Culture Action Europe & Michael Culture Association

AI & Digital Action Group

November 2024

Considerations regarding the implementation of the European Union's Artificial Intelligence Act

1. Artificial Intelligence (AI), especially generative AI, depends on culture and uses cultural data.

The development of artificial intelligence, praised by the European Union as 'an opportunity for EU industrial players to boost their competitiveness' (Draghi report), heavily relies on high-quality, diverse, human-generated data. Generative AI models, designed to produce new content, are general-purpose AI (GPAI) models that create without task-specific specialisations but derive patterns and structures from massive cultural data pools (photos, videos, texts, music, cultural heritage (meta)data, etc.).

As the demand for cultural data grows, high-quality data shortages lead to unauthorised usage of creative content to fuel these systems.

2. Copyright concerns persist despite the AI Act allowing for opt-outs from data mining for AI training purposes.

While the recently adopted AI Act and ruling in the Kneschke vs LAION case recognise that crawling data for AI training falls under the text and data mining (TDM) exception, allowing rightsholders to opt out of having their data used for AI training in line with Article 4(3) of the Copyright Directive (EU) 2019/790, significant concerns remain within the cultural and creative sectors.

GPAI model training lacks direct 'opt-in' authorisation from creators and other rightsholders whose copyrighted works are used without permission.

In recent and ongoing lawsuits worldwide, plaintiffs accuse big tech companies of copyright infringement, alleging they ignore opt-out mechanisms and allow crawlers to bypass restrictions and access copyrighted content for AI training.

Many GPAI models lack transparency and do not reveal their training data sources in detail.

Moreover, the training of generative GPAI models relies not only on semantic analysis of the works but also extracts syntactic information, including elements of copyright-protected expression.

3. Fair compensation for the use of cultural data in AI training remains unresolved under the AI Act.

The right to property, including intellectual property, is a fundamental right according to the Charter of Fundamental Rights of the European Union. Whenever someone is deprived of their possessions in the public interest and in the cases and under the conditions provided for by law, their loss should be subject to fair compensation being paid in good time.

However, the AI Act does not offer a comprehensive framework addressing the need for explicit authorisation ('opt-in') or fair remuneration for the usage of copyrighted works in generative GPAI training.

The lack of provisions enabling direct compensation or authorisation poses challenges to creators' rights and intensifies copyright debates, as highlighted by the Society of Audiovisual Authors and other organisations representing writers, performers, composers, songwriters, screen directors, screenwriters, visual artists, journalists, translators and other creative workers.

4. Cultural sector's contributions to AI development must be valued and included in decision-making.

While recognising the public interest in AI and AI's contribution to scientific progress and economic growth, we stress the importance of properly acknowledging the role of the cultural sector and human creators in AI development. Their contributions should be valued both for their instrumental role (providing data for AI models training) and deontological significance (contributing to the safe, secure and trustworthy development and use of human-centric AI that respects human rights, creators' works, and cultural diversity). The cultural sectors should have institutionalised acknowledgement in the decision-making process regarding AI development.

5. A sufficiently detailed summary of the content used for AI training, policies to identify and comply with reservations of rights from text and data mining, and labelling of AI-generated content are three key obligations for AI providers that directly impact the cultural sector.

Culture Action Europe and Michael Culture Association place particular emphasis on the following provisions of the AI Act, which will have a substantial impact on the cultural and creative sectors:

- **Article 53 (1c):** the obligation for providers of GPAI models to put in place a policy to comply with Union law on copyright and related rights, and in particular to identify and comply with, including through state-of-the-art technologies, a reservation of rights expressed pursuant to Article 4(3) of Directive (EU) 2019/790;
- **Article 53 (1d):** the obligation for providers of GPAI models to draw up and make publicly available a sufficiently detailed summary about the content used for training of the general purpose AI model, according to a template provided by the AI Office;

- **Article 50 (4):** the transparency obligation for providers and deployers of the AI systems that generate or manipulate image, audio or video content constituting a deep fake, to disclose that the content has been artificially generated or manipulated.

6. Considerations regarding the opt-out recognition policies.

With regard to Article 53(1c), we consider it necessary that the Code of Practice and related documentation developed for the implementation of the AI Act reflect the following key considerations:

- **machine-readable opt-out standards should be user-friendly, resilient to content sharing and alteration**, and easily accessible for rightsholders, especially small and independent content creators. AI providers should commit to using generic user agents (crawlers), ensuring that rightsholders are not forced to continually adapt to evolving crawlers. We stress that implementing complex, tech-heavy opt-out mechanisms would disproportionately disadvantage small and independent creators. If only rightsholders with substantial technical and legal resources can effectively protect their rights, it will disproportionately benefit larger players, exacerbating existing inequalities in the cultural and creative sectors;

- the Code of Practice needs to ensure that **the terms and conditions on websites are sufficient for opting out**, in line with Recital 18 of the Copyright Directive (EU) 2019/790, which states that rights can be reserved using machine-readable means such as metadata and website terms and conditions. Moreover, in a recent ruling on the case of Robert Kneschke vs. LAION (case No. 310 O 227/23), the German court was inclined to consider an opt-out published in natural language within the terms of use of a website an effective machine-readable TDM opt-out;

- **the Code of Practice should address the use of both location-based and unit-based identifiers**. The former is more relevant for text works published online, while the latter is better suited for media files distributed as independent units. For unit-based identifiers, the ISCC standard (International Standard Content Code) is suggested;

- **a standardised vocabulary** should be established to allow rightsholders to specify permitted uses of their content, whether for AI training, general data mining, or specific research contexts.

