### **Critical Humanities**

Manuscript 1051

## The Coming Wave by Mustafa Suleyman: A Review

Puspa Damai

Follow this and additional works at: https://mds.marshall.edu/criticalhumanities

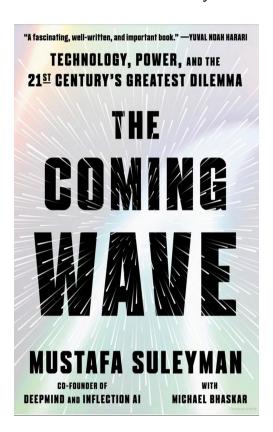


Part of the History Commons

# The Coming Wave by Mustafa Suleyman: A Review

### Puspa Damai

Marshall University



Book Review: *The Coming Wave: Technology, Power, and the 21<sup>st</sup> Century's Greatest Dilemma* by Mustafa Suleyman with Michael Bhaskar (New York: Crown, 2023)

**Puspa Damai** is an Associate Professor of English at Marshall University, Huntington, WV.

Mustafa Suleyman, the co-founder of DeepMind and CEO of Inflection AI, presents a captivating case on the promises and perils of Artificial Intelligence in his thought-provoking and compelling book, The Coming Wave: Technology, Power, and the Twenty-First Century's Greatest Dilemma. Suleyman asserts that we stand at the precipice of a technological explosion, with Artificial Intelligence, biotechnology, and other emerging technologies. While these advancements hold the promise of immense prosperity, they also pose existential threats to humanity. Our cultural and political institutions, including the nation-state, which forms the foundation of the current global political order, face the risk of being swept away by the powerful forces of ultra-modern technologies.

The argument in Suleyman's book is straightforward: the emergence ChatGPT and other large language models (LLMs) is a precursor to a bigger and far more consequential "coming wave" of AI and synthetic biology. This wave is destined to "usher in a new dawn for humanity, creating wealth and surplus unlike anything ever seen. And yet their rapid proliferation also threatens to empower a diverse array of bad actors to unleash disruption, instability, and even catastrophe on an unimaginable scale."1 What makes this impending catastrophe inevitable and inescapable is not only that our future depends on the technologies it is imperiled by but also because there seems to be no predictable or possible way of containing and preventing the catastrophe these technologies are set to unleash upon the world.

The book opens with a self-disarming statement: "This is how an AI sees it" followed by a somber reflection on the future of humanity itself. The solemn tone of the prologue foreshadows the apocalyptic temper of the book according to which "the fate of humanity hangs in the balance" due to the advent of Artificial Intelligence. This cataclysmic worldview is further amplified by the title of the first chapter itself: "Cont nment is not Possible" from which the letters "AI" seem to have escaped and achieved a level of sinister autonomy and self-determination, as they are left hovering ominously at the top of the page.

By opening the narrative from the point of view of an AI and by candidly spelling out the dangers posed by this technology, Suleyman at once exposes two key myths surrounding AI and its makers. Unlike his peers in the tech-corporate world who are brazenly promoting AI as the greatest equalizer and solver of the totality of ills plaguing humanity today, Suleyman underscores the threats AI poses to the very existence of the world as we know it. At the same time, he deconstructs the myth that AI as a tool either merely holds a mirror up to our society as AI learns exclusively from what input we feed it or seeks to dominate and overwhelm us like a totalitarian dictator. Suleyman's deconstruction of these two myths in the book institutes a rare

\_

<sup>&</sup>lt;sup>1</sup> Suleyman *The Coming*, 7

inscription of critique at the heart of AI and techno-capitalism.

Suleyman dedicates a significant portion of this three hundred pages long book to expressing his disillusionment and detachment from the world of technology corporations. Early on in the book, he reminisces about a gathering of influential figures in the tech industry on the West Coast of the United States during the 2010s. It was during this event that he first voiced his concerns regarding artificial intelligence (AI), emphasizing that AI had brought forth a host of threats requiring "proactive responses" without which we might have "massive invasions of privacy or . . . a misinformation apocalypse." <sup>2</sup> Despite these warnings based on currently available factual and logical data, the room remained unmoved, and his message did not land.

