



Suggestions for the EU Artificial Intelligence Act

(Version of 14 June 2023)

in addition to "IU Position Paper - Generative AI: Copyright status and recommendations for action"

The Copyright Initiative strongly supports the European Parliament's compromise proposal for the AI Act of June 14, 2023. However, some regulations should still be clarified and supplemented:

Expanding Transparency Requirements

The Copyright Initiative welcomes the inclusion of transparency requirements in Article 28b, Paragraph 4 of the AI Act. Providers of generative AI models are obliged to document and make publicly accessible a sufficiently detailed summary of the use of copyrighted training data. However, from the perspective of the Copyright Initiative, a summary of training data alone is not sufficient to ensure that authors and other related rights holders can enforce their claims. Instead, a comprehensive and up-to-date list of the content used by generative AI for training, input, storage, or any other purpose is required. This is essential to ensure the unambiguous identification of each specific piece of content. By obliging providers to "consider the state of the art and the relevant harmonised standards and specifications", the opportunity should be created to impose on providers the integration of content identifiers that are still under development, such as the ISCC standard¹.

Mandatory Approval for AI Training

The Copyright Initiative assumes that training for generative AI does not fall under the definition of text and data mining according to Article 2, No. 2 of the DSM Directive, because the reproductions are not made for the purpose of generating information such as patterns, trends, correlations or the like. Furthermore, Article 7, Paragraph 2 of the DSM Directive in conjunction with Article 5, Paragraph 5 of the InfoSoc Directive, contains the so-called Three-Step Test. The Three-Step Test stipulates, at the first and second stage, that an exception shall only be applied in certain special cases which do not conflict with a normal exploitation of the work or other subject matter. In accordance with the settled case law of the ECJ, the provisions of a directive which derogate from a general principle established by that directive must be interpreted strictly². Finally, the European legislator states in Recital 6 of the DSM Directive:

"The exceptions and limitations provided for in this Directive seek to achieve a fair balance between the rights and interests of authors and other rightholders, on the one hand, and of users on the other. They can be applied only in certain special cases that do not conflict with the normal exploitation of the works or other subject matter and do not unreasonably prejudice the legitimate interests of the rightholders."

¹ Information about ISCC can be found at: <https://iscc.codes> /// ISCC currently has the status of "Draft International Standard" ISO/DIS 24138, the ISO project page: <https://www.iso.org/standard/77899.html>

² ECJ, Judgment of 10 April 2014, ACI Adam BV and others v. Stichting de ThuisKopie, Stichting Onderhandeligen Thuskopie vergoeding, C-435/12, ECLI:EU:C:2014:254, paragraph 22.



However, the mass incorporation of copyrighted works for training purposes in the context of generative AI leads to exactly that. It affects the normal exploitation of copyrighted works because it creates the conditions of replacing authors in many cases, as demonstrated by the wave of layoffs justified with the use of AI at the Axel Springer publishing company (Germany), for example. At the very least, the exploitation of the work is significantly impeded by a free competitive offering.

As a result, the use or other processing of works for the purpose of training by generative AI should only be permissible with the prior explicit permission of the respective rightholder. The EU legislator should clarify this within the framework of the AI Act.

In the event that the courts nevertheless conclude that AI training constitutes text and data mining, there must be a practical opt-out option for AI training that can also be explained through terms and conditions. Therefore, both legal interpretations of text and data mining must be addressed and regulated in the AI Act. Furthermore, in this case, the implementation of compensation for AI training is necessary. Without appropriate regulations, there could be a massive quantitative and qualitative collapse in the creative industry in the foreseeable future, with unpredictable consequences for the democratic opinion-forming process.

Opt-Out Regulation and Compensation Requirement

Authors and related rights holders must not be discriminated against for the use of this opt-out option (e.g., in terms of discoverability through other online platforms or search engines). Such a regulation is necessary in view of the monopoly position of major platform and search engine operators and their various connections with providers of AI tools.