7. Considerations regarding a sufficiently detailed summary of the content used for AI training.

With regard to Article 53(1d), we endorse the Blueprint of the template for the summary of content used to train general-purpose AI models. We consider it necessary that a sufficiently detailed summary of the content used for training the general-purpose AI model include the following elements related to diversity and

copyright, inter alia:

- policies implemented to respect opt-outs under Article 4 of the Copyright Directive (EU) 2019/790;
- proportions of data from relevant categories (e.g., linguistic, regional) used in the training data to reflect the diversity of data;
- steps taken to ensure diversity and representativeness of the training data across categories such as demographics and languages;
- detailed information on the sampling methods used to select training data.

The template should be designed with the primary intention of being practical for those seeking to enforce their rights using it.

8. A financial contribution scheme where AI providers support and compensate small and independent creators should be considered.

Given the frequent use of copyrighted content for AI training without consent, especially from small and independent creators, it is worth considering the establishment of a fund or financial contribution from AI model providers and deployers to support European small and independent creators and cultural and creative sector representatives. The specifics of this scheme could be further defined through a dialogue among the cultural and creative sectors, AI providers, EU institutions, and national governments.

A proposed financial contribution scheme would mitigate the so-called economic 'displacement effects' caused by AI systems capable of mimicking human creativity.

It's important to note that financial contributions should not be seen as a pay-off but rather as compensation for the previous use of cultural content. This approach should be supplemented by appropriate licensing agreements to ensure fair and lawful use going forward.

9. Rightsholders from the cultural and creative sectors must be integrated into the AI Act governance framework at both EU and national levels.

Given how the AI Act merges copyright law (a private law domain where rightsholders can seek remedies for infringement in civil courts) with AI regulation (a public law domain where enforcement
is carried out by public authorities), cultural sector stakeholders should become part of the European Artificial Intelligence Board, the Advisory Forum, the Scientific Panel of Independent Experts, as well as within national competent authorities (notifying authorities and market surveillance authorities).

Having the cultural sector properly represented in AI Act-related bodies means we can address concerns quickly and avoid the drawn-out process of taking each case

through individual legal proceedings before these issues even reach the relevant authorities. Additionally, given the non-binding nature of the Code of Practice, the cultural sector should actively participate in further consultations with the European Standards Organisations on the harmonised standards drafting.

10. AI system providers should bear copyright obligations and be held accountable for copyright infringements.

The AI Act distinguishes between providers of AI *models* and providers of AI *systems* based on these models.

More clarity is needed regarding the scope of obligations and the entities to which they apply. It is important that copyright-related obligations extend to downstream providers, particularly when GPAI models—originally developed by non-profit organisations for scientific purposes under the protections of Article 3 of the Copyright Directive (EU) 2019/790—are subsequently used in AI systems for commercial purposes.

Licensing agreements between GPAI model providers and system providers should extend obligations across the entire value chain to the downstream industry.

11. The AI Act governance framework should include a mechanism that enables rightsholders to swiftly report copyright infringement by AI models and systems providers to the relevant authorities.

Article 85 of the AI Act grants any natural or legal person the right to file a complaint with a market surveillance authority regarding violations of the AI Act's provisions. These complaints are to be addressed in accordance with Regulation (EU) 2019/1020. However, this Regulation does not cover the protection of intellectual property rights (Recital 17).

At the same time, the AI Act allows the AI Office to 'investigate possible infringements of the rules on providers of general-purpose AI models both on its own initiative, following the results of its monitoring activities, or upon request from market surveillance authorities in line with the conditions set out in this Regulation' (Recital 162 of the AI Act).

Therefore, it is crucial to establish a straightforward mechanism specifically for creators and other rightsholders to report issues concerning the use of their data. We propose a one-stop-shop digital desk solution within national competent authorities, the AI Office, or other relevant EU-level bodies to streamline and facilitate this process.

12. The AI regulatory sandbox envisioned in Article 57 of the AI Act should actively explore solutions to protect rightsholders' rights.

Article 57 of the AI Act provides for the establishment of regulatory sandboxes in EU

member states. A regulatory sandbox is a controlled environment which allows for the development, training, testing, and validation of innovative AI systems for a limited time before they are launched on the market or put into service. National competent authorities could encourage AI providers to explore copyright protection and fair remuneration technologies as part of regulatory sandboxes.

For example, blockchain can be used to track the use of *creators'* and other rightsholders' content in AI models. Similar to how streaming platforms monitor music plays to distribute royalties to artists,
these systems could be adapted to ensure proper compensation for the use of creative content in AI development.

13. The cultural sector's response to the AI Act should go hand in hand with proactive, systemic efforts to elevate culture's place in the AI landscape.

This could involve developing AI tools tailored to the sector's specific needs, establishing new licensing frameworks, and further creating and promoting data spaces that enable creators to monetise and control their own data on their own terms.

The cultural sector's position should be reflected in the forthcoming EU strategic documents and legislative initiatives, such as the AI strategy for cultural and creative industries, Apply AI Strategy, European AI Research Council, and EU Cloud and AI Development Act, as outlined in the mission letters to the Executive Vice-President-designate for Tech Sovereignty, Security and Democracy and the Commissioner-designate for Intergenerational Fairness, Youth, Culture and Sport.

With the upcoming evaluation of the Copyright Directive in 2026, it is crucial to examine the potential for an opt-in model that ensures creators provide explicit, verifiable consent before their works are used in AI training.

For decades, European progress in science and innovation has rested on two key pillars: respect for intellectual property, which protects creators' autonomy and livelihoods, and the free exchange of ideas, essential for democratic governance. With the rise of digital technologies and AI, this balance is shifting. We acknowledge that the EU's AI policy needs to account for different public interest policy objectives but this must in no way infringe on or override the rights of the authors and performers to their work and the fair remuneration of their work as they are the ones who drive Europe's advancements in science, culture, heritage, and democracy.