Contribution ID: ceb51729-15b7-414a-bf9d-28e7d641cc0a
Date: 17/09/2024 12:49:36

Multi-stakeholder Consultation FUTURE-PROOF AI ACT: TRUSTWORTHY GENERAL-PURPOSE AI

Fields marked with * are mandatory.

Multi-stakeholder Consultation FUTURE-PROOF AI ACT: TRUSTWORTHY GENERAL-PURPOSE AI

The European AI Office is launching this multi-stakeholder consultation on trustworthy general-purpose AI models in the context of the AI Act. We invite submissions from all stakeholders with relevant expertise and perspectives, particularly from academia, independent experts, industry representatives such as general-purpose AI model providers or downstream providers integrating the general-purpose AI model into their AI system, civil society organisations, rightsholders organisations, and public authorities.

This is an opportunity for all stakeholders to have their say on the topics covered by the first Code of Practice on detailing out rules for providers of general-purpose AI models in the context of the AI Act. It will also inform related work of the AI Office, in particular on the template for the summary about the model training data and accompanying guidance.

Details about the AI Act rules for providers of general-purpose AI models, the Code of Practice, and related work by the AI Office can be found in the background documents available here.

The consultation is available in English and responses can be submitted via this form over a period of seven weeks. Submissions must be completed by Wednesday, 18 September 2024, 18:00 CET.* We encourage

early submissions.

In parallel, stakeholders who wish to participate in the entire process of drawing-up the first Code of Practice can express their interest here by Sunday, 25 August 2024, 18:00 CET.

The questionnaire for this consultation is structured along 3 sections

1. General-purpose AI models: transparency and copyright
    A. Information and documentation to providers of AI systems
    B. Technical documentation to the AI Office and the national competent authorities
    C. Policy to respect Union copyright law
    D. Summary about content used for the training of general-purpose AI models
2. General-purpose AI models with systemic risk
    A. Risk taxonomy
    B. Risk identification and assessment
    C. Technical risk mitigation
    D. Internal risk management and governance for general-purpose AI model providers
3. Reviewing and monitoring the General-Purpose AI Code of Practice

We welcome full or partial replies from all respondents based on their expertise and perspective.

At the end of the questionnaire, you have the option to upload one document to share further information with the AI Office. We provide a template which aligns with the topics covered in the Code of Practice and follows the structure of the Plenary Working Groups. Based on the submissions and answers to the targeted questions, a first draft of the Code of Practice will be developed.

All contributions to this consultation may be made publicly available. Therefore, please do not share any confidential information in your contribution. For organisations, their organisation details would be published while

respondent details can be requested to be anonymised. Individuals can request to have their contribution fully anonymised.

The AI Office will publish a summary of the results of the consultation. Results will be based on aggregated data and respondents will not be directly quoted.

Please allow enough time to submit your application before the deadline to avoid any issues. In case you experience technical problems which prevent you from submitting your application within the deadline, please take screenshots of the issue and the time it occurred.

In case you face any technical difficulties or would like to ask a question, please contact: CNECT-AIOFFICE-CODES-OF-PRACTICE@ec.europa.eu

**The AI Office has announced an extension of the consultation period for the Code of Practice concerning general-purpose AI models, as part of the ongoing implementation of the AI Act. The new deadline, set for 18 September 2024, replaces the previous 10 September cutoff. This will grant stakeholders overall seven weeks to submit their feedback.*

About you

* 1. Do you represent one or more organisations (e.g., industry organisation or civil society organisation) or act in your personal capacity (e.g., independent expert)?

- (●) Organisation(s)
- ( ) In a personal capacity

*

Please specify the name(s) of the organisation(s):

European Writers' Council (EWC)

* First name

Nicole

* Surname

Pfister Fetz

* E-Mail address (this won't be published)

nicole.pfisterfetz@europeanwriterscouncil.eu

* Is your organisation headquartered in the EU?

- (●) Yes
- ( ) No
- ( ) Other (e.g. multiple organisations)

* EU member states

- ( ) AT - Austria
- (●) BE - Belgium
- ( ) BG - Bulgaria
- ( ) HR - Croatia
- ( ) CY - Cyprus
- ( ) CZ - Czechia
- ( ) DK - Denmark
- ( ) EE - Estonia
- ( ) FI - Finland
- ( ) FR - France
- ( ) DE - Germany
- ( ) EL - Greece
- ( ) HU - Hungary
- ( ) IE - Ireland
- ( ) IT - Italy
- ( ) LV - Latvia
- ( ) LT - Lithuania
- ( ) LU - Luxembourg
- ( ) MT - Malta
- ( ) NL - Netherlands
- ( ) PL - Poland
- ( ) PT - Portugal
- ( ) RO - Romania

