CMSC 848Y: Paper Response I

Why Do States Build AI? Sagan's Framework and the Modern AI Landscape

Minhal Khan

The standard account of international competition over artificial intelligence is largely predicated on security concerns. States pursue frontier AI because their rivals do, and whoever falls behind loses. Putin's 2017 claim that the leader in AI "will become the ruler of the world" (CNBC, 2017) and Trump's recent clamor of "WHOEVER WINS AI, WINS!" (Trump, 2026) both encapsulate this idea, revealing the same logic underlying national AI strategies, export controls on chips, and the arms-race framing that the Horowitz reading examines. This paper argues that the security explanation deserves the treatment Scott Sagan gave its nuclear counterpart thirty years ago. Sagan showed that the security model for why states build the bomb, though often applicable, was "dangerously inadequate because nuclear weapons programs also serve other, more parochial and less obvious objectives" (Sagan, 1996, p. 55). As bureaucratic interests and status-seeking also drove nuclear proliferation, so too do they drive AI development today. Prima facie, Sagan's three models transfer readily from nuclear weapons to AI not because they are models of a weapon, but because they explain why any state acquires costly, strategic, and symbolically loaded technologies.

Sagan's framework is, of course, not the only one available. Horowitz's adoption capacity theory explains how military innovations spread and which organizations can adopt them, and he notes that commercial incentives drive much of AI development. His question, however, is one of adoption rather than demand, and where he addresses organizational politics at all, it is as a brake on adoption rather than a source of it. Sagan asks whether a state's *purported* reason for wanting

the technology is its *genuine* reason and whether the actors pushing for it within the state have interests of their own. This distinction matters for policy. If national security genuinely drives AI proliferation, traditional tools like deterrence and export controls can effectively address this cause. However, if AI development is instead propelled by internal pressures (such as private-sector lobbying or the pursuit of international prestige), these outward-facing security measures will fail to address the true sources of demand.

According to the neorealist logic underpinning Sagan's security model, states in an anarchic international system must rely on themselves to ensure their survival, thus transforming the pursuit of frontier AI into a strategic imperative driven by external threats. Just as nuclear weapons triggered a "strategic chain reaction" (Sagan, 1996, p. 58) among rival powers, breakthroughs in AI research provoke similar balancing behaviors today. Because AI functions as a broad "enabling" (Horowitz, 2018, p. 41) technology akin to electricity rather than a standalone weapon, it is integrated into various military systems to address specific vulnerabilities and ensure nations are not outpaced by their adversaries. Russia, for instance, is developing autonomous vehicles to guard missile bases (Hambling, 2014) and designing a nuclear-capable autonomous submarine (Horowitz, 2018, p. 45). Along contested borders, states employ AI to offset geographic or numerical disadvantages; South Korea developed the SGR-A1 semi-autonomous system to protect the Demilitarized Zone (Kumagai, 2007), while Israel leverages AI to predict rocket launches and rapidly analyze surveillance footage (Agence France-Presse, 2026). Furthermore, wealthy middle powers like Australia view AI as a critical investment to overcome the military limitations of smaller populations and recruiting shortages (Horowitz, 2018, p. 46; Ryan, 2017). These diverse global applications illustrate that states pursue AI not

necessarily for grand offensive ambitions, but as a mandatory tool for self-preservation and to even the odds in the global arena.

In contrast to the survival-driven logic of the security model, Sagan's norms model posits that technological acquisition serves profound symbolic functions, shaping and reflecting a state's identity on the global stage. Powerful technologies are thus pursued because they endow a state with a sense of modernity, unity, and legitimate statehood. One striking contemporary example comes from France, with President Emmanuel Macron's championing of domestic AI startup Mistral as a state vehicle for technological sovereignty and paving a "third way" in global AI development to challenge the duopoly of U.S. and Chinese tech (Chavez, 2025). Similar homegrown efforts can be seen in the Nordic sphere with Finland's Viking models (Silo AI, 2024), Sweden's GPT-SW3 (AI Sweden, 2023), and Norway's Borealis (National Library of Norway AI Lab, 2026). Even more interesting is the push to sustain and nourish Nordic values and customs *through* these models, as the Nordic Council of Ministers announced a joint effort in 2025 to "ensure that the development of language models reflects the Nordic countries' unique linguistic and cultural traits" (Nordic Council of Ministers, 2025). This is perhaps best demonstrated by the fact that Borealis and its variations are specifically fine-tuned for Norwegian language standards such as Bokmål and Nynorsk.

Yet, neither the fear of foreign adversaries nor the pursuit of global prestige fully captures why states invest so heavily in AI. A more complete picture unfolds when we look at the domestic level and the actors therein. Sagan's domestic politics model posits that nuclear programs were driven by the parochial interests of scientists, military units, and politicians, functioning as "solutions looking for a problem to which to attach themselves so as to justify their existence" (Sagan, 1996, p. 65). Shifting this framework towards artificial intelligence, however, exposes a

key discrepancy: nuclear weapons were not the products of private firms, while the most formidable AI models of today decidedly are. Because artificial intelligence has vast, highly profitable civilian applications, the push for algorithmic supremacy is overwhelmingly driven by commercial incentives rather than state-directed military research.

Development has run ahead of the state, not behind it. OpenAI and Anthropic, two of the largest players in the AI space, were founded relatively recently and without state direction, and ChatGPT was a viral consumer product long before it became a matter of national policy. The scale of this commercial engine is staggering; private AI investment in the United States reached $109 billion in 2024, roughly twelve times that of China (Maslej et al., 2025). Indeed, commercial demand for talent is intense enough that, as Horowitz notes, some argue it holds back military development and may force militaries to turn to civilian firms (Horowitz, 2018, p. 50). Consequently, it is private firms pushing and pulling against one another—competing fiercely for market share, talent, and capital—that forms the truest, fastest, and most chaotic engine of global AI proliferation.

