Introduction

The Coalition for Creativity (C4C) – more details below – welcomes the opportunity to share its views with the Irish Department of Enterprise, Trade and Employment (DETE) on the policy questionnaire on the relationship between generative Artificial Intelligence and copyright and related rights (document 11575/24) circulated by the Hungarian Council Presidency.

About C4C

C4C is a broad-based coalition that seeks an informed debate on how copyright can more effectively promote innovation, access, and creativity. C4C brings together libraries, scientific and research institutions, digital rights groups, technology businesses, and educational and cultural heritage institutions that share a common view on copyright. See our members here: https://coalition4creativity.org/about-us/.

Detailed Response

Please find below our comments on selected questions.

2. Training of AI Models

Questions

> 2.1. Even though the EU legal framework provides legal certainty as regards the rules applicable to the training of AI, are there still, in your views, questions or doubts related to the use of copyright protected content in the training of AI models?

We agree that the AI Act confirms the applicability of the TDM exception and consider that the training process always falls under the conditions of Articles 3 and 4. It is also important to understand and enforce that the opt-out possibility foreseen under Article 4 is not a blanket prohibition: it does not override copyright exceptions or limitations.

Questions do remain about the balance between protecting copyright holders and fostering innovation. A balanced copyright framework is critical to allow responsible AI innovation to thrive, while also giving creators the opportunity to protect their works. The EU Directive on Copyright in the Digital Single Market (CDSM Directive) has paved the way by providing copyright exceptions for text and data mining (TDM), which allows AI developers to use copyrighted material unless the rightsholders opt-out. This framework has enabled innovation to flourish by providing clarity around what is permissible under the law.

However, if researchers were limited to using only public domain data sets, the input for machine learning would be restricted, leading to potentially biassed or skewed outputs. Therefore, while legal certainty exists, challenges arise in ensuring the balance between these objectives in terms of implementation and enforcement.

This balance is especially important when looking at the implementation of the TDM provisions, as too often they seem linked solely to discussions about generative AI, even though TDM applies to a much broader spectrum of AI applications and uses.

2.2. Do you think that practical means such as the introduction of certain standards or the development of an EU-wide database, etc., could be introduced in order to provide more legal certainty within the EU regarding the functioning of the opt-out system? If yes, what practical tools would you envisage in this field?

Practical tools such as machine-readable mechanisms for opting out of text and data mining (TDM) activities already exist but more nuance can be added to them to avoid black and white situations where rightholders are either "in or out" of training datasets: there is hence a need for more opt-out granularity, e.g. between full TDM opt-out (no-tdm) and opt-out specific to generative AI training (no-generative-ai).

It should also be noted that opt-outs should not be interpreted as being retroactive: policies should acknowledge that existing AI models cannot be "un-trained".

The creation of an EU-wide database of copyrighted works has been suggested as a means of offering more transparency. However, setting up such a database would be complex and insufficient in comparison to machine-readable opt-out mechanisms. Additionally, a centralised database would not solve the practical challenges faced by AI developers, who would still need to ensure compliance with the rights reserved by individual rightsholders.

It is also important to avoid introducing a one-size-fits-all approach too early, as this could hinder innovation by not allowing the necessary flexibility to accommodate various use cases.

2.3. Do you have in place or plan to introduce any measures in order to ensure the diversity of the training data set used in the context of generative AI?

We fully agree with the statement that "In the longer term the lack of certain data and the more frequent use of generative AI tools might result in a detrimental effect on cultural diversity and less data available for innovation and public good". We would just add that a data winter is already starting to appear and that mid- to short-term effects of data shortages should be expected if policies are too restrictive.

Indeed, ensuring diversity in AI training datasets is crucial. A wide variety of data is needed to allow AI systems to learn from different perspectives, cultural contexts, and forms of language. This is essential to avoid bias and ensure fairness in AI-generated outputs.

3. The Protection of Outputs

Questions

3.2. In your view, what aspects might be considered when assessing the eligibility for copyright or related right protection of a subject matter created by AI tools?

