Is AI going to become superhuman and destroy human civilization? How do we achieve a balance between innovation and safety?

From planning permaculture gardens, specialty diets and exercise routines, to helping with Excel formulas, poetry and term papers, to transcribing medical visits into clinical notes, to deciphering intricate aspects of healthcare, finance, agriculture, and education, to writing code without knowing coding, to designing parts of a combat aircraft or spaceship, Artificial Intelligence (AI) technologies have permeated, snaked in if you prefer, every aspect of our lives. AI language powered, generative tools like LaMDA, Elicit, Copilot, Chinchilla, PaLM, and AI-powered, conversational interfaces like Bard, Bing and ChatGPT are revolutionizing the field of natural language processing, search, and conversational AI. Though these AI-powered models, are presently incapable of true consciousness or independent thought, and the responses are based on statistical patterns learned from training data, they perform tasks that, hitherto, had been the exclusively preserve of human intelligence.

Trained to generate human-like responses to text input, these AI-powered tools are literally mesmerizing and hypnotizing us with the sheer potentiality and limitlessness of what they can achieve. And, at the bedrock of AI-powered technologies like Bard, Bing, Alexa, Google Assistant, IBM Watson, Apple Siri, Salesforce Einstein, Bert, DuckDuckGo and ChatGPT is the ability to collect, process and analyze colossal amounts of data, known as 'Big Data', and use natural language processing (NLP) and neural networks (types of mathematical systems that learn skills by analyzing data) to recognize statistical patterns and relationships in data and generate human-like responses and predictions, to user input. The big data is analyzed and made sense of utilizing specialized tools and techniques such as data mining, machine learning, and AI.

Equipped with a rapidly growing ability to handle copiously gigantic amounts of data, these neural networks, including Large Language Models (LLMs), are constantly learning, evolving and pushing the boundaries of possibilities in natural language processing and other fields. And as the amount of data increases, their processing power grows exponentially. Indubitably, big data is driving innovation in the field of AI.

But even as AI become ubiquitous in our lives, it is important to ensure that AI is developed and used in an ethical and responsible manner, with proper oversight and regulations firmly in place.

Opined The Economist in an article titled 'Large, creative AI models will transform lives and labor markets': "(For example) ChatGPT embodies more knowledge than any human has ever known. Running alongside this excitement is deep concern, inside the tech industry and beyond, that generative AI models are being developed too quickly." (*Large, Creative AI Models Will Transform Lives and Labour Markets*, n.d.). What are the ethical implications of these systems? For starters, the effectiveness of these AI driven models hinges on the 4 Vs - Volume (large amounts of data), Velocity (fast-paced data streams), Variety (different types of data from various sources), and Veracity (data quality and reliability) - that they are trained on.

Today, the totality of all human knowledge is available online, but along with the promise comes peril. And questions need to be asked on the ethical implications of how this data is being collected, used, and shared. Questions that hinge on privacy, bias, transparency, and accountability. For example, the use of biased algorithms can lead to negative consequences for individuals and society as a whole. AI has also been used in producing deepfakes – convincing fake videos made using AI.

AI, which Demis Hassabis, the CEO and co-founder of DeepMind - a subsidiary of Google's parent company, Alphabet - and one of the world's leading AI labs, described, during a broadcast on CBS' 60 Minutes, as "the most important invention that humanity will ever make" (*Is Artificial Intelligence Advancing Too Quickly? What AI Leaders at Google Say*, n.d.), in an earlier interview with Time, promises that although DeepMind works on making machines smart, they (at DeepMind) "wanted to keep humanity at the center" of what they are doing there. (*DeepMind CEO Demis Hassabis Urges Caution on AI*, 2023)

Tech pioneers like Hassabis are caught on a cleft stick. While they argue that AI is now on the cusp of being able to make tools that could be deeply damaging to human civilization and urge competitors to proceed with more caution than before, competitive pressure is propelling humanity into the future ready or not. (*DeepMind CEO Demis Hassabis Urges Caution on AI*, 2023) Hassabis and his 1000 colleagues at DeepMind are already as quoted in Time working "toward a much grander ambition: creating artificial general intelligence, or AGI, by building machines that can think, learn, and be set to solve humanity's toughest problems". (*DeepMind CEO Demis Hassabis Urges Caution on AI*, 2023) "AGI, Hassabis believes, will be like the harnessing of electricity—that will change the very fabric of human life." (*DeepMind CEO Demis Hassabis Urges Caution on AI*, 2023)

In an article published over five years in Forbes, titled 'Is It Too Late For Big Data Ethics?' the author Kalev Leetaru commented: "The tide in the world of big data and AI research seems to have turned decidedly against the notion of outside ethical review. Instead of asking what questions we should be asking that would better human society, data scientists today all-too-often ask what questions are possible with the data and tools at hand and especially what questions would generate the most attention (and hence publication prestige and grant funding)." (*Leetaru, n.d.)* This view has dramatically changed, with the calamitous risk

associated with AI models moving too fast becoming more imaginable and realistic.

