

The Quest for Supremacy and Sagacity in AI

A Book Review of "Four Battlegrounds: Power in the Age of Artificial Intelligence" by Paul Scharre

Anushka Saxena*

The world is undergoing geopolitical flux, and rivalling quests for power have critical and emerging technologies at their core. Artificial Intelligence (AI) is one of them. And even as basic and applied research on AI continues to expand, how the technology works and why its deployment matters to the layperson remains comprehensible only to those sitting behind the golden doors of Silicon Valley, and occasionally, to some working behind the high walls of military-industrial complexes.

Through his latest book, *Four Battlegrounds: Power in the Age of Artificial Intelligence*, Paul Scharre manages to bring down these walls, bringing to the layperson a comprehensive analysis of the four battlegrounds where the war for supremacy in AI is being fought – data, computing power, talent, and institutions.

The research featured in the book, and its writing style, are methodological and conversational at the same time. Scharre has conducted extensive interviews with field experts, including both researchers and practitioners, and has analysed their first-hand accounts of key events that have shaped the American and Chinese AI industries.

For example, the author references his discussions with Liz O’Sullivan regarding the private sector’s sentiments against the deployment of AI in the US Department of Defense (DoD), which is akin to hearing testimony straight from the horse’s mouth. In 2019, O’Sullivan, a former employee of the AI firm ‘Clarifai’, sent a letter to her boss protesting against the company’s involvement in the US DoD’s AI programme. It set off the “revolt” in the American private industry against the use of AI in lethal weapons systems.

Similarly, to explain the AI-enabled panopticon the Chinese party-state has deployed to surveil the people of Xinjiang, Scharre has interviewed Maya Wang of Human Rights Watch, who broke the story about the use of China’s “algorithm of repression,” the Integrated Joint Operations Platform (IJOP).

In discussing China’s AI panopticon and surveillance ecosystem, Scharre covers not just the IJOP deployed in Xinjiang, but other ominous aspects that make the Chinese ecosystem resemble a techno-nationalistic, dystopian society, akin to those envisioned in sci-fi movies. From the nationwide

* Anushka Saxena is a Research Analyst with the Takshashila Institution’s Indo-Pacific Studies Programme.

deployment of 500 million facial recognition cameras and various ‘Police Cloud’ computing data centers, to the use of AI for management and effective use of DNA databases, the Chinese Communist Party is leveraging machine learning and deep neural networks to perfect a digital authoritarian state.

In many aspects, the primary difference that emerges between the US and Chinese ecosystems is that the American government has to create incentives for things that the Chinese party-state can do coercively. As Scharre notes, “*Were Chinese tech employees to openly criticize the government, as employees did at Google, Amazon, and Microsoft of the U.S. government, they would risk jail time or worse.*” What’s more, multiple other countries with authoritarian regimes (such as Zimbabwe, Tanzania, Libya, Egypt, Saudi Arabia and Thailand) are willing to import the Chinese model.

Each chapter in Scharre’s book is a story alluding to one aspect of the AI ecosystem, such as a stakeholder, an event, or a phenomenon. For example, while highlighting the potential implications of machine learning for network security and mobile communications, Scharre narrates the story of ‘Modulate’, an AI software that switches the tone and tenor of an individual’s voice on the phone to that of another individual of their choice. Similarly, to illustrate the truth-altering capabilities of Large Language Models (LLMs) and Generative AI, the author delves into his meeting with Jack Clark (former Policy Director of OpenAI), who demonstrated to him the innovative text generation capabilities of the GPT-2 LLM (which has now turned into GPT-4).

Aside from describing in great detail, almost like a primer, the use cases of AI and key actors engaged in its development and deployment, ‘Four Battlegrounds’ lists key differences between the authoritarian uses of AI, as witnessed in China, and the more democratic and knowledge-building-related uses of AI, as witnessed in the US. He explains how talent has a strategic role to play in this, because even though the American AI value chain is disaggregated, the US has a unique capability to attract talent from around the world: they can receive higher education in the US, continue research work on algorithms, and eventually outperform other actors (like China) in the race for technological power.

At the same time, his recommendation is also that governments invest in algorithmic evolution and computing power on their own accord, without relying solely on the developments made by private companies and researchers. This, he says, is necessary because with time, AI computing power and training data requirements are becoming bigger and bigger, and after a certain point in time, governments’ pockets will only be the ones deep enough to invest in them.