When determining whether a work created by AI should be eligible for copyright protection, the key consideration is the level of human creative intervention involved. Copyright law is designed to incentivise and reward human creativity. As AI systems do not require an incentive to create, extending copyright protection to AI-generated works without sufficient human involvement may not align with the fundamental principles of copyright. This suggests that any protection granted should be based on a case-by-case analysis, focusing on the extent of human intervention in the creative process.

3.3. In your opinion, is the rationale behind incentivising human creativity by providing exclusive rights in the form of intellectual property protection less relevant in the case of the creation of AI-assisted works?

Yes, the rationale for incentivising human creativity through intellectual property rights is less relevant in cases where AI plays a significant role in the creative process. Copyright is meant to encourage human innovation by rewarding the time, effort, and creativity involved. However, when AI tools are used to generate or assist in the creation of works, the need for such incentives diminishes, as AI systems do not require motivation in the same way humans do.

This implies that a case-by-case approach will be necessary, whereby the extent of human involvement in the creative process is a key factor to determine eligibility for protection.

3.4. According to your view, would it be adequate to introduce new copyright rules on AI-generated and/or AI-assisted works, such as creating a sui generis right, or other specific related rights in this context? If yes, what features might such protection entail?

No, this is not appropriate.

The creation of such additional exclusive, proprietary rights prove to be detrimental, as evidenced by the negative effects flowing out of the adoption of the Database Directive's sui generis right. Therefore, we strongly oppose the creation of sui generis rights or other specific related rights.

The Database Directive has illustrated that once new rights are attributed they are almost impossible to repeal, even if they end-up falling short of their objective.

We consider that additional layers of protection of information create disproportionate burdens for those willing to utilise this information, as it requires more efforts from them to negotiate licences or clear rights, and thus unnecessarily hinders access and increases costs. Moreover, such additional restrictions often do not generate any benefits, quite the contrary.

Professor James Boyle, of the Duke University School of Law, has cautioned that "setting intellectual property rights too high can actually stunt innovation."

4. Transparency

Questions

4.1. Taking into account the labelling and watermarking provisions of the AI Act, do you consider that specific additional obligations should be considered for deployers (users such as creators using generative AI) to label other types of generative AI content and what would be the justification for this?

While there are provisions for labelling and watermarking AI-generated content, introducing further obligations could lead to challenges. The distinction between content generated by AI and human-created content is becoming increasingly blurred. For example, many photo retouching tools already use AI to enhance images, making it difficult to draw a clear line between AI-generated and non-AI-generated content. Similarly, in most webinars, speakers replace their background with AI generated alternatives. Should this be labelled as AI?

Over-labelling content could also cause confusion and lead to phenomena like "notification fatigue," where users begin to ignore labels. There is also the risk of the "implied truth effect," where users may assume that unlabeled content is more trustworthy than labelled content. Therefore, while labelling may have a role in enhancing transparency, it must be implemented carefully to avoid these unintended consequences.

4.2. When considering such possible obligation, do you think it is necessary to differentiate between various types of works and other subject matter (e.g., audio, visual, audiovisual or textual content) from a watermarking/labelling perspective? Are there specific areas where such labelling (e.g., in the case of performances) would be more justified from a copyright perspective?

Yes, there should be differentiation between various types of content and a nuanced approach is needed to ensure that labelling enhances transparency without stifling creativity or confusing users.

5. Remuneration

Questions

5.1. In your view, what measures could be taken at EU level to facilitate the conclusion of licenses between rightholders and AI developers?

First, for the sake of clarity, the TDM regime under the DCDSM is not a licensing regime or a remuneration scheme.

Secondly, at this stage, introducing new measures to facilitate licensing between rightholders and AI developers seems premature. The current copyright regime, particularly the exceptions for TDM, already provides a balanced framework that enables the use of copyrighted works in AI training while protecting the rights of creators. Licensing requirements could create unnecessary burdens and hinder innovation, especially given the large amount of data required for AI training. Any licensing regime would also disproportionately affect smaller AI developers who lack the resources to engage in large-scale licensing negotiations, and hence runs the risk of creating AI database gatekeepers. We refer you to our answer to question 2.3 as regards the need for diversity in data.

