y. LeCun, Y. Bengio, and G. Hinton. "Deep Learning". In: *Nature* 521.7553 (2015), pp. 436–444. DOI: 10.1038/nature14539.

[25] S. Minaee, T. Mikolov, N. Nikzad, M. Chenaghlu, R. Socher, X. Amatriain, and J. Gao. *Large Language Models: A Survey*. Feb. 2024. arXiv: 2402.06196 [cs.CL]. URL: https://arxiv.org/abs/2402.06196.

[26] Y. Bengio, P. Lamblin, D. Popovici, and H. Larochelle. "Greedy Layer-Wise Training of Deep Networks". In: *Advances in Neural Information Processing Systems*. Ed. by B. Schölkopf, J. Platt, and T. Hofmann. Vol. 19. Cambridge, MA: MIT Press, 2006, pp. 153–160. URL: https://proceedings.neurips.cc/paper/3048-greedy-layer-wise-training-of-deep-networks.pdf.

[27] A. Krizhevsky, I. Sutskever, and G. E. Hinton. "ImageNet Classification with Deep Convolutional Neural Networks". In: *Advances in Neural Information Processing Systems*. Vol. 25. 2012, pp. 1097–1105. URL: https://proceedings.neurips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf.

[28] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis. "Human-level Control through Deep Reinforcement Learning". In: *Nature* 518.7540 (Feb. 2015), pp. 529–533. DOI: 10.1038/nature14236.

[29] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin. "Attention Is All You Need". In: *Advances in Neural Information Processing Systems*. Vol. 30. 2017, pp. 5998–6008. URL: https://papers.neurips.cc/paper/7181-attention-is-all-you-need.pdf.

[30] R. Bommasani, D. A. Hudson, E. Adeli, R. Altman, S. Arora, S. von Arx, M. S. Bernstein, et al. *On the Opportunities and Risks of Foundation Models*. Tech. rep. Stanford University, Center for Research on Foundation Models (CRFM), 2021. arXiv: 2108.07258. URL: https://arxiv.org/abs/2108.07258.

[31] D. Halawi, F. Zhang, Y.-H. Chen, and J. Steinhardt. "Approaching Human-Level Forecasting with Language Models". In: *Advances in Neural Information Processing Systems*. Vol. 37. 2024. URL: https://proceedings.neurips.cc/paper_files/paper/2024/file/5a5acfd0876c940d81619c1dc60e7748-Paper-Conference.pdf.

[32] J. Lu. *Evaluating LLMs on Real-World Forecasting Against Human Superforecasters*. 2025. arXiv: 2507.04562 [cs.CL]. URL: https://arxiv.org/abs/2507.04562.

[33] OpenAI. *How people are using ChatGPT*. Sept. 2025. URL: https://openai.com/index/how-people-are-using-chatgpt/.

[34] C. Cortes, L. D. Jackel, S. A. Solla, V. Vapnik, and J. S. Denker. "Learning Curves: Asymptotic Values and Rate of Convergence". In: *Advances in Neural Information Processing Systems*. Vol. 6. 1994, pp. 327–334. URL: https://proceedings.neurips.cc/paper/1993/hash/1aa48fc4880bb0c9b8a3bf979d3b917e-Abstract.html.

[35] J. Sevilla and E. Roldán. *Training compute of frontier AI models grows by 4–5× per year*. May 2024. URL: https://epoch.ai/blog/training-compute-of-frontier-ai-models-grows-by-4-5x-per-year.

[36] J. Kaplan, S. McCandlish, T. Henighan, T. B. Brown, B. Chess, R. Child, S. Gray, A. Radford, J. Wu, and D. Amodei. *Scaling Laws for Neural Language Models*. 2020. arXiv: 2001.08361 [cs.LG]. URL: https://arxiv.org/abs/2001.08361.

[37] OpenAI. *Introducing OpenAI o1-preview*. Sept. 2024. URL: https://openai.com/index/introducing-openai-o1-preview/.

[38] *The Artificial General Intelligence Race and International Security*. PE-A4155-1. RAND Corporation, Sept. 2025. DOI: 10.7249/PEA4155-1. URL: https://www.rand.org/pubs/perspectives/PEA4155-1.html.

