Ricky's Afterthought:

Does AI pose an existential threat?

A.C. (Ricky) Metaxas

Life Fellow St John's College Cambridge UK
Contact E-mail: acm33@cam.ac.uk

It is a surprise that almost 40 years after publication of the "yellow bible" I still get royalties, not enough to buy a Lamborghini but enough to raise a smile.

I thought that would continue for ever but I had a shock the other day when I asked ChatGPT the simple question, "what is microwave heating?".

Within seconds the answer appeared on my mobile, "Microwave heating is a process that uses high frequency electromagnetic waves, called microwaves to produce heat within a material, etc." and went on for a couple of paragraphs, eventually asking me whether I wished for the programme to continue generating more sentences.

So is the time approaching when technical or indeed any books become obsolete?

AI Start-ups

ChatGPT is an artificial intelligence (AI) chatbot developed by OpenAI in San Francisco and released at the end of 2022. The name "ChatGPT" combines "Chat", referring to its chatbot functionality, and "GPT", which stands for Generative Pre-trained Transformer, a type of Large Language Model (LLM). The latest models expect input formatted in a specific chat-like transcript format, and return a completion that represents a model-written message in the chat. While this format was designed specifically for multi-turn conversations, it can also work for non-chat situations. Rivals include Googles' Bard, Anthropic's Claude and Inflection's Pi (personal intelligence).

The amount invested in AI Start-ups is staggering with some $20 billion muted to have been raised during 2023 alone. There are countless AI start-ups but which one will eventually emerge as "the one" is hard to tell. Hence the reason why shrewd entrepreneurs with lots cash at their disposal invest in a range of AI's hoping that "the one to emerge on top" will be amongst them. All the big entrepreneurs such as Bill Gates, Google's ex-chief Erik Schmidt and the venture capitalist firm Andreessen Horowitz all have stakes in such schemes.

Open Source AI Start-ups

In this field, dominated by Microsoft with its $10 billion alliance with OpenAI and Google who own Alphabet, we now have a newcomer in the form of StabilityAI the billion dollar start-up based in London with some additional tools such as Stable Diffusion and Deep Floyd IF (an open-source model that supports text generation of images) and StableLM, which is able to carry out life like conversations and give answers such as ChatGPT can.

The man behind Stability is Emad Mostaque; born in Jordan the family then moved to Bangladesh and finally emigrated to London and settled in Walthamstow. A degree in Computer Science and Maths at Oxford was followed by working in venture capital. His approach to developing AI platforms is fundamentally different from his competitors in that his programmes are "open source" meaning he is allowing researchers and companies working in the field to access Stability's codes and adapt them as they wish. The other entrepreneur who follows the "open source" route is Mark Zuckerberg launching

Meta's Llama similar the Chart GPT Large Language Model.

AI Server Chips

It is understood that one needs specialised processors to perform various tasks in AI systems. Emad Mostaque realised that the most important tool for developing AI systems in the future rests in the powerful server chips that are inherent in these systems. So he invested heavily in companies that have access to these chips.

Microsoft Azure is one such example, a cloud computing platform which assists researchers with software, platform or infrastructure as a service through their global data centres.

Nvidia, founded by engineers Jensen Huang, Curtis Priem and Chris Melachowsky in 1993, pioneered the GPU (Graphics Processing Unit) which enhances computer graphics. At the heart of this electronic circuitry are specialised chips. Huang and his colleagues too realised that to win the race of developing AI systems one needed these special chips and put all their resources in perfecting them and now are way ahead of any rival chip manufacturer. Two months ago Nvidia announced that it sold $10.3billion worth of data centre GPUs during the second quarter. The activity in this sector is staggering where AI server chips are expected to reach $53 billion at the end of 2023 and expected to double over the next four years. Despite such large numbers there is a shortage of Nvidia server chips which hit many smaller start-ups who rely on getting the chips when they need them rather than commit their companies to more expensive long-term agreements which will secure ready availability. In June of this year the valuation of Nvidia reached $1 trillion matching Meta's and Amazon's valuations.

Training of AI Systems

The AI systems are learning from what is on the internet and something which is harmful must be eliminated. OpenAI set up a collaborative project in a country in Africa where employees spent countless hours removing harmful data from the internet so that AI systems should not learn from, the likes of murder, abuse, bestiality and so on. However, the work of content moderation should be done humanly and willingly and that project was abandoned after just two years as it was not run properly.

Recent AI systems are under development not only to read a language but to read minds and emotions by observing tiny facial muscle movements and aiming at automatic decision making without a human input.

AI in Medicine

There are undoubtedly benefits in some areas such medicine where robots controlled by AI have already assisted in surgery and other specific tasks. The UK government has announced that it is laying aside £21 million for the National Health Service Trusts which could be used for designing AI tools to assist in medical imaging and diagnosis of various cancers, including lung cancer. A specific example is University College London Hospital's Cadu system which analyses images inside a patient's oesophagus to detect the onset of cancer.

Systems based on AI are already in use, which help to control patient waiting lists. Dr Katharine Halliday, President of the Royal College of Radiologists welcomed the debate of the impact of AI in medicine and added, "At a time when diagnostic services are under strain, it is critical that we embrace innovation that could boost capacity and so we welcome the government's announcement of the £21m fund to purchase and deploy AI diagnostic tools".