For Suleyman, not only his peers, but people in general seem to be affected by what he calls "pessimism aversion," an emotional response that refuses to accept the possibility of negative outcomes. Suleyman argues that the book seeks to confront the overwhelming emotion of pessimism aversion, but it also addresses the "core dilemma" or "terrible bind" that AI has presented to us that "sooner or later, a powerful generation of technology leads humanity toward either catastrophic or dystopian outcomes." <sup>3</sup>

The Coming Wave comprises fourteen chapters, thirteen of which are grouped into four parts titled "Homo Technologicus," "The Next Wave," States of Failure," and "Through the Wave." The first part of the book, "Homo Technologicus" narrates the story of technology in conjunction with human evolution. In this section of the book, Suleyman demonstrates how in the story of our evolution from primates to the planet's dominant force, the waves of technology and humanity have evolved in symbiosis, hence the moniker, humans as technological animals. He argues that throughout history, "population size and innovation levels are linked. New tools and techniques give rise to larger populations,"4 which in turn leads to greater specialization, emergence, and proliferation of general-purpose technologies. This proliferation of technology has resulted in what he calls "the containment problem." He concludes the first part of the book with the proposition that the key truth about Homo Technologicus is that while historically humans faced the challenge of creating technology and unleashing its power, "the challenge of technology today is about containing its unleashed power, ensuring it continues to serve us and our planet."5

Part II, "The Next Wave," is the longest and perhaps the most important section of the book. It consists of five chapters that delve into the key topics and issues explored throughout the book. The opening chapter of this section, "The Technology of

<sup>&</sup>lt;sup>2</sup> Suleyman *The Coming*, 11

<sup>&</sup>lt;sup>3</sup> Suleyman *The Coming*, 11

<sup>&</sup>lt;sup>4</sup> Suleyman The Coming, 27

<sup>&</sup>lt;sup>5</sup> Suleyman *The Coming*, 48

Intelligence," recounts the author's personal involvement with the development and his interaction with AI since his cofounding of DeepMind. Suleyman provides an inside scoop on how AI operates by taking us behind the scenes back in 2012 when he and his team at DeepMind were striving to develop AGI (artificial general intelligence) through the creation of an algorithm called DQN (Deep Q Network) to train it so that it can learn the game called Breakout. Initially, Suleyman recalls that DQN appeared to be utterly incapable of performing the task. However, as more data on the game was supplied and more repeated plays were conducted, DQN quickly recovered and something extraordinary happened. Suleyman aptly describes this event as "remarkable":

For the first time, I'd witnessed a very simple, very elegant system that could learn valuable knowledge, arguably a strategy that wasn't obvious to many humans. It was an electrifying moment, a breakthrough in which an AI agent demonstrated an early indication that it could discover new knowledge.<sup>6</sup>

Suleyman proceeds to draw a parallel between technology and language or chemistry. He notes that like a language or chemistry, AI is "not a set of independent entities and practices, but a comingling set of

parts to combine and recombine."7 This intermingling characteristic of AI, resembling linguistic or chemical processes, is effectively utilized by modern language models (LLMs) like ChatGPT through the implementation of an "attention map." Sulevman further discloses his involvement in the development of a previous LLM project called LaMDA, which predated ChatGPT. Interestingly, LaMDA was recognized as having a sense of sentience by Google's Blake Lemoine, who unfortunately faced termination due to his tendency to anthropomorphize AI, which Google argued violated security regulations.

The rest of the chapters in part II illustrate and expand on his notion of waves of technology by describing how AI is aiding the technology of life or synthetic biology and robotics. It is in this part that he fully fleshes out his notion of wave, which, according to him, has four key features: i) asymmetrical impact ("It won't just be Ukrainian soldiers using weaponized drones. It will be anyone who wants to," [as the term "wave" implies, it has initiated] colossal transfer of power from traditional states and militaries anyone with the capacity and motivation . . . and one point of failure [would] go quickly cascading around the world"); 8 ii) hyper-evolution (innovation in the 'real world' moving at a digital pace); 9 iii) omni-use (like electricity, AI will be an on-demand utility); 10

<sup>&</sup>lt;sup>6</sup> Suleyman *The Coming*, 52

<sup>&</sup>lt;sup>7</sup> Suleyman *The Coming*, 56.

<sup>&</sup>lt;sup>8</sup> Suleyman *The Coming*, 106, 107

<sup>&</sup>lt;sup>9</sup> Suleyman *The Coming*, 108

<sup>&</sup>lt;sup>10</sup> Suleyman *The Coming*, 111

iv) *autonomy* (where constant intervention and oversight are increasingly unnecessary).<sup>11</sup>

After enumerating the threats that AI poses to humanity, especially to our economic and medical welfare, in part III of the book, Suleyman goes on to examine the potential imperilment to the nation-state that AI might cause if not urgently and effectively addressed. He notes that the nation-state as the foundation of our present political order has already been rocked by the lack of trust in the government and the rise of authoritarianism across the globe. He contends that AI will only further "amplify" these crises by opening the floodgates of misinformation and by unleashing new forms of violence, both of which could lead to catastrophic accidents.

