

Written Evidence Submitted by Creative Commons (CC)

(GAI0015)

Creative Commons (CC) is an international nonprofit organisation dedicated to helping build and sustain a thriving commons of shared knowledge and culture. Together with an extensive member network and multiple partners, we build capacity and infrastructure, develop practical solutions, and advocate for open sharing. CC's long focus on open sharing has evolved into a focus on better sharing: We must pursue a commons of knowledge and culture that is inclusive, just, and which inspires reciprocity — a commons that serves the public interest and ensures transparency on how personal data will be applied. Creative Commons works with various international groups including the Open Source Initiative, Open Knowledge and the Open Data Institute (ODI) to remove barriers to information and promote the sharing of data and knowledge by educators, activists, advocates, librarians, educators, lawyers, technologists, scientists and more who may be able to assist with monitoring the effectiveness of the UK Government's approach to AI regulation.

AI is an evolving concept and set of technologies, and any policy or legal intervention addressing AI should be based on high quality, reliable evidence and conceptual certainty, especially given the fast-paced development of these technologies. A contextual approach to AI regulation based on clear principles should enable the UK's regulatory scheme to be able to adapt to new developments in how AI is designed and implemented for different use cases. Harmful uses are best addressed through a contextual approach, based on clear principles that can adapt to new developments in how AI is designed and implemented for different use cases based on their risks to society.

Such a tailored approach can and should avoid overbroad restrictions on general purpose artificial intelligence (GPAI) and the sharing and use of open source tools. Open source tools can play a crucial role in lowering the barriers to development of beneficial technology, ensuring that the capacity to develop and use AI to support better sharing should be widely distributed, rather than concentrated among a narrow few.

Moreover, different sectors have different needs. For example, it is important to provide assurances for patient privacy while developing the applications for AI in healthcare settings. Any system that regulates AI must be able to adapt to evolution in AI. It is difficult to regulate a sector that is rapidly changing and trying to account for risks that may arise e.g. deep fakes and democracy. Accordingly, AI regulation should be flexible and adaptable to evolving applications of the technology. Flexible rules, as opposed to prescriptive regulation, can best adapt to changing technologies.

As the UK considers forms of governance, it is essential for the UK to set up systems that do not conflict with similar measures from other jurisdictions like the EU and USA. Conflicting regulation could make the UK an unattractive location for the global industry, and in general create undue barriers to helpful innovations. The EU is currently debating legislation to regulate AI, and the UK has the opportunity to be a leader in this space by addressing AI based on agreed-upon global standards. For instance, in defining AI itself, the UK may look to broadly agreed upon definitions like the OECD's.

In the context of AI, Creative Commons is primarily focused on the intersection of AI and sharing of knowledge and culture, including as it pertains to copyright and intellectual property law. In particular, it is important to address how copyright impacts AI-training data (AI inputs) and AI-generated content (AI outputs).

AI researchers often need to use various types of data to test and train AI systems. The data used for these purposes can take many different forms, including works that are protected by copyright law. Because of this, copyright law could slow innovation and development in AI by blocking access to and use of those data sets. Accordingly, it is essential to the success of the UK's AI regulation goals to clearly exempt AI training data from copyright's scope.

Lack of high quality data to train AI can result in harmful consequences. In particular, it can encode and replicate biases in society against minority and underrepresented communities, and lead to discrimination in critical areas affecting people's lives like employment and housing. Even when there is access to high quality data to train AI systems, there is a lack of transparency with the public over how their data will be used. Supporting the creation of high quality, open data sets is one way to help address these challenges.

To that end, to the extent that data sets are protected by copyright, Creative Commons encourages providers of such data to make it available under a CC0 license, renouncing any existing rights and dedicating the data to the public domain. This license ensures that training data can be used without any copyright restrictions in the development of AI technologies. As noted, ideally any legal framework would already ensure that such uses are permitted. But to the extent that national intellectual property laws lack such clarity, Creative Commons licenses can play an enabling role.

At the same time, Creative Commons recognises that both the training and implementation of AI systems may pose some risks to individuals and to groups. It is imperative for the UK to establish regulations that promote innovation and support business development while safeguarding against any potential harms that may arise from the development or use of AI systems. The UK should empower individuals with access to and control over their data to enable them to make informed decisions about how their data is used.

When it comes to AI-generated content – such autonomously created output should be presumed to be in the public domain, insofar as the content was generated autonomously without human, creative choices. Originality and human authorship must remain essential to the granting of copyright or other related exclusive rights over creative works, and AI-generated content by itself does not meet those standards. Copyright exists to incentivize human creativity and, where the hallmarks of creative choices don't exist, protection should not arise. Doing otherwise would impede copyright's core purposes and shrink the commons without any recognizable public benefit in the form of bolstered creativity. As with training data, Creative Commons also strongly encourages users of AI to generate content to support broad re-use of that content through the CC0 license – providing legal clarity and more permissions to the extent such content may be covered by copyright.

(November 2022)