

COPYRIGHT AND GENERATIVE AI

*Thomas Balch**
IP and genAI Consultant, UK

ABSTRACT

This is an opinion piece looking at the UK legal copyright framework, human creativity, generative artificial intelligence, and if you need to get permission from the creators or owners of on-line content such as writing, photos, music and film to scrape their content to build AI systems. There is huge concern in the UK's economically vital and “soft power” driving creative industries, especially when the AI output competes directly with the very content that is used to make the generative AI tool work in the first place. There are essentially two opposing views and this article presents one side.

Keywords: copyright, genAI, photojournalism, fake news, human creativity

The law, in a democracy, is predicated on ethics. Its fair or efficient application may be impacted by the imperfections of bureaucracy, the imbalance of wealth and the fickle chance of human circumstance, but essentially laws exist to protect a nation's citizens, to allow its industries and institutions to function properly, and to uphold the values a society holds dear.

Copyright law democratises innovation and creativity. It removes the reliance on independent wealth, taxpayers' money, or corporate patronage for a nation's citizens to create and innovate. It protects freedom of expression; it supports freedom of speech. Furthermore, a robust copyright framework is vital to the digital, culture, media and sports sectors, all significant UK economic drivers. Diluting, devaluing or disregarding copyright would be at best a backward step to an archaic time of cultural patronage. At worst it presages a world where corporate or state backing is vital not only to creative arts but how reality is disseminated, how truth is controlled.

This article is about generative AI, copyright, and how synthetic imagery is produced. It is not about wider AI developments such as the exciting advances in

*Corresponding author: e-mail: corky@corkybalch.com

Disclaimer: The information and data presented in this article were accurate and up-to-date at the time of drafting in June 2024. Subsequent developments may have occurred that could affect the accuracy and relevance of the content.

medicine where most can agree there are clear humanitarian benefits. It is also noteworthy that AI applications are already used to great effect by many creative businesses, maximizing efficiency as well as supplementing human artistry.

The term generative AI refers to digital tools that allow the user to create content such as images, video and music from text prompts. For example, you type black dog running in a field into your computer, and up pops... well, an image of a black dog running in a field. The software “knows” what image to create because through scraping billions of images online a machine has “looked” at black dogs running and “learned” what it is by the accompanying keywords. The italics stress a fallacious argument used by some AI developers who continue a wider and damaging trend of anthropomorphising machines. The fact is that the machines scraping the Internet didn’t have eyes; they didn’t look or see in the way that humans look and see. A machine isn’t influenced by imagery in the way that humans, for example, see art and then create new art that evolves over time. A machine by its nature has to copy the information it scans online to “learn”.

The UK Intellectual Property Office (IPO) recently proposed four options that included changing UK copyright law to allow AI companies to use online content and data (text and data mining, or TDM) more freely and for any purpose. This option posed an existential threat to hundreds of thousands of livelihoods, and risked destabilising a creative industry worth over one hundred billion pounds annually to the UK economy.

TDM copyright exceptions already exist for research uses. If you want to make money from Artificial Intelligence it seems ethical and fair that you should pay the owners of the content that you rely on to build effective AI. It isn’t free for the owners, the copyright holders, to develop, create, host and manage their content and data. Indeed for many it is a significant cost of business.

Mark Getty said twenty years ago that Intellectual Property is the oil of the 21st Century. An arguable assertion, but there’s an important point here with regard the proposed blanket TDM exception. In an online world all copyrighted content is essentially data, and is largely publicly available. Importantly however, “publicly available” does not mean it is free to use however you want, unless specifically stated by the content owner. Copyright law protects the content (or data) creator from third parties using it without their permission.

Copyright is a cornerstone of our democracy and civilisation, not only from a purely artistic perspective. Take news photography. Without copyright laws protecting the content they create photographers, or their licensing agents, can’t afford to cover events, which means in turn that news publishers of all sizes can’t report on the events so vividly and effectively. An image tells a thousand words. Filming and photographing the Olympics, the European Championships, film festivals, concerts and premieres, and covering global news, is specialised, uniquely skilled, and often expensive. Ensuring copyright is respected and its importance understood is essential to creating the content

that brings stories to life and engages and accurately informs the world. We need to illustrate reality with impact, whether it is shocking or upsetting, if it's affirming, momentous, or simply one of the diverse and myriad human-interest stories that account for much of the content we find online.

