Response to the public consultation of the UKIPO on Copyright and Artificial Intelligence

25 February 2025

Intellectual Property and Innovation Research Group at Brunel University of London

Drafting Committee: Dr. Zoi Krokida, Lecturer in AI, Innovation and Law

Dr. Francesca Mazzi, Lecturer in AI, Innovation and Law

Dr Paula Westenberger, Senior Lecturer in IP Law/BRAID Research Fellow

We welcome the UKIPO's initiative of investigating the implications of Artificial Intelligence for copyright law. Below, we address the issues of the public consultation with suggestions for amendments and proposals in specific areas, relating to questions 5, 16, and 28-39.

Dr Paula Westenberger responded to questions 5, 16, 28, 29; Dr Francesca Mazzi responded to questions 30-37; and Dr Zoi Krokida responded to questions 38-39.

Views are our own as individual researchers, and do not intend to reflect those of any institutions, colleagues, research participants or project partners. The views and opinions on each question are those of the researcher who prepared the corresponding answer. Researchers who have not contributed to a particular answer do not take a view on that specific point.

Question 5. What influence, positive or negative, would the introduction of an exception along these lines have on you or your organisation? Please provide quantitative information where possible.

Dr Paula Westenberger:

I am concerned about the potential impact of the proposed exception in relation TDM research, in view of the current issues with s29A CDPA (TDM for non-commercial research), which are discussed in Question 28 below and my forthcoming paper 'Artificial Intelligence for Cultural Heritage Research: the challenges in UK copyright law and policy' (by Dr Paula Westenberger and Dr Despoina Farmaki)

We should avoid the situation that research uses (that should be covered by the non-commercial research exception, which cannot be overridden by contract) end up being regulated under a commercial TDM exception with opt-out, as this could have the possible result of biasing and affecting the quality of research as a result of incomplete datasets. The boundaries between any new commercial TDM exception and the non-commercial research exception must be carefully and clearly delineated. S. 29A must be clarified and expanded before any new TDM exception is included in legislation, as discussed in Question 28 below.

Question 16. Are you aware of any individuals or bodies with specific licensing needs that should be taken into account?

Dr Paula Westenberger:

My research highlights that the heritage sector (in its full breadth and diversity) and researchers in such contexts should be taken into further consideration in AI and copyright policy, as further developed in my paper (cited above) and Questions 28 and 29 below.

The heritage sector is diverse, formed not only by libraries, archives, galleries and museums, but also natural heritage institutions, heritage sites, historic houses, intangible heritage stakeholders, and others, from different sizes and regions. Different institutions house different materials, and each kind of collection bears special characteristics that can have different and complex copyright implications. It is thus important to engage with the sector in its full diversity to understand the different copyright issues facing these stakeholders.

Bespoke and non-commercial AI models used as tools for research or collections management in heritage contexts present different copyright law implications - particularly regarding "non-commercial research" - compared to commercial generative AI tools. The current policy focus on commercial generative AI presents a missed opportunity to infuse crucial considerations on ethics, bias and cultural impact into AI and copyright regulation, that heritage practitioners are well equipped to reflect on. Heritage stakeholders (including researchers) can thus provide important insights to shape better policies and regulation.

Understanding the risk-averse nature of the heritage sector is also an important consideration, as this reflects on attitudes towards legal uncertainties, which I explore further below (and in my forthcoming paper).

Question 28. Does the existing data mining exception for non-commercial research remain fit for purpose?

Dr Paula Westenberger:

The current TDM exception in s 29A is not fit for purpose for non-commercial research, particularly in heritage contexts.