Other areas where AI is currently being adopted are finance, city management, such as transportation, manufacturing, politics, social media and many more.

European Initiative

The European Commission has set some strict rules for companies using AI initiatives. Margrethe Vestager, the European Commission executive Vice President who oversees digital policy for the 27-nation bloc, said the following: "On artificial intelligence, trust is a must, not a nice-to-have. With these landmark rules, the E.U. is spearheading the development of new global norms to make sure AI can be trusted".

The companies that she had in mind are not just huge conglomerates such as Amazon, Facebook, Google and Microsoft but other companies that have

poured billions of dollars in developing schemes based on AI. The European Union released a document in an attempt to regulate an emerging technology before it becomes the norm. The rules have far-reaching implications for major technology companies which are developing artificial intelligence in various walks of life such as medicine, taxation and transport. The new laws that will eventually come on stream would ask companies to provide evidence that their systems are safe and used responsibly and what the user sees is computer generated rather than coming from actual humans. Further risk assessment on using the technology would be paramount in order to give the user confidence that the system can be trusted.

Cautionary remarks

Pioneers in the field of AI are warning of proceeding with developing codes and platforms without considering the dangers ahead. Dr Geoffrey Hinton, who developed a neural net in 2012, quit his job with Google to be able to warn colleagues about surging ahead with developing platforms run by AI and asked them to pause and consider the consequences. Some think it is already too late but governments are beginning to take this issue very seriously and are attempting to introduce a regulatory frame. In the UK the Competition and Markets Authority gave warning that they were looking into this area with the view to publishing its findings within a few months. The UK government has launched the AI & Digital Regulation Service which in effect is a cross-regulatory advisory service supporting developers and adopters of AI and digital technologies. It provides the guidance, evaluation and data governance pathways. It will bring benefits to the health and social landscape through safer and more effective use of this technology.

Dr Hinton added, "I've come to the conclusion that the kind of intelligence we're developing is very different from the intelligence we have and can be exploited by bad actors." Jeff Dean, Google's chief scientist responded, "As one of the first companies to publish AI Principles, we remain committed to a responsible approach to AI. We're continually learning to understand emerging risks while also innovating boldly." Elon Musk was a cosignatory in a letter published in April 2023 with thousands of others urging caution and pause in some significant areas of AI development.

Sam Altman who developed ChatGPT fears that eventually this will present major problems for humanity. One of his recent comments at an economics event was, I quote "What I lose the most sleep over is the hypothetical idea that we already have done something really bad by launching ChatGPT. That, maybe, there was something hard and complicated in there [the system] that we didn't understand and have now already kicked it off." Another quote goes like this, "A misaligned super-intelligent AGI (Artificial General Intelligence) could cause grievous harm to the world; an autocratic regime with a decisive superintelligence lead could do that too" and "think people should be happy that we are a little bit scared of this." Some have claimed that, "Models will confidently state things as if they were facts that are entirely made up."

Ian Bremmer, President and founder of Eurasia &Gzero media and Mustafa Soleyman, the CEO and co-founder of Inflection said in an interview recently, "AI governance must also be as impermeable as possible. Unlike climate change mitigation, where success will be determined by the sum of all individual efforts, AI safety is determined by the lowest common denominator: a single breakout algorithm could cause untold damage."

Sir Patrick Vallance, the outgoing Chief Scientific Advisor in the UK warned that AI may cause uncertainty in what to believe is true when there is doubt and the impact AI may have on future jobs. Governments must ensure that retraining would be necessary.

Anna Korhonen, Professor of Language Processing and Director of Human-Inspired Artificial Intelligence at Cambridge University stated in an article recently, I quote, "…from a technical standpoint there is no basis for the current hype for AI outsmarting humans or taking over. We are still far from human-level AI." However, she goes on to state a word of caution, "…of course the technologies can also be used for malicious purposes be that financial fraud, toxic language, fake media and misinformation. So the suggestion that AI will take over from the human race is not the biggest current concern; that it will perpetuate global inequalities and increase the threat to privacy and

security are just a couple of examples that are of more pressing concern."

The centre for AI safety, a San Francisco based non-profit organisation stated recently, "Mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war." This statement was issued as encouragement to AI experts, journalists, policymakers and the public to talk more about urgent risks relating to artificial intelligence. Above all else AI systems must be safe, secure and transparent.

So does AI pose an existential threat? Many argue that it does. Only time will tell.

New AMPERE member responsible for Social Media

During the AMPERE 2023 OGA, Feyisara Adekunle offered to look after the AMPERE presence on social media. A short bio of Feyisara is reported below.

Feyisara Adekunle is a seasoned Community and Engagement Lead with over 5 years of experience in mental health and community outreach. She holds a bachelor's degree from the University of Grant MacEwan and pursued a career in development therapy before joining the team at Aurora Hydrogen. Feyisara has successfully led multiple high-impact community outreach programs and events throughout her career. She is an active member of several non-profit outreach organizations, including A Village of Wellness, Munangwa Foundations, AbbyCare Health and Wellness Centre, and The African Arts Foundation of Edmonton.

Feyisara is deeply passionate about community engagement and adopts a holistic approach to care, which is evident in her co-creation of 'A Divine Experience (ADE),' a lifespan development and entertainment business, alongside her brother. Beyond her professional life, Feyisara channels her creativity primarily through writing. Her mission is to continue spreading awareness and knowledge through different creative outlets.