So, when over a 1,000 technology leaders, researchers and their ilk including a wide variety of people from industry and academia and their ilk dashed of a letter - 'Pause Giant AI Experiments: An Open Letter'(*Future of Life Institute, 2023*) calling on all AI labs to immediately pause for at least six months the development of the most advanced systems and the training of large AI language models more powerful than ChatGPT, the chatbot introduced in March by the scrappy, startup research lab OpenAI, the world was not surprised.

The letter released by the nonprofit 'Future of Life Institute' on March 22 succinctly and rather bluntly warned that AI tools present a "profound risks to society and humanity." And that AI developers are "locked in an out-of-control race to develop and deploy ever more powerful digital minds that no one — not even their creators — can understand, predict or reliably control." Signed among others by billionaire tech mogul Elon Musk, who incidentally cofounded OpenAI, and recently in a discussion on the dangers of AI with Fox News' Carlson Tucker opined: "AI is more dangerous than, say, mismanaged aircraft design or production maintenance or bad car production. In the sense that it has the potential, however small one may regard that probability, but it is non-trivial, it has the potential of civilization destruction." In the same interview Musk added that AI is everywhere, affects the future fundamentally and that humanity would "have to be careful (with AI)." (*Elon Musk to Tucker Carlson: AI Has the Potential to Destroy Civilization | Fox News Video*, 2023)

Sundar Pichai, Chief Executive at Google reiterated Musk's thinking, telling journalists Kevin Roose and Casey Newton, in Hard Fork, a New York Times' podcast, titled "Google C.E.O. Sundar Pichai on Bard, A.I. 'Whiplash' and Competing With ChatGPT" that "AI is the most profound technology humanity

will ever work on". And that AI "will get to the essence of what humanity is". Added Pichai: We (at Google) are definitely working with technology, which is going to be incredibly beneficial, but clearly has the potential to cause harm in a deep way. And so I think it's very important that we are all responsible in how we approach it." (*Roose et al., 2023*)

The time is now for both a pause and government regulation. Of course, there are those who argue that a pause or a moratorium will not be implementable, and that state actors from other nations will continue to ramp up AI technologies.

Writing in the Financial Times, nearly a decade ago Tim Harford was prescient when he said: "Statisticians have spent the past 200 years figuring out what traps lie in wait when we try to understand the world through data. The data are bigger, faster and cheaper these days – but we must not pretend that the traps have all been made safe. They have not." (*Harford, 2014*)

References

1. April 22, 2023, 'Large, creative AI models will transform lives and labor markets', The Economist, https://www.economist.com/interactive/science-and-technology/2023/04/22/large-creative-ai-models-will-transform-how-we-live-and-work?utm_campaign=a.the-economist-this-week&utm_medium=email.internalnewsletter.np&utm_source=salesforce-marketing-cloud&utm_term=4/20/2023&utm_id=1569410
2. Pelley, S, April 16, 2023, 'Is artificial intelligence advancing too quickly? What AI leaders at Google say, April 16, 2023, 7.02 pm, https://www.cbsnews.com/news/google-artificial-intelligence-future-60-minutes-transcript-2023-04-16/
3. Perrigo, B, January 12, 2023, 'DeepMind's CEO Helped Take AI Mainstream. Now He's Urging Caution' https://time.com/6246119/demis-hassabis-deepmind-interview/
4. Kalev,L, October 16, 2017, 'Is It Too Late For Big Data Ethics?'Forbes, https://www.forbes.com/sites/kalevleetaru/2017/10/16/is-it-too-late-for-big-data-ethics/?sh=2e5067733a6d
5. 'Future of Life Institute' March 22, 2023, https://futureoflife.org/open-letter/pause-giant-ai-experiments/
6. Elon Musk to Tucker Carlson: AI has the potential to destroy civilization | Fox News Video. (2023, April 14). Fox News. https://www.foxnews.com/video/6325033515112
7. Roose, K and Newton C (Hosts), March 31, 2023, Bard, A.I. 'Whiplash' and Competing With ChatGPT, Hard Fork, Technology, https://www.nytimes.com/2023/03/31/podcasts/hard-fork-sundar.html
8. Harford, T, Financial Times. March 28, 2014 Big data: Are we making a big mistake? https://www.ft.com/content/21a6e7d8-b479-11e3-a09a-00144feabdc0