Adobe recently made AI generated images of the Gaza conflict available for licensing, which posed some challenging questions. Not least, will generative AI eventually replace news photojournalism? Will consumers accept synthetic imagery that is representative of a story, rather than the moment itself shot by a human with a camera? It could mitigate exposure to risk in conflict zones, but do photographers and cameramen not accept risk as part of their vocation? The bottom line is it would be cheaper, which isn't necessarily a net positive for humanity. What is the ultimate cost beyond short-term profit? "Fake" photos have been around for a hundred years and more, from the Cottingley Fairies to the advent of Photoshop and everywhere in between. A journalist on the ground could generate imagery with text prompts and ensure it is representative of what they see with their own eyes, so it is not "fake news", but where does news reporting go from there? Why the need for reporters to cover events or news at all?

Users of unique and powerful, copyrighted content often plead poverty when it comes to their photo and film licensing budgets, even large production companies and established publishers, broadcasters, and media outlets. AI developers point at the cost and complexity of licensing copyright and data, when in truth it can be cost effective, provide accurate data provenance, and is relatively straightforward to manage. The value and wider importance of copyright must not be lost. It should be acknowledged as a necessary cost of business, and paid for accordingly. Wider exceptions for commercial uses are the thin end of the wedge, and nominal or gratis copyright licensing set a bad precedent in troubling times. We need to protect copyright, value it, and ensure creators and rights holders can charge fair fees for their work, and data. In doing so we protect our freedom of speech; we protect the fundamental human right to show and speak truth. In light of the disinformation wars being waged online, this principle is as urgent as ever.

AI and synthetic imagery will replace some and complement other traditional content production and distribution, everywhere from current affairs to art. Maintaining a robust copyright framework and establishing licensing protocols does not impede these technological advances. Copyright licensing protects UK and global AI developers and consumers with data provenance and transparency.

Softening of copyright obligations to allow AI to evolve quicker has wide and far-reaching implications. It puts the whole field of human art and creativity at risk. In a hundred years new generations may be amazed that people created music, art or literature at all. Is this progress? What is the value, beyond profit, of arts and culture to society and civilisation? Technological progress should not be embraced as an imperative, and seen as an inescapable fate. We need to

carefully assess the risks vs. the benefits of any new application. It would be wise to remember the maxim, just because we ‘can’ doesn’t mean we ‘should’. The AI race between nations is as potentially catastrophic to humanity as any arms race.

I am not “anti-tech”. I want to use generative AI as much as the next person and I use a host of tech products for various (and largely very amateurish) creative pursuits. The creative and tech industries have always been symbiotic, especially in the digital age. However, I want original creators whose work has taught and enabled generative AI to be paid if I use it to generate new content. Intellectual Property underpins AI and the endeavours of original creators should be acknowledged and permission sought. If this is not a legal obligation under UK copyright laws then what copyright is for, and whom it is for, must be urgently reviewed. It may be fanciful to expect a wide public debate on copyright but it affects the majority of UK citizens, even if indirectly.

The future for traditional human creators looks bleak if no permission or recompense is required for the use of the content and data that generative AI is built on. The outlook is even bleaker if AI output can be protected under the Copyright Designs and Patents Act 1988 (ref clause 9.3 regarding who or what can be considered the ‘Author’ i.e. the copyright owner), which is unclear and ambiguous. Monopolies, competition, laws and regulation have always weaved a merry dance within industrial sectors, but this is a new beast entirely. One industry is effectively assimilating another.

The IPO’s option for a blanket TDM exception is currently off the table. Copyright reliant businesses must be vigilant and ensure it remains that way. Laws and their applications are imperfect and one can sit on both sides of a legal debate and still distinguish between right and wrong, and what is fair. We shall see if the UK’s copyright framework is fit for purpose in the face of generative AI; if the law isn’t supporting fairness and competition then the law does indeed bray.