

British Academy & DSIT Roundtable: Possibilities of AI for the public good

October 2023

Possibilities of AI for the public good

Summary Note

On 1st & 2nd November 2023, the UK will host the first global summit on Artificial Intelligence (AI) safety at Bletchley Park. As the scope of the Summit is focused on the safety risks of frontier AI, the UK government has been undertaking wider engagement to ensure that other important issues are also discussed. This included a joint roundtable on AI and the public good between the Department for Science, Innovation and Technology (DSIT) and the British Academy. The private roundtable convened DSIT officials and expert perspectives from the SHAPE (Social Sciences, Humanities and Arts for People and the Economy) disciplines to feed into planning in the run-up to the Summit.

This note summarises some of the central themes and messages that emerged from the discussion:

Short and medium-term risks and benefits of AI must be considered alongside the long-term opportunities and existential risks

- While the Summit's focus primarily lies with frontier AI safety, it is important for policymakers to pay attention to the near-term and medium-term risks and opportunities of AI in addition to the long-term risks and benefits.
- Primarily spotlighting speculative, long-term risks of AI may undermine public confidence that the present risks of AI are being attended to and prioritised, and therefore generate public pessimism and fear about these technologies.
- SHAPE researchers are encouraging a more holistic view of AI that includes not only frontier AI but also recognises important research and investments that have already been made in the area of AI safety and responsibility, such as existing long-term research projects that have examined relevant challenges and produced insights in this space.
- This is an area in which lessons from history will be highly valuable – for example, we can draw lessons from past instances of system change brought about by innovation, such as the social, physical, and regulatory infrastructure that came with the development, adoption, and diffusion of automobiles.

Geopolitical opportunities and challenges around AI

- The Summit represents a real opportunity for the UK to model what safe and responsible AI looks like – following the Summit, the UK should seek to become world-leading in responsible and safe AI. While the UK is unlikely to be the very top global leader in AI innovation, it could be world-leading on AI responsibility and safety.
- The UK cannot control how AI is deployed globally but should look to use its international position to influence how AI is developed and deployed responsibly.

The challenges of developing an adequate regulatory framework that is equipped to keep pace with the rapid developments in AI

- Many different groups of people are already using AI technologies and experiencing first-hand the effects of advancements in AI. For example, large language models (LLMs) have the potential to severely exacerbate existing risks within digital environments, such as the threat of misinformation, disinformation (particularly in the context of elections or portrayals of armed conflicts), harmful or abusive online content, financial crime, and cybercrime. Furthermore, there have already been cases of existing societal inequalities being reinforced through biases being built into such AI models.
- While the UK has a very-high level reporting and transparency regime, and UK regulators and supervisors have much more access to information than other contexts around the world, the existence of different regulatory regimes across the globe mean that it can be difficult to see where and how specific risks might emerge. Policymakers need to consider how to structure the infrastructure that regulators and supervisors will require to access and understand AI usage data.
- It has taken several years for the UK to develop the Online Harms Bill. Given the rapid pace of change in AI advancements, government and regulators need mechanisms that allow them to respond in timely fashion, and often at pace, to emerging threats and opportunities. Doing so requires having the necessary expertise to hand in relevant government and regulatory bodies, and appropriate powers to intervene.
- AI technologies function in different ways, have different applications, and have varying impacts across sectoral and societal contexts, which means that a 'one-size-fits-all' regulatory approach will not be appropriate. There will need to be tailored forms of regulation to address different kinds of AI development and usage in different contexts.
- Given the huge computing power and resources required to build and train LLMs over many months, only certain companies are likely to be able to undertake such innovation. It may be untenable for public bodies or research organisations to create such systems. Regulation must therefore consider the potential risks that could be created by the problem of the development of large monopolies or oligopolies in AI design and development within the private sector, much of which is based outside the UK.
- Similarly, the public sector will need to ensure it has the necessary expertise or resources to take advantage of and deploy AI, particularly when AI technologies and services are being procured externally.

The importance of using public engagement to inform and shape policy around AI

- Policymakers will need to be aware of the ways in which the public routinely think about and experience AI in their day-to-day lives, including how it relates to people's most pressing needs and concerns (such as accessing healthcare, or impacts upon their work opportunities). Strategies to use AI for the 'public good' will be far more effective if they are delivering tangible benefits that the public want (e.g., upgrading social housing, improving food systems, generating equitable solutions to climate change). Consistent public engagement is needed to identify the range of these potential benefits, and to inform and shape policies that aim to deliver them.
- It is important to clarify and ensure the possible near-term benefits of deploying AI technologies in the delivery of public services and reassure the public around both near-term and long-term risks. There are crucial lessons to be learned from mistakes of the past, such as the backlash created by ineffective approaches to public engagement around GMO technologies. Similar mistakes could prevent policy from using AI technologies in ways that realise the greatest range of benefits for the public, and instead leave different publics feeling hesitant to, and disenfranchised by, policies around AI.

Implications of AI for different sectors, the wider economy, and the future of work

- A number of sectors have already been significantly impacted by the adoption of AI technologies (e.g., the creative industries sector), and many others have the potential to be dramatically impacted (e.g., the transportation sector). Technologies such as LLMs have the potential to transform industries. For instance, the UK's creative industries are world leading. It is therefore crucial for government to address both the challenges and opportunities related to the impact of AI in the creative industries, including the issues it poses for intellectual property, licensing, and copyright. We must have a clear policy on how to protect and nurture these industries with the coming developments.
- For us to understand and address the possibilities, opportunities, and risks of AI, it is crucial that we have a diversity of input from across disciplines and sectors. Expertise from both SHAPE and STEM (Science, technology, engineering, and mathematics) fields is needed if we are to grasp how people are already experiencing AI, and what the social and economic impacts of these technologies might be for different sectors (e.g., health, education, transport, etc.). For example, rigorous quantitative and qualitative social science research will be required to remain abreast of unexpected societal effects of AI as they emerge, and researchers will need to be equipped with the resources, time, and funding to collect such data.

About the British Academy

The British Academy is the UK's national academy for the humanities and social sciences. We mobilise these disciplines to understand the world and shape a brighter future. From artificial intelligence to climate change, from building prosperity to improving well-being – today's complex challenges can only be resolved by deepening our insight into people, cultures and societies. We invest in researchers and projects across the UK and overseas, engage the public with fresh thinking and debates, and bring together scholars, government, business and civil society to influence policy for the benefit of everyone.

If you would like further information about this work or would like to engage with us, please contact: digitalsociety@thebritishacademy.ac.uk