As noted in the *Living With Machines* project book, the exception has proved difficult to use in innovative research involving diverse datasets, limiting its effectiveness in supporting national priority research in the intersection of technology and culture.[1]

The response below summarises some of the arguments made in my forthcoming paper 'Artificial Intelligence for Cultural Heritage Research: the challenges in UK copyright law and policy' (by Dr Paula Westenberger and Dr Despoina Farmaki), the abstract for which I include below:

> *"Artificial intelligence (AI) is revolutionising our relationship with cultural heritage, enhancing access to, engagement with and preservation of collections and heritage sites. AI is also being used as a valuable research tool in the context of heritage collections. However, as materials protected by copyright may be used in AI development, training and use, copyright law can become an obstacle to important AI deployments in the heritage sector, an area which is currently understudied from the United Kingdom (UK) perspective. This article explores the intricate interplay between cultural heritage, AI and copyright law, demonstrating the main copyright law and policy challenges facing cultural heritage professionals and researchers in using AI in the UK for heritage research. It highlights the complexity and uncertainties as regards the current Text and Data Mining exception in the UK Copyright Designs and Patents Act, emphasising the need for an improved legal framework that balances copyright protection with the benefits of*

[1] Ahnert R, Griffin E, Ridge M, and Tolfo G. Collaborative Historical Research in the age of big data. Cambridge University Press (2023) p. 28. Available at: https://www.cambridge.org/core/elements/collaborative-historical-research-in-the-age-of-big-data/839C422CCAA6C1699DE8D353B3A1960D

AI for cultural heritage research and management. It also reveals the underrepresentation of the heritage sector in AI regulation and copyright policy discussions in the UK. This exploration underscores the imperative for an inclusive policy dialogue that considers the perspectives and evidence of the cultural heritage sector in its full breadth and diversity (including related researchers) in shaping copyright law reform and AI regulation, and for further research to be carried out in this field."

The main issues, as argued in my paper referred to above, are:

- Definition of lawful access is unclear;
- Lawful access requirements may restrict the enjoyment of the exception and may represent contractual override in practice (which is not permitted by the exception);
- Prohibition to transfer copies in CDPA, s 29(A)(2)(a) is unclear and unsuitable for current collaborative forms of AI research: organisations holding relevant data often partner with those possessing computational resources/expertise needed for AI projects. I believe the TDM research exception should support this practice by allowing transfer of datasets between research partners in such cases;
- Web mining: it is unclear the extent to which content placed online falls under lawful access;
- Risk-averse attitudes and lack of resources in the sector: this can essentially deter organisations from engaging in important AI research, especially in view of the lack of clarity of the exception;
- Issues of bias: relying on copyright permissions, or the copyright status of materials, to select research datasets may compromise AI research quality;
- "Non-commercial" research: boundaries between commercial and non-commercial can become blurred (including in the heritage sector, where public-private partnerships can occur in digitisation projects). Focusing on non-profit organisations/uses may be an alternative solution.

This excerpt from our paper summarizes the issues, particularly in light of the new TDM exception being proposed in section C.1 of the consultation:

> "Notably, the consultation proposes a new exception for commercial TDM including an opt-out provision for rights holders, allowing them to exclude their works from AI training datasets. While this aims to protect copyright interests, the possibility of opt-outs raises concerns about potential biases, omissions, and incomplete datasets that could skew and compromise AI research. This is particularly problematic if this provision was to apply in research and heritage contexts, which could become the case in light of the uncertainties regarding the current non-commercial research TDM exception and risk-averse attitudes of heritage stakeholders ... It is unclear how the proposed new exception will impact the existing TDM exception for non-commercial research, which we believe is not fit for purpose for heritage research, and should be clarified and expanded … before the introduction of any new exception. The boundaries between any new commercial TDM exception and the non-commercial research exception must be carefully and clearly delineated, to resolve existing issues and protect and promote AI research in heritage contexts." (Westenberger and Farmaki, forthcoming).

Question 29. Should copyright rules relating to AI consider factors such as the purpose of an AI model, or the size of an AI firm?

Dr Paula Westenberger:

Yes, I believe the purpose of an AI model should be considered. It is important to note, however, that looking at purpose alone might not provide enough clarity depending on how the purpose is framed: see the discussion above and in our paper on the blurred boundaries between commercial versus non-commercial purposes. It might be important to also look at the beneficiary of the provision. In my view, heritage and research uses that are in the public interest, and public heritage and research institutions, deserve particular consideration.