[39] Z. Burdette, D. Phillips, J. L. Heim, E. Geist, D. R. Frelinger, C. Heitzenrater, and K. P. Mueller. *How Artificial Intelligence Could Reshape Four Essential Competitions in Future Warfare*. RR-A4316-1. RAND Corporation, 2026. DOI: 10.7249/RRA4316-1. URL: https://www.rand.org/pubs/research_reports/RRA4316-1.html.

[40] Y. N. Harari. *Nexus: A Brief History of Information Networks from the Stone Age to AI*. New York: Random House, 2024.

[41] L. Aschenbrenner. *Situational Awareness: The Decade Ahead*. Online essay series. June 2024. URL: https://situational-awareness.ai/.

[42] D. Kokotajlo, S. Alexander, T. Larsen, E. Lifland, and R. Dean. *AI 2027*. Scenario report. Apr. 2025. URL: https://ai-futures.org/.

[43] D. Hendrycks, E. Schmidt, and A. Wang. *Superintelligence Strategy: Expert Version*. Online report. 2025. URL: https://www.nationalsecurity.ai/.

[44] J. Matheny. *Challenges to U.S. National Security and Competitiveness Posed by Artificial Intelligence*. Testimony before the Senate Committee on Homeland Security and Governmental Affairs. Mar. 2023. URL: https://www.hsgac.senate.gov/.

[45] National Academies of Sciences, Engineering, and Medicine. *Biodefense in the Age of Synthetic Biology*. Washington, DC: National Academies Press, 2018. DOI: 10.17226/24890.

[46] A. F. Krepinevich. *The Origins of Victory: How Disruptive Military Innovation Determines the Fates of Great Powers*. New Haven, CT: Yale University Press, 2023.

[47] V. Bush. *Science—The Endless Frontier*. Report to the President of the United States. 1945.

[48] Congressional Research Service. *Government Expenditures on Defense Research and Development by the United States and Other OECD Countries: Fact Sheet*. Tech. rep. R45441. Washington, DC: Congressional Research Service, 2018. URL: https://crsreports.congress.gov/product/pdf/R/R45441.

[49] U.S. Department of Defense, Office of the Under Secretary of Defense (Comptroller). *Defense Budget Overview: Fiscal Year 2025 Budget Request*. Budget overview book. 2024. URL: https://comptroller.defense.gov/.

[50] NATO Science & Technology Organization. *Science & Technology Trends: 2023–2043*. Brussels: NATO STO, 2023.

[51] United Nations Institute for Disarmament Research. *Quantum Technology, Peace and Security: A Primer*. 2024. URL: https://unidir.org/publication/quantum-technology-peace-and-security-a-primer.

[52] M. J. Mazarr, T. R. Heath, A. S. Cevallos, A. Radin, and G. Casey. *Disrupting Deterrence: Examining the Effects of Emerging Technologies on Strategic Deterrence and Stability*. Santa Monica, CA: RAND Corporation, 2022. URL: https://www.rand.org/pubs/research_reports/RRA595-1.html.

[53] *Black Sea battle: how Ukraine's drones overpowered the Russian navy*. 2025. URL: https://www.navylookout.com/black-sea-battle-how-ukraines-drones-overpowered-the-russian-navy/.

[54] M. C. Horowitz. *Ukraine's Operation Spider's Web Shows Future of Drone Warfare*. 2025. URL: https://www.cfr.org/article/ukraines-operation-spiders-web-shows-future-drone-warfare.

[55] Z. Bai, Y. Wang, C. Wang, C. Yu, D. Lukyanenko, I. Stepanova, and A. G. Yagola. "Joint Gravity and Magnetic Inversion Using CNNs' Deep Learning". In: *Remote Sensing* 16.7 (2024), p. 1115. DOI: 10.3390/rs16071115.

[56] D. Wines, T. Xie, and K. Choudhary. "Inverse Design of Next-Generation Superconductors Using Data-Driven Deep Generative Models". In: *The Journal of Physical Chemistry Letters* 14.29 (2023), pp. 6630–6638. DOI: 10.1021/acs.jpclett.3c01260.