- SK - Slovak Republic
- SI - Slovenia
- ES - Spain
- SE - Sweden

* What is the size of your organisation?
- Micro (1 to 9 employees)
- Small (10 to 49 employees)
- Medium (50 to 249 employees)
- Large (250 or more employees)
- Other (e.g. multiple organisations)

* Which stakeholder category would you consider yourself in?
- Provider of a general-purpose AI model, or acting on behalf of such providers
- Downstream provider of an AI system based on general-purpose AI models, or acting on behalf of such providers
- Other industry organisation, or acting on behalf of such organisations
- Academia
- Civil Society Organisation
- Rightsholder or a collective management organisation (CMO) or an independent management organisation (IME) or the representative of an organisation acting on behalf of rightsholders (other than a CMO or IME)
- Public authority
- Others

* If you indicated to be a rightsholder, in which sector do you operate?
- Music
- Audiovisual
- Publishing
- Visual Arts
- Video games
- Other

* Please briefly describe the activities of your organisation or yourself:

1000 character(s) maximum

The European Writers' Council (EWC) is the world's largest representation of writers in the book sector, constituted as federation of 50 organisations from 32 countries of the EU, the EEA, and the European non-EU areas. The EWC represents 220,000 writers, publishing in 34 languages and in all genres, also worldwide.
We respond with the explicit proviso that we do not agree with the premise of this consultation. The exception for commercial text and data mining (Art. 4, 2019/790 CDSM Directive) is inapplicable: The statutory language and text of the provision, its conception, and the ratio of the exception indicate that it must not be applied to the training of generative AI models. Hence, the training of generative AI models without the authors consent can be classified as both a copyright infringement and a violation of duties in the AI Act. We call on the AI Office to get the scope of the TDM Art. 3 and Art. 4 exception clarified before the Code of Practice is drafted.

* **Availability for a follow-up conversation**

We may follow up with you for clarification or further discussion if your submission prompts additional interest.

I agree to be contacted by the AI Office for a follow-up conversation to my submission.

- (●) Yes
- ( ) No

All contributions to this consultation may be made publicly available. Therefore, please do not share any confidential information in your contribution. For organisations, their organisation details would be published while respondent details can be requested to be anonymised. Individuals can request to have their contribution fully anonymised. Your e-mail address will never be published.

Please select the privacy option that best suits you. Privacy options default based on the type of respondent selected.

* **Contribution publication privacy settings**

If you represent one or more organisations: All contributions to this consultation may be made publicly available. You can choose whether you would like respondent details to be made public or to remain anonymous.

- **Anonymous.** Only organisation details are published: The type of respondent that you responded to this consultation as, the name of the organisation on whose behalf you reply as well as its size, its presence in or outside the EU and your contribution will be published as received. Your name will not be published. Please do not include any personal data in the contribution itself if you want to remain anonymous.
- **Public.** Organisation details and respondent details are published: The type of respondent that you responded to this consultation as, the name of the organisation on whose behalf you reply as well as its size, its presence in or outside the EU and your contribution will be published as received. Your name will also be published.

Privacy statement

☑ I acknowledge the attached privacy statement.

privacy_statement.pdf

Section 1. General-purpose AI models: transparency and copyright-related rules

A. Information and documentation by general-purpose AI model providers to providers of AI systems

Providers of general-purpose AI models have a particular role and responsibility along the AI value chain, as the models they provide may form the basis for a range of downstream systems, often provided by downstream providers that necessitate a good understanding of the models and their capabilities, both to enable the integration of such models into their products, and to fulfil their obligations under the AI Act or other regulations. Therefore, model providers should draw up, keep up-to-date and make available information and documentation to providers of AI systems who intend to integrate the general-purpose AI model into their AI system. Widely adopted documentation practices include model cards and data sheets.

A minimal set of elements of information and documentation by general-purpose AI model providers to providers of AI systems is already set out in AI Act Annex XII.

1. In the **current state of the art**, for which elements of **information and documentation** by general-purpose AI model providers to providers of AI systems do **practices** exist that, in your view, achieve the **above-mentioned purpose**?

From the list below following AI Act Annex XII, please select all relevant elements.
If such practices exist, please provide **links to relevant material** substantiating your reply, such as model cards, data sheets or templates.

A general description of the general-purpose AI model including:

- [x] **The tasks that the model is intended to perform and the type and nature of AI systems into which it can be integrated;**
- [x] **The acceptable use policies applicable;**
- [x] **The date of release and methods of distribution;**
- [x] **How the model interacts, or can be used to interact, with hardware or software that is not part of the model itself, where applicable;**
- [x] **The versions of relevant software related to the use of the general-purpose AI model, where applicable;**
- [x] **The architecture and number of parameters;**
- [x] **The modality (e.g., text, image) and format of inputs and outputs;**
- [x] **The licence for the model.**

A description of the elements of the model and of the process for its development, including:

- [x] **The technical means (e.g., instructions for use, infrastructure, tools) required for the general-purpose AI model to be integrated into AI systems;**
- [ ] **The modality (e.g., text, image, etc.) and format of the inputs and outputs and their maximum size (e.g., context window length, etc.);**
- [ ] **Information on the data used for training, testing and validation, where applicable, including the type and provenance of data and curation methodologies.**

Alternatively:

- [ ] No practices for any of the listed elements exist that achieve the above-mentioned purpose.