General Disclosure Requirement

Finally, in addition to the disclosure requirement for deep fakes, there should be a general labelling requirement for AI-generated content. Users have the right to know who they are dealing with, in order to prevent a loss of trust in any content. The Copyright Initiative also hopes to create a competitive advantage for publishers who work with humans rather than robots. Nevertheless, people who create must be able to use artificial intelligence to optimise their works. Therefore, editing one's own work with AI tools should not require labelling. This distinction also appears sensible in terms of the work of collecting societies, as they also need to know whether an author or related rights holder can assert claims or whether it is not a human work.

In addition, a general disclosure requirement helps prevent "model collapse" or the breeding of a "degenerative AI," which can occur when AI systems are trained with AI-generated content, essentially self-replicating.



Suggested amendments to the AI Act

Art. 28b (new)	Amendment
<p>Obligations of the provider of a foundation model</p> <p>[Paragraphs 1-3: no changes]</p> <p>4. Providers of foundation models used in AI systems specifically intended to generate, with varying levels of autonomy, content such as complex text, images, audio, or video (“generative AI”) and providers who specialise a foundation model into a generative AI system, shall in addition</p> <p>a) comply with the transparency obligations outlined in Article 52 (1),</p> <p>b) train, and where applicable, design and develop the foundation model in such a way as to ensure adequate safeguards against the generation of content in breach of Union law in line with the generally-acknowledged state of the art, and without prejudice to fundamental rights, including the freedom of expression,</p> <p>c) without prejudice to Union or national or Union legislation on copyright, document and make publicly available a sufficiently detailed summary of the use of training data protected under copyright law.</p>	<p>Obligations of the provider of a foundation model</p> <p>[Paragraphs 1-3: no changes]</p> <p>4. Providers of foundation models used in AI systems specifically intended to generate, with varying levels of autonomy, content such as complex text, images, audio, or video (“generative AI”) and providers who specialise a foundation model into a generative AI system, shall in addition</p> <p>a) comply with the transparency obligations outlined in Article 52 (1),</p> <p>b) train, and where applicable, design and develop the foundation model in such a way as to ensure adequate safeguards against the generation of content in breach of Union law in line with the generally-acknowledged state of the art, and without prejudice to fundamental rights, including the freedom of expression,</p> <p>c) without prejudice to Union or national or Union legislation on copyright, document and regularly make publicly available a sufficiently detailed summary of the use of training data protected under copyright law complete, comprehensive and up-to-date list of content that has been used by generative AI for training, input, storage or any other purpose. This information shall include clear identification of any individual piece of content, such as the exact uniform resource locator (URL), and,</p>



	<p>where necessary, additional information enabling the identification of that content, including the name of the publisher or other provider of that content concerned. Providers shall take into account the generally accepted state of the art and the relevant harmonized standards and specifications.</p> <p>5. Any use or other processing of works or other subject matter protected by intellectual property rights for training, input, storage or any other purpose by generative AI shall only be permitted under the condition of a prior expressed authorisation of the respective rightholder. As regards text and data mining pursuant to Article 4(3) of Directive (EU) 2019/790 the reservation of use of works and other subject matter by their rightholders can be expressed by the rightholder concerned in any way, including via machine-readable means or in the terms and conditions of their website or service.</p> <p>Generative AI providers shall maintain effective, non bypassable, and non-discriminatory technical means necessary for them to implement the reservation of rights to use works and other subject matter. The reservation of rights shall not result in any disadvantage or discrimination for the rightholder.</p> <p>6. Providers of generative AI must ensure fair and appropriate remuneration for any use or other processing of works and other subject matter for training, input, storage or any other purpose to their respective rightholders. Rightholders may not waive their right to remuneration.</p>
--	--