[57] X.-Q. Han, Z. Ouyang, P.-J. Guo, H. Sun, Z.-F. Gao, and Z.-Y. Lu. *InvDesFlow: An AI Search Engine to Explore Possible High-Temperature Superconductors*. 2024. arXiv: 2409.08065 [cond-mat.supr-con]. URL: https://arxiv.org/abs/2409.08065.

[58] M. Krelina. *Military and Security Dimensions of Quantum Technologies: A Primer*. Tech. rep. Stockholm: Stockholm International Peace Research Institute (SIPRI), July 2025. URL: https://www.sipri.org/publications/2025/other-publications/military-and-security-dimensions-quantum-technologies-primer.

[59] Y. Zhang, S. A. Khan, A. Mahmud, H. Yang, A. Lavin, M. Levin, J. Frey, J. Dunnmon, J. Evans, A. Bundy, S. Džeroski, J. Tegner, and H. Zenil. "Exploring the role of large language models in the scientific method: from hypothesis to discovery". In: *npj Artificial Intelligence* 1 (2025), p. 14. DOI: 10.1038/s44387-025-00019-5.

[60] K. Duraisamy. *Active Inference AI Systems for Scientific Discovery*. 2025. arXiv: 2506.21329 [cs.AI]. URL: https://arxiv.org/abs/2506.21329.

[61] A. R. Doshi and O. P. Hauser. "Generative AI Enhances Individual Creativity but Reduces the Collective Diversity of Novel Content". In: *Science Advances* 10.28 (2024). DOI: 10.1126/sciadv.adn5290.

[62] J. A. Byrne et al. "A Call for Research to Address the Threat of Paper Mills". In: *PLOS Biology* 22.11 (2024), e3002931. DOI: 10.1371/journal.pbio.3002931.

[63] A. W. Ding and S. Li. "Generative AI lacks the human creativity to achieve scientific discovery from scratch". In: *Scientific Reports* 15 (2025), p. 9587. DOI: 10.1038/s41598-025-93794-9.

[64] J. Sourati and J. A. Evans. "Accelerating Science with Human-Aware Artificial Intelligence". In: *Nature Human Behaviour* 7.10 (2023), pp. 1682–1696. DOI: 10.1038/s41562-023-01648-z.

[65] H. Wang, T. Fu, Y. Du, et al. "Scientific Discovery in the Age of Artificial Intelligence". In: *Nature* 620.7972 (2023), pp. 47–60. DOI: 10.1038/s41586-023-06221-2.

[66] B. Romera-Paredes, M. Barekatain, A. Novikov, M. Balog, M. P. Kumar, E. Dupont, F. J. R. Ruiz, P. Wang, P. Kohli, A. Fawzi, J. S. Ellenberg, and O. Fawzi. "Mathematical discoveries from program search with large language models". In: *Nature* 625.7995 (2024), pp. 468–475. DOI: 10.1038/s41586-023-06924-6. URL: https://www.nature.com/articles/s41586-023-06924-6.

[67] D. A. Boiko, R. MacKnight, B. Kline, and G. Gomes. "Autonomous chemical research with large language models". In: *Nature* 624 (2023), pp. 570–578. DOI: 10.1038/s41586-023-06792-0.

[68] X. Luo, A. Rechardt, G. Sun, et al. "Large language models surpass human experts in predicting neuroscience results". In: *Nature Human Behaviour* 9 (2025), pp. 305–315. DOI: 10.1038/s41562-024-02046-9.

[69] T. Besiroglu, N. Emery-Xu, and N. Thompson. "Economic impacts of AI-augmented R&D". In: *Research Policy* 53.7 (2024). DOI: 10.1016/j.respol.2024.105037.

[70] S. Bubeck et al. *Early science acceleration experiments with GPT-5*. Nov. 2025. arXiv: 2511.16072 [cs.AI]. URL: https://arxiv.org/abs/2511.16072.

[71] M. C. Horowitz, G. C. Allen, E. Saravalle, A. Cho, K. Frederick, and P. Scharre. *Artificial Intelligence and International Security*. Washington, DC: Center for a New American Security, July 2018. URL: https://www.cnas.org/publications/reports/artificial-intelligence-and-international-security.