Other exceptions are specifically tailored to the needs and characteristics of the heritage and research sectors, so it is not something unfamiliar to copyright law to look at specific purposes or beneficiaries,

particularly when discussing exceptions. What seems important is that those rules are clearly crafted, and in close consultation with the relevant stakeholders, considering the practical issues facing the relevant sector. Suggesting "a more rigorous and empirically focused framework through which to assess … the drafting of copyright exceptions", and focusing specifically on cultural institutions, see Hudson E. *Drafting Copyright Exceptions: From the Law in Books to the Law in Action* (Cambridge University Press 2020) p.13.

It is problematic that this question is framed as enquiring on the size of the "AI firm". In a discussion on s 29A, aimed at TDM for non-commercial research, the question should, importantly, in my view, enquire on the size of the research or heritage organisation. Consultation documents tend to overemphasize AI firms and creative industries, failing to address important public sector organisations and stakeholders such as those in heritage.

More proactive and substantial engagement with the heritage and research sectors is required in copyright and AI policy making. Public consultations such as this, albeit a great opportunity to contribute to discussions, may not necessarily be accessible/feasible to many such stakeholders. Consultations so far are being drafted in a manner that in my view does not appropriately invite heritage sector contributions, and provide too short of a response timeframe (including over the festive period). As such, it does not allow sufficient time for stakeholder engagement events and research to be carried out that could reveal important evidence to the questions posed. Therefore, I believe that post-consultation engagement strategies, including targeted events such as roundtables with appropriate and diverse representation, would be important to try and capture the views and evidence of the heritage sector in its full breadth and diversity.

See also my response to Question 16 above.

Question 30. Are you in favour of maintaining current protection for computer-generated works? If yes, please explain whether and how you currently rely on this provision.

Dr. Francesca Mazzi:

The provision should be maintained. There is no sufficient evidence of detrimental effect from an economic standpoint, and the moral argument that only human-created works deserve protection, and that protection of CGWs is to the detriment of human creation is flawed. The role of humans in computer-generated works, whether with AI or with other technologies, is a determinant one. Should we reach a point where autonomous systems can evaluate at their discretion whether to create and what to create, then we should probably evaluate other legal issues that can lead to copyright protection of autonomously generated works, such as legal personhood, insurance, liability. But CGW is a provision that allows to protect all those work in the middle between that stage of autonomy and now, where users and developers still have a determinative function underlying, for example, Gen AI creations. Also, the fact that it is not clear whether or not the provision has been used to far, and to what extent, does not tell much about its potential in the future, especially considering that some creators might have avoided relying on it due to legal uncertainty.

Question 31. Do you have views on how the provision should be interpreted?

Dr. Francesca Mazzi:

Guidelines on how to seek protection for CGW could be a good way to ensure more legal clarity, if the decision is to allow the protection of CGW for Gen AI works. In that sense, suggesting a "copyright by design" tool for Gen AI systems, that delivers a proof of interaction between the user and the system (prompt, refinement, selection of the image to protect) could be a good option to show human input in the creative process. [2]

[2] Francesca Mazzi, 'Authorship in Artificial Intelligence-Generated Works: Exploring Originality in Text Prompts and Artificial Intelligence Outputs through Philosophical Foundations of Copyright and Collage Protection' (2024) 27 The Journal of World Intellectual Property 410.

In that case, particular attention should be paid to automation of copyright claims, i.e. the creation of automated system to prove fake interactions between AI and a user, to create a lot of copyrighted works that do to actually have a real human input in it. Something similar happened in relation to patent to ruin prior art by automatically assembling data (see for example "All Cloems"). Therefore, the feasibility of this option should be considered from a technical standpoint, and from a policy standpoint it might be useful to evaluate whether registration for CGW could be an option (considering the associated costs and bureaucratic burdens).

Question 32. Would computer-generated works legislation benefit from greater legal clarity, for example to clarify the originality requirement? If so, how should it be clarified?

