

Dear AI Act negotiators,

The undersigned represent libraries, scientific and research institutions, digital rights groups, and educational institutions.

Our concern - Yes to General Transparency, No to Copyright Specific Provisions

We write to you today to express our deep concerns regarding Article 28b (4) (c) as included in the European Parliament text, and which mandates that providers of foundation models for generative AI systems document and publicly disclose a detailed summary of training data protected under copyright law.

While the intent behind this requirement is to enhance transparency for rightsholders, we firmly believe that it is fundamentally flawed and should be removed for several compelling reasons.

Effective Recently Adopted Existing Framework - Article 4 EUCD

First and foremost, the existing Directive 2019/790 (EUCD) provides a robust framework that effectively addresses the concerns of rightsholders. It establishes transparent and manageable text and data mining (TDM) obligations, offering a balance between protecting rightsholders' interests and facilitating AI-related activities. Commercial TDM activities fall under the exception clause in Article 4 EUCD unless rightsholders actively opt out, ensuring they retain control over the use of their works.

Unreasonable Shift of Responsibility and Bypass of Existing Safeguards

In stark contrast, the proposed transparency provision in Article 28b (4) (c) unfairly shifts the responsibility of identifying copyright-protected data from rightsholders to AI model providers, creating an unworkable and unreasonable burden. There is simply no comprehensive, up-to-date information to identify underlying rights ownership due to the absence of registration requirements under international and EU copyright laws, inconsistent metadata, and a low threshold for copyright protection. This shift also bypasses the checks and balances established by the existing copyright enforcement framework, violating the principles of proportionality and fairness.

Infeasible "Sufficiently Detailed" Summaries

Moreover, the requirement for a "sufficiently detailed" summary of the use of training data is infeasible for datasets trained on web-crawled content, as it effectively entails disclosing the entirety of web content. The dynamic nature of the Internet further complicates matters, rendering summaries obsolete within a short timeframe.

Conclusion: Remove Article 28b (4) (c) and Any Other Copyright References and Focus on Implementing Article 4 of the EUCD

In light of these arguments, we urge EU lawmakers to reconsider and remove Article 28b (4) (c) from the final text of the AI Act. The existing Article 4 EUCD provides an effective solution for rightsholders' interests in transparency and control, as confirmed by the European Commission in its answer to a Parliamentary question¹. Work must be undertaken outside of the AI Act to enable and support stakeholder efforts on opt-out mechanisms that allow rights holders to express rights reservation relating to commercial text and data mining (TDM) practices in a machine readable manner.

We therefore ask you to refrain from any intervention that interferes with the carefully balanced European copyright framework.

Sincerely,

[Coalition for Creativity](#) (C4C) - Brings together over 40 organisations representing libraries, scientific and research institutions, digital rights groups, technology businesses, and educational and cultural heritage institutions that share a common view on copyright (*)

[D3 - Defesa dos Direitos Digitais](#) - Portuguese digital rights organisation

[Electronic Frontier Foundation](#) (EFF) - The leading nonprofit organisation defending civil liberties in the digital world, founded in 1990

[Europa Aperta](#) - Italian association created to promote and spread a pro-European culture based on progressive values

[LIBER Europe](#) - The Association of European Research Libraries (LIBER) is the voice of Europe's research library community, representing more than 420 national, university and other libraries

[Xnet](#), Institute for Democratic Digitalisation - Spanish digital rights organisation

#noisiamorete - Italian digital rights association

¹ See parliamentary question

https://www.europarl.europa.eu/doceo/document/E-9-2023-000479_EN.html and reply

https://www.europarl.europa.eu/doceo/document/E-9-2023-000479-ASW_EN.html

(*): This letter was not supported by European Digital Rights (EDRi)