[72] E. Miguel, S. Satyanath, and E. Sergenti. "Economic Shocks and Civil Conflict: An Instrumental Variables Approach". In: *Journal of Political Economy* 112.4 (Aug. 2004), pp. 725–753. DOI: 10.1086/421174.

[73] M. Gerlich. "Brace for Impact: Facing the AI Revolution and Geopolitical Shifts in a Future Societal Scenario for 2025–2040". In: *Societies* 14.9 (2024), p. 180. DOI: 10.3390/soc14090180.

[74] U.S. Department of Energy. *Critical Minerals: Roles for Artificial Intelligence in Support of FECM RDD&D Priorities*. Tech. rep. Mar. 2023. URL: https://www.energy.gov/sites/default/files/2023-03/ai-role-in-critical-minerals.pdf.

[75] S. Itani, Y. Zhang, and J. Zang. "The northeast materials database for magnetic materials". In: *Nature Communications* 16 (2025), p. 9415. DOI: 10.1038/s41467-025-64458-z. URL: https://doi.org/10.1038/s41467-025-64458-z.

[76] J. W. Erisman, M. A. Sutton, J. Galloway, Z. Klimont, and W. Winiwarter. "How a Century of Ammonia Synthesis Changed the World". In: *Nature Geoscience* 1 (2008), pp. 636–639. DOI: 10.1038/ngeo325.

[77] State Council of the People's Republic of China. *New Generation Artificial Intelligence Development Plan*. Beijing, July 2017. URL: https://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm.

[78] HM Government. *National AI Strategy*. London, Sept. 2021. URL: https://www.gov.uk/government/publications/national-ai-strategy.

[79] *Removing Barriers to American Leadership in Artificial Intelligence*. Executive Order 14179. Jan. 2025. URL: https://www.govinfo.gov/app/details/FR-2025-01-31/2025-02172.

[80] The White House. *Winning the AI Race: America's AI Action Plan*. White House policy document. Washington, D.C., July 2025. URL: https://www.whitehouse.gov/wp-content/uploads/2025/07/Americas-AI-Action-Plan.pdf.

[81] J. Sevilla, T. Besiroglu, B. Cottier, J. You, E. Roldán, P. Villalobos, and E. Erdil. *Can AI Scaling Continue Through 2030?* Aug. 2024. URL: https://epoch.ai/blog/can-ai-scaling-continue-through-2030.

[82] International Energy Agency. *Energy and AI*. Paris, Apr. 2025. URL: https://www.iea.org/reports/energy-and-ai.

[83] E. Erdil, A. Potlogea, T. Besiroglu, E. Roldan, A. Ho, J. Sevilla, M. Barnett, M. Vrzla, and R. Sandler. *GATE: An Integrated Assessment Model for AI Automation*. 2025. arXiv: 2503.04941 [econ.GN]. URL: https://arxiv.org/abs/2503.04941.

[84] T. C. Schelling. *Arms and Influence*. New Haven: Yale University Press, 1966.

[85] A. Narayanan and S. Kapoor. *AI as Normal Technology*. Apr. 2025. URL: https://knightcolumbia.org/content/ai-as-normal-technology.

[86] S. Armstrong, N. Bostrom, and C. Shulman. "Racing to the Precipice: A Model of Artificial Intelligence Development". In: *AI & Society* 31.2 (2016), pp. 201–206. DOI: 10.1007/s00146-015-0590-y.

[87] Anthropic. *Disrupting the First Reported AI-Orchestrated Cyber Espionage Campaign*. Nov. 2025. URL: https://www.anthropic.com/news/disrupting-AI-espionage.

[88] E. Ho, A. Rajagopalan, A. Skvortsov, S. Arulampalam, and M. Piraveenan. "Game Theory in Defence Applications: A Review". In: *Sensors* 22.3 (2022), p. 1032. DOI: 10.3390/s22031032.

[89] S. Yao, D. Yu, J. Zhao, I. Shafran, T. L. Griffiths, Y. Cao, and K. Narasimhan. *Tree of Thoughts: Deliberate Problem Solving with Large Language Models*. 2023. arXiv: 2305.10601 [cs.CL]. URL: https://arxiv.org/abs/2305.10601.