Dr. Francesca Mazzi:

The idea of "copyright by design" would allow to get copyright protection based on the originality standard of Infopaq, author's own intellectual creation, by granting copyright protection only to a specific, selected outcome that is the result of a specific prompt/set of prompts. The prompt and the selection arguably amount to intellectual creation, as visible from photograph (one could argue that it is the same thing). This allows enough left for others to innovate (for example, using the same prompt and selecting another result, since also the same AI does not generate always the same output for the same prompt) in line with Lockean proviso of leaving enough as good for others.

Question 33. Should other changes be made to the scope of computer-generated protection?

Dr. Francesca Mazzi:

Clear guidelines should be released (and updated when necessary due to the rapid evolution of the technology) to clarify the scope of CGW protection when it comes to identify the boundaries between human creativity using AI, and AI autonomously generating, to specify that the provision intends to promote human input into AI - and not to over reward AI ownership, for the return of investments of the AI itself is another IP aspect (potentially covered by patents, trade secrets, etc) and the two should remain separated. This is in line with copyright human-centric rationale.

Question 34. Would reforming the computer-generated works provision have an impact on you or your organisation? If so, how? Please provide quantitative information where possible.

Dr. Francesca Mazzi:

Reforming computer-generated works provision would have a positive impact in terms of legal certainty. I would know the extent to which I can protect the works that I generate with AI, in personal and professional capacity, being it artistic, literary, or other types of works.

Question 35. Are you in favour of removing copyright protection for computer-generated works without a human author?

Dr. Francesca Mazzi:

The concept of computer-generated works without a human author should be clarified by the guidelines suggested above. If it is possible to prove human input that amount to author's own intellectual creation, then it should be possible to grant copyright protection to CGW. In the case of AI generating autonomously, copyright protection should be carefully considered as an entirely different scenario. The idea is to maintain the CGW provision and explain the interpretation of necessary arrangements for the work, to maintain protection for the human creativity in collaboration with AI.

Question 36. What would be the economic impact of doing this? Please provide quantitative information where possible.

Dr. Francesca Mazzi:

Quantitative information related to the economic effect of increased legal certainty require a certain level of data for modelling that is not available for the time being. Among others, different stakeholders might have different copyright strategies. Some might be interested in having copyright protection for the exploitation of works, some others might prefer to not claim it to favour spread and achieve visibility. However, what is sure is that a similar provision would be in favour of the weaker stakeholders: the users, the authors, that do not have means to develop AI and can rely on its use with more legal certainty about the future recognition of their works. This would align with copyright purpose.

Moreover, given the important role of AI developers in providing to an extent the necessary arrangement, some options could be considered: whether to always report in the authorship certificate what AI has been used (and therefore, what is the underlying company) for visibility, whether to evaluate co-authorship by default of the legal entity owning the AI, together with the user, or something similar. This policy decision should consider the fact that normally the AI company already has IP protection over the AI as an IP asset, therefore one of the proposed scheme to recognise the role of the AI in the creative process should make sure that over rewarding is avoided, since it would in turn lead to monopolies. An option could be to be co-author but with different degrees of contribution, so that the majority of potential economic revenues goes to the user.

Question 37. Would the removal of the current CGW provision affect you or your organisation? Please provide quantitative information where possible.

Dr. Francesca Mazzi:

As mentioned above, quantitative information related to the economic effect of increased legal certainty require a certain level of data for modelling that is not available for the time being. Among others, different stakeholders might have different copyright strategies. The removal of the provision is likely to create more legal certainty, yes, but unless another provision is introduced to protect AI-assisted creations (in line with what has been recently decided at the USPTO) the grey area will remain, or it will become bigger. Moreover, from the standpoint of economy of legal changes, removing a provision from the CDPA to introduce another one requires more costs than just releasing guidelines on how to define the scope of CGW.

Question 38. Does the current approach to liability in AI-generated outputs allow effective enforcement of copyright?