Finally, we would like to stress that collective licensing structures are currently not a valid option as (1) copyright collecting societies are usually limited in scope as opposed to the vast array of data present in AI models, (2) it is unclear at this stage if collecting societies have the authority to express opt-outs, as the reservation right befalls rightholders themselves.

5.2 Do you consider that there is a justified reason to introduce any kind of a specific remuneration regime in the context of generative AI activities?

Introducing a specific remuneration regime for AI activities should be approached with caution. The use of data in AI training is inherently transformative, serving a purpose distinct from the original works. The current EU copyright framework already allows rightholders to reserve their rights, which provides a mechanism for protection. Additional remuneration schemes could interfere with these existing frameworks and create unnecessary complexity in the AI development process.

It could also threaten the diversity of datasets in the EU, with consequences on smaller players (see our response to question 5.4) and on the representativity and inclusion of Europe's cultural diversity and values in the outputs generated by AI.

5.3. In your view, are there any specific sectors or aspects of the generative AI creating process where the setting up of such a remuneration scheme would be more appropriate?

No, we do not consider it appropriate for any sector.

5.4. Do you consider that specific measures would be needed in this context to ensure that small EU AI providers have access to quality data for training their models?

Yes, restrictions on access to training data could disproportionately harm smaller AI providers, who may not have the resources to secure high-quality data. Such measures could inhibit innovation by preventing smaller developers from competing with larger players who have greater access to data and licensing capabilities. Therefore, any additional measures introduced should ensure that small and medium-sized AI developers are not unfairly disadvantaged.

6. Liability for Copyright Infringements

Questions

6.1. In your view, is there a justified reason to introduce any kind of specific liability regime in relation to copyright infringement in the context of generative AI activities or are the existing general rules on infringement of copyright and related rights sufficient?

The existing rules at the EU level provide a sufficient framework for addressing copyright infringement in the context of AI.

While there are concerns that AI systems could generate content that closely mirrors existing works, research has shown that such instances are extremely rare and are the reflection of either a poorly engineered model or a purpose-driven (also called "adversarial") set of prompts by certain users. Holding AI developers strictly liable for infringing outputs could have a chilling effect on innovation, as developers would face excessive liability even after taking reasonable steps to prevent infringements. Any liability for AI-generated infringements should attach to the user who deliberately prompts the system to produce infringing content, rather than to the AI developer.

We would also like to stress that the suggestion of using the liability regime of Article 17 of the CDSM Directive as a "source of inspiration" is inappropriate as that regime was set in place for a totally different situation, namely communication to the public of copyrighted content by content-sharing platforms, whereas generative AI is very much a private one-on-one exchange between a user and an LLM.

It is actually contrary to the principles of freedom of speech and even freedom of thought for AI outputs to be censored too heavily: a recent article by Calvet-Bademunt and Mchangama points out that major chatbots routinely refuse to produce certain outputs—not necessarily because these outputs would be illegal or even harmful, but because the companies behind these tools fear backlash, negative press, or legal liabilities. The result? A form of self-censorship that limits the potential of these AI tools to serve as platforms for free expression and thought exploration.

This "brainstorming" aspect has notably been highlighted by Philippe Van Damme, Deputy Director General of the European Commission's DG DIGIT, who wrote on LinkedIn about the roll-out of the EC's internal AI-tool 'GPT@EC' that "it can help Commission staff drafting texts, summarising documents, brainstorming, developing software code and much more."

6.2. If yes, in your view, does this regime need to be introduced regarding the input or the output related infringements?

See our response to question 6.1.

7. Policy and International Context

Questions

7.1. Do you consider that, based on the discussions regarding the above topics, introducing further copyright-specific legislation on the aspects of the relationship between AI and copyright law on the EU level would be appropriate or needed?

No, introducing further copyright-specific legislation at this stage would be premature. AI and copyright are already heavily regulated through the CDSM Directive and the AI Act, and it is important to assess the impact of these existing frameworks before considering additional regulations. The CDSM Directive is not scheduled for review until 2026, which provides a reasonable timeline to evaluate how the current legal framework is working. Over-regulation could stifle innovation and limit the benefits that AI can offer.