[90] C. F. Camerer, T.-H. Ho, and J.-K. Chong. "A Cognitive Hierarchy Model of Games". In: *Quarterly Journal of Economics* 119.3 (2004), pp. 861–898. DOI: 10.1162/0033553041502225.

[91] R. Jervis. *Perception and Misperception in International Politics*. Princeton, NJ: Princeton University Press, 1976.

[92] K. A. Lieber and D. G. Press. "The New Era of Counterforce: Technological Change and the Future of Nuclear Deterrence". In: *International Security* 41.4 (2017), pp. 9–49. DOI: 10.1162/ISEC_a_00273.

[93] M. C. Horowitz and P. Scharre. *AI and International Stability: Risks and Confidence-Building Measures*. Tech. rep. Center for a New American Security (CNAS), Jan. 2021. URL: https://www.cnas.org/publications/reports/ai-and-international-stability-risks-and-confidence-building-measures.

[94] V. Chernavskikh and J. Palayer. *Impact of Military Artificial Intelligence on Nuclear Escalation Risk*. Tech. rep. Stockholm International Peace Research Institute (SIPRI), June 2025. DOI: 10.55163/FZIW8544. URL: https://www.sipri.org/publications/2025/sipri-insights-peace-and-security/impact-military-artificial-intelligence-nuclear-escalation-risk.

[95] U.S. Department of Defense. *Summary of the Joint All-Domain Command & Control (JADC2) Strategy*. Tech. rep. U.S. Department of Defense, Mar. 2022. URL: https://media.defense.gov/2022/Mar/17/2002958406/-1/-1/1/SUMMARY-OF-THE-JOINT-ALL-DOMAIN-COMMAND-AND-CONTROL-STRATEGY.pdf.

[96] B. R. Green and A. Long. "Conceal or Reveal? Managing Clandestine Military Capabilities in Peacetime Competition". In: *International Security* 44.3 (2020), pp. 48–83. DOI: 10.1162/isec_a_00367.

[97] *Classification Levels*. 2025. URL: https://www.dami.army.pentagon.mil/site/infosec/TP-levels.aspx.

[98] Y.-A. de Montjoye, C. A. Hidalgo, M. Verleysen, and V. D. Blondel. "Unique in the Crowd: The Privacy Bounds of Human Mobility". In: *Scientific Reports* 3 (2013), p. 1376. DOI: 10.1038/srep01376. URL: https://www.nature.com/articles/srep01376.

[99] L. Rocher, J. M. Hendrickx, and Y.-A. de Montjoye. "Estimating the Success of Re-identifications in Incomplete Datasets Using Generative Models". In: *Nature Communications* 10 (2019), p. 3069. DOI: 10.1038/s41467-019-10933-3. URL: https://www.nature.com/articles/s41467-019-10933-3.

[100] M. Binz, Z. Akata, M. Bethge, et al. "A Foundation Model to Predict and Capture Human Cognition". In: *Nature* 644.8078 (2025), pp. 1002–1009. DOI: 10.1038/s41586-025-09215-4. URL: https://www.nature.com/articles/s41586-025-09215-4.

[101] J. H. Ellis. *The History of Non-Secret Encryption*. Tech. rep. Cheltenham: Communications-Electronics Security Group (CESG), Government Communications Headquarters (GCHQ), 1987.

[102] National Institute of Standards and Technology. *What Is Post-Quantum Cryptography?* Aug. 2024. URL: https://www.nist.gov/cybersecurity/what-post-quantum-cryptography.

[103] A. Novikov, N. Vũ, M. Eisenberger, E. Dupont, P.-S. Huang, A. Z. Wagner, S. Shirobokov, B. Kozlovskii, F. J. R. Ruiz, A. Mehrabian, M. P. Kumar, A. See, S. Chaudhuri, G. Holland, A. Davies, S. Nowozin, P. Kohli, and M. Balog. *AlphaEvolve: A coding agent for scientific and algorithmic discovery*. 2025. arXiv: 2506.13131 [cs.AI]. URL: https://arxiv.org/abs/2506.13131.