ARTIFICIAL INTELLIGENCE: What's the hype and how is it impacting society?

Dr Devyani Gajjar
Physical Sciences and Digital Adviser
POST Parliamentary Office of Science and Technology

Everyone has been talking about AI. So, what's causing the hype? November 2022 saw the first public release of the popular chatbot ChatGPT and AI was thrust into a global limelight.

For many people, ChatGPT, and other chatbots released since then such as Bard and Claude, seemed to be producing outputs that were increasingly indistinguishable from human outputs. For example, AI seemed to be able to perform tasks, such as writing essays, at a similar level to humans, and this had implications for different areas of society, such as AI impacting peoples' jobs or AI being used by students in exams.

The Parliamentary Office of Science and Technology (POST), Parliament's in-house research team, has recently produced several impartial research briefings on various aspects of AI to inform scrutiny. Our research explains new advances in AI, drivers of these advances, and synthesises evidence around how it could impact society.

NEW ADVANCES IN AI

In the past few years, AI systems have gained two main advances compared to previous forms of AI [1].

Firstly, the ability of some AI technologies, known as generative AI, to generate realistic content such as text, images, media and audio, has improved drastically.

Secondly, recent advances mean some AI systems can be increasingly adapted to a wide range of tasks. This is in contrast to many AI technologies previously being designed to perform a specific task only. AI systems that can perform a wide range of tasks are known as Foundation Models. Large Language Models, such as ChatGPT, are a specific type of Foundation Model that can carry out a wide range of language related tasks. For example, ChatGPT can write essays, answer questions, spellcheck, and synthesise different sources.

DRIVERS OF ADVANCES IN AI

There have been four main drivers of recent advances in AI.

1) Firstly, increased data availability has meant AI models can be trained on larger and larger datasets. Large language models are often trained on billions or even trillions of bits of data. For example, the large language model underpinning ChatGPT 3.5 (the version released to the public in November 2022) was trained using 300 billion words obtained from internet text. [2]

2) Secondly, the amount of computing power used to develop and run large AI models has increased exponentially in the past half-decade. For example, a report by the Centre for Security and Emerging

Technology noted that a Foundation Model released in 2020 used 600,000 times more computing power than a noteworthy model in 2012[3].

3) Thirdly, there have been investments in computing infrastructure.

4) Fourthly, new AI algorithms developed in the past few years have greatly improved generative AI.

Only a few large private sector technology companies have developed the biggest Large Language Models due to the scale of computing power and data required. A 2023 report by the Government Office for Science predicted that, in the near future, the development of cutting-edge large language models is highly likely to be carried out by a select few companies with the required funding for computing power, skills and data[4]. These include OpenAI, Google, Anthropic and Meta. Due to high costs, concerns exist around the inaccessibility of developing Frontier models for small companies, open-source communities and academia, and the concentration of market power by a few private sector organisations.[1]

CONCERNS AND ISSUES

Alongside issues around unequal access to AI systems, AI could create other ethical challenges including discrimination and inequalities from biases in AI systems, decisions being based on incorrect information generated by AI, the spread of false information, and increased mistrust in online information. The use of AI technologies could also impact the economy with evidence suggesting potential effects could be mixed.

Discrimination and inequalities

It has been well-established that AI systems can contain biases which can manifest themselves in different ways, such as an AI developer training the system on data that leads to inappropriate AI decisions. Biased AI systems have already had real-world discriminatory impacts. For example, a 2019 study found that an algorithm used to allocate healthcare in US hospitals was less likely to refer Black people who were equally as sick as White people to healthcare programmes.[5]

Decisions being based on incorrect information generated by AI

Large Language Models, such as ChatGPT, generate text by predicting the most likely words and phrases that go together based on patterns they have seen in training data. However, they are unable to identify if the phrases they generate make sense or are accurate.[1] This can sometimes lead to inaccurate results, also known as 'hallucination' effects, where Large Language Models generate plausible sounding but inaccurate text. Hallucinations can cause problems where the results of an AI are used to take decisions without proper consideration of the risk that the results are inaccurate.

The spread of false information

AI systems can also generate realistic images and videos which can enable the creation of 'deepfakes': pictures and video that are deliberately altered to spread false information either for causing harm or political, personal or financial gain.[5] Whilst deepfakes have been around for years, they had previously required skills and time to produce. Advances in generative AI means that anyone with basic IT skills is more easily able to produce fake content. Particularly with elections around the corner, there are questions about what the impact of AI generated false information could be on elections, on public trust in online content and institutions, and on divisions in society. An example of a recent deepfake that went viral was an audio clip of the London Mayor making inflammatory remarks about Armistice Day.[6]

Impacts on the economy

Evidence is mixed on the implications of AI for the economy. An OECD report says AI is changing the nature of work and could improve productivity and some emerging academic research says AI could be linked to a loss of quality and earnings in some white-collar jobs.[5] There are concerns around AI disproportionately affecting some groups, such as AI more likely affecting clerical work which is mostly carried out by women.[5]

POTENTIALS

Alongside the risks, AI also has immense potential to benefit society. AI tools could be used to strengthen democracy by engaging the public and helping voters to understand manifestos and which candidates and political parties might best align with their priorities.

In healthcare, AI is being used to help diagnose diseases, find new drugs, and develop personalised treatments, which could lead to better health outcomes. In education, AI tools could provide different ways of learning and help educators with lesson planning, marking and other tasks.[7]

AI could be applied across various sectors, increase worker productivity and assist with decision making. A 2023 McKinsey report estimated that generative AI has the potential to add between $2.6 trillion to $4.4 trillion annually to the global economy.[8]

While there are risks and opportunities posed by AI, their societal impact will depend on how AI is used, regulations, geopolitics, access, ownership, safety measures and public attitudes.

For more information on AI, you can read our recent POST reports on the policy implications of AI, an AI explainer and the use of AI in education delivery and assessment. Subscribe to get the latest from POST delivered to your inbox, including new research.

Our AI reports were produced through POST's fellowship schemes. Over 3 months, our POST fellows learn how to write with balance and impartiality, get to experience new and exciting areas of research, and build relationships with key Parliamentary stakeholders. Find out more.

References

1 Parliamentary Office of Science and Technology (2023). POSTbrief 57: Artificial Intelligence – An explainer. UK Parliament.

2 Hughes, A. (2023). ChatGPT: Everything you need to know about OpenAI's GPT-4 tool. BBC Science Focus Magazine.

3 Lohn, A. J. et al. (2022). AI and Compute: How Much Longer Can Computing Power Drive Artificial Intelligence Progress? Centre for Science and Emerging Technology.

4 Government Office for Science (2023). Future Risks of Frontier AI.

5 Parliamentary Office of Science and Technology (2024). POSTnote 708: Policy implications of artificial intelligence. UK Parliament.

6 Sky News (2023). Deepfake audio of Sadiq Khan suggesting Remembrance weekend 'should be held next week instead' under police investigation. Sky News.

7 Parliamentary Office of Science and Technology (2024). POSTnote 712: Use of artificial intelligence in education delivery and assessment. UK Parliament.

8 Chui, M. et al. (2023). Economic potential of generative AI. McKinsey. ■