Conceding that tech companies often view the state as an obstacle to their goals of promoting and profiting off of new technology, Suleyman insists that we cannot wish the state away as AI's erosion of "the stable, sovereign borders of nation states" could lead to catastrophic consequences. By citing the 2017 NHS crisis in Britain caused by NSA's WannaCry, Suleyman underscores the urgent task of rallying behind efforts to shore up agents of containment in a world where states fail to act against a threat.

The concluding section of the book (part IV) explores the many ways "society can

begin to face the dilemma, to shake off pessimism aversion and really grapple with the containment problem." Suleyman argues that "containment of the coming wave" is "not possible in our current world," however, we can and must give it a chance. He proposes the following nine steps for that chance as a path forward through the wave:

- Technical safety: We need an Apollo-like program for AI-safety and biosafety.
- Audits for new AI tools: Red teaming or hunting for flaws is another step toward containment so that new AI tools are subjected to thorough safety and ethical assessments before being released to the public.
- Attacking choke points to buy time: Suleyman cites US export controls on selling advanced semiconductors to China as an example of not just a geostrategic play but an experiment in containment.
- Holding makers responsible: Suleyman believes that credible critics of AI must be practitioners by "actively showing the way" to the makers of AI.
- Profit + Purpose: According to Sulleyman, we need to balance profit with ethics; he cites his own

<sup>&</sup>lt;sup>11</sup> Suleyman *The Coming*, 113

<sup>&</sup>lt;sup>12</sup> Suleyman *The Coming*, 226.

<sup>&</sup>lt;sup>13</sup> Suleyman *The Coming*, 238.

- involvement in Google's "AI Principles" as an example of this.
- Governmental reform for AI: Besides surviving the current crisis, governments must employ a two-pronged approach: educating (themselves) about AI and regulating AI. They must try to learn about, invest in, and regulate AI through reformed taxation and legislation.
- International alliances: Forging international agreements to oversee
  the development and deployment
  of AI and synthetic biology is in the
  interests of countries such as China
  and the US so that groups like Aum
  Shinrikyo do not crop up everywhere.
- Culture of embracing failure gracefully: Develop a culture of ethics similar to the Hippocratic oath in medicine: *Primum non nocere* – first, do no harm.
- People Power: Change happens when people demand it – AI creators and entrepreneurs should promote civil society or popular movements for change.

Written in a reader-friendly style – perhaps Michael Bhaskar's contribution to the project – this book would remind readers of the classic *A Brief History of Time* by Stephen Hawking. Despite its brilliance, however, Suleyman's powerful retelling of the compelling story of AI has been tainted by a few (minor) blemishes. To identify one of the issues, the trope of the wave helps the

author allude to the impending deluge and apocalypse, but it also prevents him from clearly seeing the fact that the existing economic and political terrain of the world is uneven where waves crash against unsurmountable boulders of inequities. His prophetic warning against the threat posed by AI also gets blunted to some extent by this trope of the wave as he seems to suggest that technologies evolve through waves, which seamlessly connect the past and the present and crest in the future in the form of AI and synthetic biology. Even long before AI, and definitely before the coming wave Suleyman has predicted with such prescience, the world of some people was always already and continues to be threatened by technological and ideological forces such as imperialism, colonialism, and globalization. Not citing any sources for the information included in the book also gives the impression of a grand narrative originating from a single source as opposed to a robust discussion and debate consisting of multiple and often conflicting and contradictory voices and perspectives, which could have more accurately captured the state of our response to AI right now.

Overall, *The Coming Wave* is a powerful warning to the world against our pessimism aversion and our refusal to accept the existential threats posed by AI. Instead of playing the role of a doomsayer, Suleyman lists a range of potential options (listed above in this review) that the world still has to contain these threats. This book is a timely call to action for all of us to work together to ensure that the coming waves

of technology are channeled into benefiting humanity, and not destroying it.

#### References

Suleyman, Mustafa with Michael Bhaskar.

The Coming Wave: Technology,
Power and the 21st Century's Dilemma. New York: Crown, 2023.