The book makes an effort to remind readers of the expectations we have of AI systems, based on the perceptions we have developed from watching sci-fi movies and reading dystopian novels. For example, we expect that AI systems will act exactly like human brains do. However, the reality can go either way. For example, on the one hand, the 2017 poker competition between Jason Les (who has been referred to as the “human benchmark” for poker skills) and Libratus AI software (developed by Carnegie Mellon University) demonstrated that Libratus could learn techniques and poker hands that the human players couldn’t anticipate and imagine, even though they believed they were well prepared

to tackle it. On the other hand, Scharre narrates his own experience with attending an AlphaPilot drone racing league, where he argued that autonomous drones racing a human-piloted aircraft “couldn’t match even a half-decent run from a human.”

At the same time, the book sheds light on warning signs that are already apparent to humans – especially those navigating the ‘Post-Truth’ information landscape. As already mentioned above, AI has the capability to generate authentic text and skew audio-visual communications. This has severe implications for the socio-moral fabric.

Scharre gives the example of how AI-enabled face-swapping and deepfake video technologies led to the creation of an obscene video of actress Scarlett Johansson, who later remarked that “*the internet is a vast wormhole of darkness that eats itself.*” Moreover, such an information landscape, where everything is “synthetic reality,” is the greatest tool in the hands of authoritarian leaders, who thrive on manipulating narratives. Clearly, this indicates that a two-pronged investment is required towards both, advancing the capabilities of AI technologies, as well as countering their dangerous misuses.

From an Indian perspective, the book has various takeaways. To begin with, even though India, with the world’s largest population, generates massive amounts of (potential) training data, Scharre demonstrates the significance of diverse data sets when training an AI model. He states that a trustworthy algorithm needs to be trained in different contexts and with different sets of people, lest it remain restricted by regional similarities.

Also, the author demonstrates China’s increasing experimentation with AI in surveillance and weapons systems, as well as influence operations that the country has conducted using AI and social media tools like TikTok (which India has now banned, following the clash between armed forces on the two sides of the Border at Galwan Valley, Ladakh, in June 2020). These issues are interlinked with India’s national security goals, and the book serves as a guide in this regard. Scharre warns, however, that for cooperation in military AI, it is necessary for the two sides to build and deliver on existing Confidence-Building Measures (CBMs). This is easier said than done, given the unwillingness China has demonstrated in the past few years to negotiate with India’s interests in mind.

In global politics, AI has now become a tool to further national security imperatives and project power. In this regard, CBMs aren’t something for India and China to build alone – a multilateral effort is required to build global governance mechanisms on AI. However, this is not to say bilateral and minilateral arrangements towards this aren’t already emerging. Taking India’s example itself, Scharre has discussed in the book the initiatives taken collaboratively by Quad countries (India, the US, Japan and Australia) on AI ethics and governance.

One such initiative is the launch of the Quad Senior Cyber Group, and there is the promulgation of documents such as the ‘Quad Joint Principles on Cybersecurity’ and ‘Secure Software’. India and the US also concluded an Initiative on Critical and Emerging Technologies earlier this year, the impetus behind which is to secure scientific and industry-wide partnerships between the two countries to build a robust innovation ecosystem.

If both India and the US's rivalry with China is one of the factors behind the conclusion of this Initiative, the global AI ecosystem is becoming fragmented into two large bins – one that is China-led (and appeals to authoritarian regimes) and another that is led by democratic countries like the US and its allies. This is also evident from another example that Scharre discusses in this book – the support Japan and the Netherlands have extended toward US export controls against China, which essentially curb China's access to key semiconductor technology and manufacturing tools (which are crucial to the advancement of AI compute).

Overall, the book isn't essentially jargon-free, but is equipped with an expansive glossary and source list to help the reader navigate the complexity that accompanies the study of a topic such as AI. It is the second of Scharre's books on understanding the AI ecosystem, and its implications on geo-technological power, the first one being '*Army of None*' – which is another must-read.

"Four Battlegrounds: Power in the Age of Artificial Intelligence" by Paul Scharre, W.W Norton, 2023, Pages 496. ₹2,320 (Hardcover).