The national security rationale for AI thus arises *after the fact*. Eric Schmidt, for instance, chaired the National Security Commission on Artificial Intelligence, whose 2021 report framed AI almost entirely through the lens of strategic competition with China (Schmidt et al., 2021). Today, Silicon Valley giants and firms like Palantir, Anduril, and OpenAI maneuver aggressively for lucrative defense contracts, mirroring the kind of integration Project Maven pioneered. None of these state-sponsored initiatives caused frontier AI to exist in the first place, thereby reflecting the mechanism Sagan described: a powerful domestic establishment, possessing interests and capabilities of its own, retroactively finding a national security problem to justify and expand its existence.

The heavy involvement of private firms does not invalidate Sagan's domestic politics model; instead, it represents its evolution in the 21st century. Sagan explicitly included civilian scientific institutions and mining industries within his definition of the domestic coalitions that drive technological proliferation. While the presence of massive, independent tech oligarchies is predominantly an American phenomenon, the commercial outputs of these firms clearly dictate global dynamics. Because AI functions as a dual-use enabling technology, commercial breakthroughs rapidly diffuse across borders, forcing the rest of the international system to respond. Rivals like China adapt to this pressure by explicitly partnering their military apparatus with their own private enterprises, while middle powers without frontier tech giants are structurally compelled to navigate a technological environment shaped largely by foreign commercial exports. The sheer volume of private capital concentrated in a few commercial hubs accelerates global AI proliferation faster than any single government-led military program could.

Applying Sagan's framework to artificial intelligence reveals that international technological competition extends far beyond security dilemmas. While the security model explains the geopolitical arms-race rhetoric, and the norms model highlights the performative pursuit of national prestige, the domestic model arguably exposes the true underlying engine of the AI revolution: commercial enterprise. To effectively navigate this new era of global competition, states must recognize that the balance of power is increasingly dictated not just by foreign adversaries, but by the private tech companies, globalized diffusion, and commercial incentives operating from within.

References

Agence France-Presse. (2026, April 3). Israel using AI to fine-tune air raid alert system. *Courthouse News Service*. https://www.courthousenews.com/israel-using-ai-to-fine-tune-air-raid-alert-system

AI Sweden. (2023, November). *Open release of the first large Nordic language model GPT-SW3 | AI Sweden*. https://www.ai.se/en/news/open-release-first-large-nordic-language-model-gpt-sw3

Chavez, P. (2025, February 13). France Pursues an AI "Third Way." *CEPA*. https://cepa.org/article/france-pursues-an-ai-third-way/

Hambling, D. (2014, April 23). Armed Russian robocops to defend missile bases. *New Scientist*. https://www.newscientist.com/article/2001072-armed-russian-robocops-to-defend-missile-bases/

Horowitz, M. C. (2018, May 15). Artificial Intelligence, International Competition, and the Balance of Power. *Texas National Security Review*. https://tnsr.org/2018/05/artificial-intelligence-international-competition-and-the-balance-of-power/

Kumagai, J. (2007, March). *A Robotic Sentry For Korea's Demilitarized Zone—IEEE Spectrum*. https://spectrum.ieee.org/a-robotic-sentry-for-koreas-demilitarized-zone

Maslej, N., Fattorini, L., Perrault, R., Gil, Y., Parli, V., Kariuki, N., Capstick, E., Reuel, A., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Niebles, J. C., Shoham, Y., Wald, R., Walsh, T., Hamrah, A., Santarlasci, L., … Oak, S. (2025). *Artificial Intelligence Index Report 2025* (arXiv:2504.07139). arXiv. https://doi.org/10.48550/arXiv.2504.07139

National Library of Norway AI Lab. (2026, May). *A new family of Norwegian-centric models – AI-lab*. https://ai.nb.no/borealis/

Nordic Council of Ministers. (2025, October). *Nordic ministers for culture invest in joint network for AI language models*. https://www.norden.org/en/news/nordic-ministers-culture-invest-joint-network-ai-language-models

Putin: Leader in artificial intelligence will rule world. (2017, September 4). CNBC. https://www.cnbc.com/2017/09/04/putin-leader-in-artificial-intelligence-will-rule-world.html

Ryan, M. (2017, December 11). *Building a Future: Integrated Human-Machine Military Organization*. The Strategy Bridge. https://thestrategybridge.org/the-bridge/2017/12/11/building-a-future-integrated-human-machine-military-organization

Sagan, S. D. (1996). Why Do States Build Nuclear Weapons?: Three Models in Search of a Bomb. *International Security*, *21*(3), 54–86. https://doi.org/10.2307/2539273

Schmidt, E., Work, B., Catz, S., Chien, S., Darby, C., Ford, K., Griffiths, J., Horvitz, E., Jassy, A., Mark, W., Matheny, J., McFarland, K., & Moore, A. (2021). *Final Report: National Security*

Commission on Artificial Intelligence (AI). (No. AD1124333). National Security Commission on Artificial Intellegence Washington DC. https://ntrl.ntis.gov/NTRL/dashboard/searchResults/titleDetail/AD1124333.xhtml

Silo AI. (2024). *Viking 7B: The First Open LLM for the Nordic Languages*. AMD. https://www.amd.com/en/blogs/2024/viking-7b-the-first-open-llm-for-the-nordic-languages.html

Trump, D. J. (2026, September). *Donald J. Trump (@realDonaldTrump)*. Truth Social. https://truthsocial.com/@realDonaldTrump/117269745153543631