	<p>7. Any use or other processing of works or other subject matter published within the Union for training, input, or any other purpose shall be deemed to have been made within the Union.</p>
<p>Art. 52</p>	<p>Amendment</p>
<p>Transparency obligations for certain AI systems [Paragraphs 1-2: no changes]</p> <p>3. Users of an AI system that generates or manipulates text, audio or visual content that would falsely appear to be authentic or truthful and which features depictions of people appearing to say or do things they did not say or do, without their consent ('deep fake'), shall disclose in an appropriate, timely, clear and visible manner that the content has been artificially generated or manipulated, as well as, whenever possible, the name of the natural or legal person that generated or manipulated it. Disclosure shall mean labelling the content in a way that informs that the content is inauthentic and that is clearly visible for the recipient of that content. To label the content, users shall take into account the generally acknowledged state of the art and relevant harmonised standards and specifications.</p>	<p>Transparency obligations for certain AI systems [Paragraphs 1-2: no changes]</p> <p>3. Users of an AI system that generates or manipulates text, audio or visual content that would falsely appear to be authentic or truthful and which features depictions of people appearing to say or do things they did not say or do, without their consent ('deep fake'), shall disclose in an appropriate, timely, clear and visible manner that the content has been artificially generated or manipulated, as well as, whenever possible, the name of the natural or legal person that generated or manipulated it. Disclosure shall mean labelling the content in a way that informs that the content is inauthentic and that is clearly visible for the recipient of that content. To label the content, users shall take into account the generally acknowledged state of the art and relevant harmonised standards and specifications.</p> <p>AI generated text, audio, or visual output must disclose in a reasonable, timely, clear, and visible manner that the content was artificially generated. Disclosure shall mean labelling the content in a way that informs that the content was generated by AI and that is clearly visible to the recipient of that content. To label the content, users shall take into account the</p>



<p>3a. Paragraph 3 shall not apply where the use of an AI system that generates or manipulates text, audio or visual content is authorized by law or if it is necessary for the exercise of the right to freedom of expression and the right to freedom of the arts and sciences guaranteed in the Charter of Fundamental Rights of the EU, and subject to appropriate safeguards for the rights and freedoms of third parties. Where the content forms part of an evidently creative, satirical, artistic or fictional cinematographic, video games visuals and analogous work or programme, transparency obligations set out in paragraph 3 are limited to disclosing of the existence of such generated or manipulated content in an appropriate clear and visible manner that does not hamper the display of the work and disclosing the applicable copyrights, where relevant. It shall also not prevent law enforcement authorities from using AI systems intended to detect deep fakes and prevent, investigate and prosecute criminal offences linked with their use.</p>	<p>generally acknowledged state of the art and relevant harmonized standards and specifications. AI supported editing of human works and other subject matter are not subject to disclosure requirements.</p> <p>3a. The disclosure obligation according to paragraph 3 does not apply if a law explicitly provides for it or if the exercise of freedom of expression, artistic freedom or scientific freedom from the Charter of Fundamental Rights of the EU makes the absence of disclosure necessary. Where the content forms only a minor part of an evidently creative, satirical, artistic or fictional cinematographic, video games visuals, musical or analogous work or programme, transparency obligations set out in paragraph 3 are limited to disclosing of the existence of such generated or manipulated content in an appropriate clear and visible manner that does not hamper the display of the work and disclosing the applicable copyrights, where relevant. It shall also not prevent law enforcement authorities from using AI systems intended to detect deep fakes and prevent, investigate and prosecute criminal offences linked with their use.</p>
---	---

For further queries contact:

Hanna Möllers, legal adviser and Deputy General Secretary of the German journalists' association Deutscher Journalisten Verband (DJV) moe@djv.de

Berlin, 19 September 2023

Initiative Urheberrecht

Markgrafendamm 24, Haus 18 | 10245 Berlin

0160 9095 4016 | www.urheber.info

Katharina Uppenbrink, Geschäftsführerin, Initiative Urheberrecht

katharina.uppenbrink@urheber.info

Matthias Hornschuh, Komponist & Sprecher der Kreativen in der Initiative Urheberrecht

matthias.hornschuh@urheber.info