Dr Zoi Krokida:

The current approach to liability in AI-generated outputs does not permit effective copyright enforcement. Whilst the user will be subject to primary liability for copyright infringement, it seems unclear which form of liability could be imposed to the provider of an AI model.

Attempts to shape the liability of AI providers at the EU level have not been fruitful. In particular, the Proposal for an AI Liability Directive envisaged a fault-based liability regime to compensate any damage caused by AI system. It introduced new rules for damages caused by creating a 'rebuttable presumption of causality'. However, the Proposal for an AI Liability Directive was withdrawn on 11 February 2025.

At the UK level, tort law doctrines could be applicable in the context of liability for the provider of AI model for copyright infringement, either primary or secondary liability. In particular, the provider of AI model could be secondary liable through authorization as per s. 16(2) CDPA. Another option might be to hold provider of AI model liable as accessory through the common law doctrine of joint- tortfeasance using procurement or combination.

Notwithstanding the potential judicial approaches, my suggestion would be to turn on s. 19 of the E-Commerce Regulations 2002 which regulates hosting provider's liability. As outlined in my article 'Large

Language Models and EU Copyright intermediary liability: Quo vadis?',[3] which informs my response, the provider of AI model could be seen as search engine by analogy. This is due to its function, namely to generate content upon users' textual prompts. As per the line of case law,[4] search engines could fall within the definition of hosting providers. Therefore, drawing parallels with search engine, the provider of AI model would be subject to the liability immunity provisions as set forth in s. 19 of the E-Commerce Regulations 2002 and can escape from liability if it does not have actual knowledge of the infringing copyright output or if it removes the infringing copyright output upon being notified by the rights holders. In addition, copyright holders can apply for injunctive relief to enforce their rights. In particular, as per s. 97A CDPA, the provider of AI model might be subject to blocking injunction if it has actual knowledge of another person using its service to infringe copyright. The blocking injunction can be granted irrespective of the liability of the provider of AI model.[5]

Question 39. What steps should AI providers take to avoid copyright infringing outputs?

Dr Zoi Krokida:

Many AI providers resort to filtering technology to filter users' textual prompts and thus block copyright infringing outputs. Crucially, such technology performs a high margin of errors due to their lack of ability to distinguish between lawful and unlawful content.[6] However, filtering technologies are now more advanced and sophisticated, without overlooking the errors. A telling example is content ID system of YouTube. It has been estimated that its Database includes 50 million works which amount to 600 years of audio-visual material.

To mitigate potential errors, a co-regulatory framework could be established.[7] In particular, a supervisory body that will oversee the operation of providers of AI models can achieve greater legal certainty. At the EU level, the EU AI Act has established AI offices in EU member states that would request information and documentation for AI providers in order to demonstrate compliance with the obligations set in EU AI Act.[8]

At the UK level, Ofcom could be assigned to audit the filtering tools before their use, or during the operation, to ensure the correct functioning of the filtering technology. Ofcom is the regulator for online safety as per the Online Safety Act and therefore has already the know-how and the resources to address infringements in new technologies.

[3] Zoi Krokida, 'Large language models and EU intermediary copyright liability: quo vadis?' (2024) 6 European Intellectual Property Review 361.
[4] Joined Cases C-236/08, C-237/08 and C-238/08 of *Google France/Inc. v. Louis Vuitton Malletier* (2009) ECLI:EU:C:2009:569 para. 134.
[5] Case C-324/09, L'Oréal SA and Others v eBay International AG and Others [2009] ECLI:EU:C:2011:474.
[6] J Urban, J Karaganis and B Schofield, 'Notice and Takedown in everyday practice' (2017) Berkeley Law Policy Paper 4; S Jacques, K Garstka, M Hviid and J Street, 'Automated anti-piracy systems as copyright enforcement mechanism: a need to consider cultural diversity' (2018) 40 European Intellectual Property Review 218
[7] Zoi Krokida, *Internet service provider liability for copyright and trade mark infringement: Towards an EU co-regulatory framework* (Hart Publishing, 2022).
[8] Article 54 of Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence.