

EARE Amendments to MEP Axel Voss' INI Report on AI and Copyright

| INI Report | EARE's Suggestions |
|--|-------------------------|
| <p>D. whereas generative AI (GenAI) is a type of artificial intelligence that, unlike traditional AI systems that only classify or predict, creates content, such as text, images, music, videos and code, often mimicking human creativity, thereby relying on pre-existing content, including copyright-protected materials;</p> | <p>To remove</p> |
| <p style="text-align: center;"><i>Justification</i></p> <p><i>We should align the definition of an AI system with that established in the AI Act. It is risky to differentiate between generative AI and AI systems when such a distinction is not present in either the DSM Directive or the AI Act. Both legal texts provide a framework for what constitutes an AI system – introducing distinctions in this report can lead to inconsistencies in how AI tech is governed and could increase the already existing complexity. Alignment with what has been agreed so far should be pursued to ensure a coherent framework.</i></p> <p><i>Moreover, non-GenAI system can do more than classify or predict. Neural networks are used with those systems as well.</i></p> | |
| <p>F. whereas the key legal questions about the interplay between GenAI and copyright and related rights include whether the new kind of use of copyrighted works and other subject matter in training datasets is lawful under EU law and what the status of AI generated content should be;</p> | <p>To remove</p> |
| <p style="text-align: center;"><i>Justification</i></p> <p><i>We advise against revisiting a debate that has already been resolved. The question of whether Text and Data Mining (TDM) applies to Generative AI was addressed following the publication of the DSM Directive in multiple Commission's statement and subsequent discussions on the AI Act. We should focus instead on making the system workable for everyone.</i></p> | |
| <p>G. whereas the reference to the CDSM Directive in the AI Act is inadequate and</p> | <p>To remove</p> |

| | |
|--|--|
| <p>fails to provide an appropriate and proportionate solution; whereas copyright and related rights, as fundamental rights enshrined in the Charter of fundamental rights of the European Union, are not overridden by the AI Act;</p> | |
| <p style="text-align: center;"><i>Justification</i></p> <p><i>We advise against revisiting a debate that has already been resolved. The question of whether Text and Data Mining (TDM) applies to Generative AI was addressed following the publication of the DSM Directive in multiple Commission's statement and subsequent discussions on the AI Act. We should focus instead on making the system workable for everyone.</i></p> | |
| <p>H. whereas fair remuneration for the use of protected content is the backbone of the creative industry in Europe; whereas the use of content protected by copyright and related rights as training data for generative AI models, without remuneration to rights holders, creates a systemic imbalance in the copyright ecosystem to their detriment, thereby undermining the economic sustainability of the creative sector in the European Union, particularly given that, at present, rights holders cannot easily or effectively exercise their right to opt out from the exception provided for in Article 4 of the CDSM Directive, nor verify whether their opt-out has been respected;</p> | <p>H. whereas fair remuneration for the use of protected content is the backbone of the creative industry in Europe; whereas the use of content protected by copyright and related rights as training data for generative AI models is permitted under copyright exceptions and limitations in the EU and does not undermine rightsholders legitimate interests; whereas IP frameworks should be maintained to enable rightsholders the development of new business models as technology evolves, based on existing copyright laws, which protect the interest of rightsholders, while allowing the public to learn and innovate. without remuneration to rights holders, creates a systemic imbalance in the copyright ecosystem to their detriment, thereby undermining the economic sustainability of the creative sector in the European Union, particularly given that, at present, rights holders cannot easily or effectively exercise their right to opt out from the exception provided for in Article 4 of the CDSM Directive, nor verify whether their opt-out has been respected; Moreover, it is also important to both</p> |

| | |
|--|---|
| | <p>enable innovation and recognize the potential benefits that AI can bring to content creators, such as new markets and innovative tools. Further, measures in the Code of Practice require the development of standards that enable rightsholders to express an opt out for the use of their works in AI training. These standards will ensure that AI developers can reliably identify and respect such opt out signals, allowing rightsholders to effectively exercise their rights.</p> |
| <p><i>Justification</i></p> <p><i>This statement incorrectly assumes that rightsholders’ interests are undermined by AI training and that fair remuneration for the use of protected content is the main source of remuneration for the creative industry in Europe. The livelihood of creators is not solely based on remuneration for the use of protected content. The statement ignores the choice of some creators who intentionally choose to open their works or even release them under permissive licenses to maximize reach and impact. This follows the logic of open-source software, where developers often release their work for free, under licenses that allow others to use, modify, and share it while earning a living through donations, sponsorships, or offering additional services. This shows that not all creative work depends on exclusive rights. The statement also ignores the fact that many creators, including live streamers, podcasters, and digital artists often monetize through fan support and digital platforms.</i></p> <p><i>Similarly, the statement does not consider the AI potential benefits (use of AI tech by creative, exposure, creation of new markets). It implies that the inability to opt out is a significant issue – this undermines the licenses that have already been negotiated on opted out materials, or the effectiveness of this mechanism in a rapidly evolving digital landscape.</i></p> | |
| <p>I. whereas ensuring proper enforcement of the law and a level playing field across the Union requires that European rules on copyright and related rights apply uniformly to all AI providers deploying products or offering services within the European Union, irrespective of their place of</p> | <p>To remove</p> |

| | |
|---|--|
| establishment and of where any use of protected content took place prior to such deployment or offer; | |
| <p><i>Justification</i></p> <p><i>This observation could present significant challenges and complexities as it implies the extraterritoriality of EU law that could end up in conflict with jurisdictions and enforcement challenges while also restricting European researchers, innovators, and startups to access AI technology, impacting their competitiveness globally.</i></p> | |
| J. whereas Article 4 of Directive (EU) 2019/790 on copyright and related rights in the Digital Single Market (CDSM Directive) introduced an exception for the reproduction and the extraction of works and other subject matter for the purpose of text and data mining (TDM), which, under Article 2 of that Directive, is defined as ‘any automated analytical technique aimed at analysing text and data in digital form in order to generate information [...]’; whereas Article 4 was neither drafted nor intended to regulate the specific practices involved in AI training; | J. whereas Article 4 of Directive (EU) 2019/790 on copyright and related rights in the Digital Single Market (CDSM Directive) introduced an exception for the reproduction and the extraction of works and other subject matter for the purpose of text and data mining (TDM), which, under Article 2 of that Directive, is defined as ‘any automated analytical technique aimed at analysing text and data in digital form in order to generate information [...]’; whereas Article 4 was neither drafted nor intended to regulate the specific practices involved in AI training; |
| <p><i>Justification</i></p> <p><i>We advise against revisiting a debate that has already been resolved. The question of whether Text and Data Mining (TDM) applies to Generative AI was addressed following the publication of the DSM Directive in multiple Commission’s statement and subsequent discussions on the AI Act. We should focus instead on making the existing system workable for everyone.</i></p> | |
| K. whereas this new and specific form of use (GenAI training) requires a clarification of the legal conditions under which such training may be conducted; | To remove |
| <p><i>Justification</i></p> <p><i>We advise against revisiting a debate that has already been resolved. The question of whether Text and Data Mining (TDM) applies to Generative AI was addressed following the publication of the DSM Directive in multiple Commission’s statement and subsequent</i></p> | |

| | |
|--|--|
| <i>discussions on the AI Act. We should focus instead on making the system workable for everyone.</i> | |
| N. whereas, in addition to a standardised machine-readable opt-out, rights holders should also have the possibility to register such opt-out in a centralised registry, using a single technological standard and in machine-readable format, potentially managed by the European Union Intellectual Property Office (EUIPO), thereby enabling the effective exclusion of registered works from automated data crawling; | N. whereas, in addition to a standardised machine-readable opt-out, rights holders should also have the possibility to register such opt-out in a centralised registry, using a single technological standard and in machine-readable format, potentially managed by the European Union Intellectual Property Office (EUIPO), thereby enabling TDM users to act on a rights reservation when appropriate without undue burden, and ensuring that the registry respects the unconditional TDM exception that exists under Article 3 and do not overlook the existence of other relevant exceptions under the CDSM Directive; further facilitating the effective exclusion of registered works from automated data crawling |
| <p style="text-align: center;"><i>Justification</i></p> <p><i>EARE acknowledges the potential value of a centralised EU-level register of opt-outs as a tool to help researchers and innovators identify which data can be accessed for AI model development and improvement. It is paramount that any registry is consistent with existing provisions of the CDSM, including that a registry satisfies the requirements of Article 4 of the CDSM that a reservation of rights be expressed in an appropriate manner, such as a machine-readable format. It is important that such a register does not undermine the unconditional TDM exception that exists under Article 3 or overlook the existence of other relevant exceptions.</i></p> <p><i>However, a centralised system within the EUIPO could become a barrier to the diversity and richness of data for researchers and innovators, especially if there is not clear guidance on what qualifies under the TDM exceptions. Moreover, compliance burdens should be carefully assessed and mitigated. The need to continuously monitor the register or implement measures to comply with updated registry entries which are particularly burdensome, should be avoided.</i></p> | |
| O. whereas any GenAI provider should ensure full and detailed transparency concerning all copyright-protected content | To remove |

| | |
|--|--|
| <p>used to train that system, irrespective of the jurisdiction in which the copyright-relevant acts underlying the training were performed; whereas this transparency shall consist in an itemised list identifying each copyright-protected content used for training; whereas the same requirement should apply mutatis mutandis to any subsequent use of content for inference, retrieval-augmented generation or finetuning not only by providers of AI models, as currently stipulated by Article 53 AIA, but also by providers or deployers of AI systems;</p> | |
| <p style="text-align: center;"><i>Justification</i></p> <p><i>Transparency and reproducibility of research is core to the values of EARE.</i></p> <p><i>However, the approach proposed in this article does not appear proportionate and is technically infeasible. Granular disclosure requirements would be unworkable. First, the vast volume of content, with trillions of works available online, combined with the absence of a central registry of these works and the overlapping exclusive rights within a single file, makes such requirements impractical. For example, a digital file of a musical work may include multiple copyright works and authors involved including the composition, sound recording, and lyrics, each potentially involving different authors and rightsholders. Second, during the pre-training phase of AI model development, a significant portion of the collected content is not actually used for training. Moreover, content sourced from the web is often transformed or rewritten. Even if perfect copyright metadata were available, it would be extremely difficult to determine with certainty which specific works were used to train a model. Any proposed granularity measures must also consider the risks of disclosing trade secrets and other highly sensitive information, particularly during the fine-tuning process.</i></p> <p><i>Obligations which touch upon transparency and openness of AI tools need to be proportionate, and any measures should be properly understood, and balanced. It is important to note that the European Commission’s study on improving access to reuse of research results for scientific purposes, a detailed summary of the data used for training can “add a layer of compliance costs for research organizations”. Similarly, for SMEs, and startups, this document can be difficult to implement. For this reason, the European institutions should focus on working with AI providers, research organizations, SMEs, startups, and other relevant stakeholders to monitor the implementation of the template for training data. The EU should focus on the right implementation of the current AI</i></p> | |

| | |
|--|---|
| <p>Act, the Code of Practice and the template of the summary of training data. Before introducing any additional requirements on transparency, these tools should be rightly implemented.</p> | |
| <p>P. whereas such transparency could be facilitated through a trusted intermediary, such as the European Union Intellectual Property Office (EUIPO), which would be responsible for notifying rights holders of the use of their content, thereby enabling them to assert claims in relation to its use for training; such an intermediary should be endowed with the necessary powers and resources to assess whether providers and deployers comply fully with the transparency obligations;</p> | <p>To remove</p> |
| <p><i>Justification</i></p> <p><i>EARE acknowledges the idea of involving EUIPO as a trusted intermediary. However, researchers and startups might face increased bureaucracy and compliance requirements, when using datasets for training AI models. This could slow down research timelines. As the European Commission has embraced the reduction of administrative barriers and EU's competitiveness, the involvement of the EUIPO as a trusted intermediary should be rightly evaluated.</i></p> | |
| <p>Q. whereas, as an alternative to the aforementioned EUIPO register, transparency could also be achieved by enabling rights holders to watermark their works and other protected subject matter, and by requiring AI providers to make available search tools that allow for the detection of such watermarks among the materials used for training;</p> | <p>To remove</p> |
| <p><i>Justification</i></p> <p><i>Imposing watermark detection requirements on researchers would create operational burdens and cost and thereby limit their access to data. EARE considers that the EU should implement the current framework first, including the Code of Practice and the template for training data.</i></p> | |
| <p>R. whereas, in addition to the obligation of full transparency concerning copyright</p> | <p>R. whereas, in addition to the obligation of full transparency concerning</p> |

| | |
|--|--|
| <p>protected works and other protected subject matter, there is a need to establish a mechanism whereby, under certain conditions, the failure by AI providers or deployers to provide complete transparency shall give rise to an irrebuttable presumption that any relevant copyrighted work or other protected subject matter has been used for training purposes, thereby triggering all applicable legal consequences under Union and national law for the infringement of copyright or related rights; whereas, where a court finds in favour of a rights holder on the basis of either such a presumption or of submitted evidence, all reasonable and proportionate legal costs and other expenses shall be borne by the AI provider;</p> | <p>copyrightprotected works and other protected subject matter, there is a need to establish a mechanism whereby, under certain conditions, the failure by AI providers or deployers to provide complete transparency shall give rise to an irrebuttable presumption that any relevant copyrighted work or other protected subject matter has been used for training purposes, thereby triggering all applicable legal consequences under Union and national law for the infringement of copyright or related rights; whereas, where a court finds in favour of a rights holder on the basis of either such a presumption or of submitted evidence, all reasonable and proportionate legal costs and other expenses shall be borne by the AI provider; the establishment of a mechanism to support transparency should be clear, proportionate, and technically feasible, whereby, when a rightsholder succeeds in legal proceedings by demonstrating an infringing use of protected works, appropriate remedies including reasonable and proportionate costs may be borne by the AI provider</p> |
| <p style="text-align: center;"><i>Justification</i></p> <p><i>The irrebuttable presumption introduced in the INI report, which assumes that copyrighted works have been used in training AI models if transparency requirements are not fully met, could negatively impact access to data for researchers and innovators. This approach risks disproportionately affecting researchers and start-ups located in the EU, particularly those lacking the resources to ensure full compliance or who may be unaware of these obligations despite being willing to comply. Since the presumption is irrebuttable, it does not allow small developers or researchers to prove otherwise, even if they have not included opted out works to train AI models. This presumption could further deter researchers and smaller AI startups to develop AI models. Similarly, since the irrebuttable presumption would apply to any form of TDM, it could impact the work of researchers and innovators by discouraging the use of data and potentially facilitating bias into research results. Instead, the EU should focus on the right implementation of the current provisions on transparency established in the AI Act, and the future Code of practice.</i></p> | |

| | |
|---|--|
| <p>Z. whereas, to the extent permitted by international law, the existing principle of territoriality needs to be adapted for the training of GenAI systems in order to ensure that training with European content is subject to European law even if it is realised outside the EU;</p> | <p>To remove</p> |
| <p style="text-align: center;"><i>Justification</i></p> <p><i>This observation could present significant challenges and complexities as it implies the extraterritoriality of EU Law that could end up in conflict with jurisdictions and enforcement challenges while also restricting European researchers, innovators, and startups to access AI technology, impacting their competitiveness globally.</i></p> <p><i>This observation is also a significant departure from existing copyright law by ignoring national legal protections and exceptions. It also makes it nearly impossible to engage with copyrighted works that involve multiple authors or when parties hold rights to such works on a territory-by-territory basis. It would upend any negotiations where the rights are carved up by territory.</i></p> | |
| <p>1. Recommends that the Commission, independently of its planned review of the copyright framework and the CDSM Directive and without presupposing the need for legislative revision, urgently conduct a thorough assessment of whether the existing EU copyright acquis adequately addresses the legal uncertainty and competitive effects associated with the use of protected works and other subject matter for the training of generative AI systems, as well as the dissemination of AI-generated content that may substitute human-created expression.</p> | <p>1. Recommends that the Commission, as part independently of its planned review of the copyright framework and the CDSM Directive and without presupposing the need for legislative revision, urgently conducts a thorough assessment of whether the existing EU copyright acquis. This assessment should include a holistic approach which considers the needs of all actors, including researchers, universities, libraries, cultural organizations, European AI startups, news outlets, and the creative sector when using and developing AI. addresses the legal uncertainty and competitive effects associated with the use of protected works and other subject matter for the training of generative AI systems, as well as the dissemination of</p> |

| | |
|--|--|
| | AI-generated content that may substitute human-created expression. |
| <p style="text-align: center;"><i>Justification</i></p> <p><i>EARE considers that the upcoming review of the EU copyright framework, expected in 2026 should be conducted with a comprehensive and inclusive approach. It is essential that the perspectives and needs of all relevant stakeholders are taken into account, including researchers, universities, libraries, cultural institutions, European AI startups, news media, and the creative sector.</i></p> <p><i>Particular attention should be given to the research ecosystem, whose voices are often underrepresented in discussions on AI and copyright. Researchers and innovators are not only central to the EU's competitiveness but also key drivers of scientific and cultural development.</i></p> <p><i>The review should prioritize the effective implementation of the current copyright framework and the text and data (TDM) exceptions, which are vital tools for enabling research, innovation, and cultural development in Europe.</i></p> <p><i>This process should be given sufficient time to allow for thoughtful reflection and evidence-based policymaking. Rather than rushing into legislative changes, the goal should be to find balanced and sustainable solutions that work for all actors involved, using the unique framework created in the European Union.</i></p> | |
| 2. Further recommends that such assessment aims to uphold a framework in which fair remuneration mechanisms enable the generation of the resources needed for European artistic and creative production to thrive in the context of AI-driven global transformation; | 2. Further recommends that such assessment aims to uphold a framework in which fair investigate remuneration frameworks to understand if mechanisms enable the generation of the resources needed for European artistic and creative production can to thrive in the context of AI-driven global transformation; |
| 2.bis. Further encourages that the assessment considers the special needs of researchers and innovators, as well as the critical role of the text and data mining (TDM) exception in supporting their work and advancing Europe's competitiveness and future discoveries. | To add |
| | |

| | |
|--|---|
| <p style="text-align: center;"><i>Justification</i></p> <p><i>EARE considers that the upcoming review of the EU copyright framework, expected in 2026 should be conducted with a comprehensive and inclusive approach. Particular attention should be given to the research ecosystem, whose voices are often underrepresented in discussions on AI and copyright. Researchers and innovators are not only central to the EU's competitiveness but also key drivers of scientific and cultural development.</i></p> <p><i>The review should prioritize the effective enforcement of the text and data (TDM) exceptions, which are vital tools for enabling research, innovation, and cultural development in Europe.</i></p> <p><i>The ultimate goal should be to find balanced and sustainable solutions that work for all actors involved.</i></p> | |
| <p>4. Calls on the Commission to immediately impose a remuneration obligation on providers of general-purpose AI models and systems in respect of the novel use of content protected by copyright or related rights, with such obligation applying until the reforms envisaged in this report are enacted;</p> | <p>To remove</p> |
| <p style="text-align: center;"><i>Justification</i></p> <p><i>This temporary provision unravels years of practice – since the transposition deadlines of June 2021. It is, moreover, unnecessary, unclear, and risks undermining data access for researchers, cultural institutions and innovators. The TDM exception clarifies that the use of a copyrighted work to train an AI model is not a copyright infringement and therefore does not require remuneration. Further, immediate remuneration (whatever that may mean) obligations would make it harder or more expensive for researchers and European innovators to access large datasets, compromising scientific research and innovation.</i></p> <p><i>Similarly, many research projects involve public-private partnerships. If the provision is not clear enough, these partnerships might be treated as commercial, even if their purpose is for non-commercial scientific research. This could also prevent researchers and innovators to be involved in public-private partnerships.</i></p> <p><i>This immediate obligation may also distort the future review of the copyright framework expected in 2026.</i></p> | |
| <p>6. Supports the clarification of the TDM exception under Article 4 CDSM as regards the main flaws and ambiguities detected thus far in its application, especially as concerns the establishment of a clear</p> | <p>6. Supports the clarification of the TDM exception under Article 4 CDSM through the development of agreed standards or guidelines for implementing opt-outs that are not overly burdensome to</p> |

| | |
|---|--|
| <p>machine-readable standard for the opt-out and the concept of ‘lawful access’;</p> | <p>implement for researchers and innovators, as regards the main flaws and ambiguities detected thus far in its application, especially as concerns the establishment of a clear machine-readable standard for the opt-out, and the concept of ‘lawful access’,</p> |
| <p style="text-align: center;"><i>Justification</i></p> <p><i>Today, researchers and innovators face an inconsistent application of the CDSM Directive across member states, including TDM exceptions and Article 4 CDSM. This legal fragmentation impedes the use of AI and machine learning. Researchers too often refrain from using research tools such as AI and machine learning "because they are afraid of copyright infringement" (see here).</i></p> <p><i>Moreover, the delineation between commercial and non-commercial uses creates legal uncertainty for researchers and AI providers, particularly in the context of public-private partnerships and the TDM exceptions.</i></p> <p><i>EARE supports the development of opt-out standards in a way that benefits innovation and research. Any opt-out should be clear defined, machine readable, and not lead to ambiguity. Policy makers should support industry efforts between rightsholders and AI developers to create standard technological solutions for declaring and reading machine readable opt-outs, with special attention to SMEs and start-ups.</i></p> <p><i>Clarifications do not presuppose additional legislation, and may instead take the form of guidance directed at Member States or relevant actors, such as researchers.</i></p> | |
| <p>7. Believes that a legal framework for GenAI should be established either through the introduction of a dedicated exception to the exclusive rights to reproduction and extraction, distinct from that provided for TDM under Article 4 of the CDSM Directive, or by expanding the scope of that provision to explicitly encompass the training of GenAI, which is currently not covered; stresses that rights holders shall have the right to opt out through a standardised, machine-readable mechanism;</p> | <p>To remove</p> |
| <p style="text-align: center;"><i>Justification</i></p> <p><i>A new legal framework for GenAI is not needed. TDM exceptions were intended to train AI systems, including generative AI. The opt-outs within Article 4 of the DSM Directive were</i></p> | |

expressly intended to allow rightsholders to reserve their rights on TDM activities for commercial purposes. The DSM Directive defines TDM as “any automated analytical technique aimed at analysing text and data in digital form to generate information”. This already covers training AI models, including those used for generative AI, which involves analysing large datasets to extract patterns.

The application of the TDM exception to train AI systems is not a new issue for policymakers and EU institutions and Member States have been aware of these discussions for long time. This is reflected in the negotiations of the DSM Directive and the AI Act. Similarly, Articles 3 and 4 of the DSM Directive as well as the Recital 105 of the AI Act explicitly confirm that TDM exceptions apply when training general purpose AI models. The new interpretation established in the report adds to this uncertainty, creating further confusion for researchers and innovators when using data.

*The introduction of a dedicated exception, distinct from that provided for TDM under Article 4 of the DSM Directive, or expanding the scope of the provision to encompass the training of genAI as proposed in the INI report, would add **unnecessary complexity and confusion for researchers and innovators**, further complicating an already complex regulatory framework. This would prevent researchers and innovators from using content they already have access to, renegotiating access to content for AI purposes, reducing the data available, increasing bias in research, raising costs, and undermining the willingness among researchers and innovators to use data. Ultimately, this would negatively impact the quality of scientific research in the EU.*

8. Recommends that the Commission ensures the compatibility of this new GenAI legal framework with the three-step test of Article 5(5) InfoSoc Directive;

To remove

*Justification
[Same as above]*

9. Recommends assigning the EUIPO responsibility for setting up and managing a central register of opt-outs and, where necessary, for mediating the licensing process, so as to streamline relations between GenAI providers and rights holders, establishing a workable, innovation-friendly framework that supports the Union’s competitiveness without unduly hindering the development

9. Suggests ~~Recommends assigning~~ **assessing the feasibility of designating the EUIPO as the responsible party** ~~responsibility~~ for setting up and managing a central register of opt-outs ~~and, where necessary, for mediating the licensing process,~~ so as to streamline relations between GenAI providers and rights holders, establishing a workable, innovation-friendly framework that

| | |
|---|---|
| <p>of AI technologies; further recommends that both opt-out declarations and licence offers be recorded in machine-readable form in the same register;</p> | <p>supports the Union's competitiveness without unduly hindering the development of AI technologies; further recommends that both opt-out declarations and licence offers be recorded in machine-readable form in the same register in a manner that enables TDM users to act on a rights reservation when appropriate without undue burden; further calls for ensuring that the registry respects the unconditional TDM exception that exists under Article 3 and do not overlook the existence of other relevant exceptions under the CDSM Directive.</p> |
| <p style="text-align: center;"><i>Justification</i></p> <p><i>EARE acknowledges the potential value of a centralised EU-level register of opt-outs as a tool to help researchers and innovators identify which data can be accessed for AI model development and improvement. It is paramount that any registry is consistent with existing provisions of the CDSM, including that a registry satisfies the requirements of Article 4 of the CDSM that a reservation of rights be expressed in an appropriate manner, such as a machine-readable format. It is important that such a register does not undermine the unconditional TDM exception that exists under Article 3 or overlook the existence of other relevant exceptions.</i></p> <p><i>However, a centralised system within the EUIPO could become a barrier to the diversity and richness of data for researchers and innovators, especially if it creates any compliance burdens. Moreover, compliance burdens should be carefully assessed and mitigated. The need to continuously monitor the register or implement measures to comply with updated registry entries which are particularly burdensome, should be avoided.</i></p> <p><i>EARE also acknowledges the idea of involving EUIPO as a trusted intermediary. However, researchers and startups might face increased bureaucracy and compliance requirements, when using datasets for training AI models. This could slow down research timelines. As the European Commission has embraced the reduction of administrative barriers and EU's competitiveness, the involvement of the EUIPO as a trusted intermediary should be rightly evaluated.</i></p> | |
| <p>10. Calls on the Commission to propose the full, actionable transparency and source documentation by providers and</p> | <p>10. Calls on the Commission to focus on the implementation of the transparency requirements of the AI Act, including the</p> |

| | |
|---|---|
| <p>deployers of general-purpose AI models and systems, with regard to the use of any copyright-protected work or other protected subject matter for any purpose, including for inferencing, retrieval-augmented generation, or finetuning, taking into due account the need to protect trade secrets and confidential business information;</p> | <p>Code of Practice and the template to help general-purpose AI providers summarise the content used to train their model; recommends the European Commission to continue working with AI providers and relevant stakeholders to improve the template of training data and consider a simplified version of the template for SMEs, startups, mid-caps, and research organizations, propose the full, actionable transparency and source documentation by providers and deployers of general-purpose AI models and systems; with regard to the use of any copyright-protected work or other protected subject matter for any purpose, including for inferencing, retrieval-augmented generation, or finetuning, taking into due account the need to protect trade secrets and confidential business information.</p> |
| <p style="text-align: center;">Justification</p> <p><i>Transparency is core to the values of EARE. Our organization considers that the European Commission should focus on the right implementation of the transparency requirements established in the AI Act, including the Code of Practice and the template to help general-purpose AI providers summarise the content used to train their model. Before introducing any additional requirements regarding transparency, these tools should be rightly implemented.</i></p> <p><i>We also note that, as recently highlighted in European Commission’s study on improving access to reuse of research results for scientific purposes, a detailed summary of the data used for training can “add a layer of compliance costs for research organizations”. Similarly, for SMEs, and startups, this document can be difficult to implement. For this reason, the impact of this detailed summary on research organizations, SMEs, and startups should be carefully monitored. As the template may be reviewed before the entry into application of the enforcement powers of the AI Office on 2 August 2026, the European Commission should focus on working with AI providers, research organizations, SMEs, startups, and other stakeholders to improve it. Potential improvements could consider exceptions or a simplified version of the template for research organizations, European startups, SMEs, or mid-caps. The EU Commission should also apply a proportionate approach, especially for SMEs, startups, mid-caps and research organizations when assessing the submissions of the training data.</i></p> | |

Similarly, it is essential for SMEs and startups the protection of trade secrets to secure funding. When working with the AI Office to assess the submissions from SMEs and startups, the AI Office should ensure that sensitive or proprietary information is not disclosed.

11. Calls on the Commission to propose the establishment of an irrebuttable presumption that, for any general-purpose AI (GenAI) model or system placed on the Union market, works and other subject matter protected by copyright or related rights have been used for its training where the statutory transparency obligations set out in this resolution have not been fully complied with; further recommends that, where a rights holder succeeds in legal proceedings either on the basis of this presumption or through submitted evidence, any reasonable and proportionate legal costs and other expenses incurred in enforcing such rights shall be borne by the provider of the AI model or system;

11. Calls on the Commission to propose ~~the establishment of an irrebuttable presumption that, for any general-purpose AI (GenAI) model or system placed on the Union market, works and other subject matter protected by copyright or related rights have been used for its training where the statutory transparency obligations set out in this resolution have not been fully complied with;~~ **to ensure that transparency obligations for any general purpose AI (GenAI) model or system placed on the Union market, are clear, proportionate and technical feasible;** further recommends that, where a rights holder succeeds in legal proceedings **to demonstrate that the provider of the AI model or systems has infringed copyright of a protected work, through submitted evidence, appropriate remedies including reasonable and proportionate costs may be borne by the provider of the AI model or system.** ~~either on the basis of this presumption or through submitted evidence, any reasonable and proportionate legal costs and other expenses incurred in enforcing such rights shall be borne by the provider of the AI model or system;~~

Justification

*The **irrebuttable presumption** introduced in the INI report, which assumes that copyrighted works have been used in training AI models if transparency requirements are not fully met, could **negatively impact access to data for researchers and innovators**. This approach risks disproportionately affecting researchers and start-ups located in the EU, particularly those lacking the resources to ensure full compliance or who may be unaware of these obligations despite being willing to comply. Since the presumption is irrebuttable, it does not allow small developers or researchers to prove otherwise, even if*

they have not included opted out works to train AI models. This presumption could further deter researchers and smaller AI startups to develop AI models. Similarly, since the irrebuttable presumption would apply to any form of TDM, it could impact the work of researchers and innovators by discouraging the use of data and potentially facilitating bias into research results. Instead, the EU should focus on the right implementation of the current provisions on transparency established in the AI Act, and the future Code of practice.

14. Notes that nothing in this proposal should be intended to affect scientific or research activities.

To add

Justification

EARE considers it necessary to include an article explicitly stating that none of the provisions in the proposal should be intended to affect scientific or research activities, nor any public-private partnerships already established within the scientific and research ecosystem. This safeguard is essential to ensure that the EU continues to lead scientific research and protect researchers and innovators.