☐ I don't know

Links to relevant material

a) The data provenance initiative: A large scale audit of dataset licensing & attribution in AI, MIT in collaboration with Harvard University, October 2023:
https://arxiv.org/pdf/2310.16787
Quote: „We also observe frequent miscategorization of licenses on widely used dataset hosting sites, with license omission of 70%+ and error rates of 50%+. This points to a crisis in misattribution and informed use of the most popular datasets driving many recent breakthroughs."

b) The Foundation Model Transparency Index v1.1: May 2024, Stanford and Princeton Universities:
https://arxiv.org/pdf/2407.12929
Quote: „Foundation models are increasingly consequential yet extremely opaque. To characterize the status quo, the Foundation Model Transparency Index was launched in October 2023 to measure the transparency of leading foundation model developers. The October 2023 Index (v1.0) assessed 10 major foundation model developers (e.g. OpenAI, Google) on 100 transparency indicators (e.g. does the developer disclose the wages it pays for data labor?). At the time, developers publicly disclosed very limited information with the average score being 37 out of 100."

2. Beyond the minimal set of elements listed in the previous question, are there **other elements** that should be included in **information and documentation** by general-purpose AI model providers to providers of AI systems to achieve the above-mentioned purpose?

- ◉ Yes
- ○ No
- ○ I don't know

Please specify

700 character(s) maximum

Information on: a) whether the model was developed for scientific research by a research organisation; b) how authors/rightsholders can opt-out their works used in development and how to license; c) T&C under which outputs can be used & must be labelled by end-users incl. guidelines for developers interacting with the model, ensuring that the rights of authors are respected; d) internal guidelines, incl. considerations for anonymisation, bias, transparency of sources and curation, incl. mitigation of risks of copyright infringement, of data privacy, moral and personality rights violations; e) confirmation of legal access of data; f) confirmation of trusted data set entities.

Links to relevant material

A Survey on Data Selection for Language Models (University Standford in cooperation with UC Santa Barbara, February 2024, on the taxonomy of works and data sets and relevant information:
https://arxiv.org/pdf/2402.16827v1

B. Technical documentation by general-purpose AI model providers to the AI Office and the national competent authorities

In addition to the provision of information on the general-purpose AI model for its usage by the downstream providers, technical documentation should be prepared and kept up to date by the general-purpose AI model provider for the purpose of making it available, upon request, to the AI Office and the national competent authorities.

A minimal set of elements of such technical documentation of the general-purpose AI model to be made available by providers, upon request, to the AI Office and the national competent authorities is already set out in AI Act Annex XI.

3. In the **current state of the art**, for which elements of **documentation** by general-purpose AI model providers do practices exist that, in your view, provide a **necessary level of information for the above-mentioned purpose**?

From the list below following AI Act Annex XI, please select all relevant elements.
If such practices exist, please provide **links to relevant material** substantiating your reply, such as model cards, data sheets or templates.

A general description of the general-purpose AI model including:

- ☑ **The tasks that the model is intended to perform and the type and nature of AI systems into which it can be integrated**;
- ☐ **The acceptable use policies applicable;**
- ☑ **The date of release and methods of distribution;**
- ☑ **The architecture and number of parameters;**
- ☑ **The modality (e.g., text, image) and format of inputs and outputs;**
- ☐ **The licence.**

A description of the elements of the model, and relevant information of the process for the development, including:

- ☑ **The technical means (e.g., instructions for use, infrastructure, tools) required for the general-purpose AI model to be integrated into AI systems;**
- ☐

The design specifications of the model and training process, including training methodologies and techniques, the key design choices including the rationale and assumptions made; what the model is designed to optimise for and the relevance of the different parameters, as applicable;

- [ ] **Information on the data used for training, testing and validation**, where applicable, including the type and provenance of data and curation methodologies (e.g. cleaning, filtering etc), the number of data points, their scope and main characteristics; how the data was obtained and selected as well as all other measures to detect the unsuitability of data sources and methods to detect identifiable biases, where applicable;
- [x] **the computational resources used to train the model** (e.g. number of floating point operations), training time, and other relevant details related to the training;
- [ ] **known or estimated energy consumption of the model.**

Additional information to be provided by providers of general-purpose AI models with systemic risk:

- [ ] **A detailed description of the evaluation strategies**, including evaluation results, on the basis of available public evaluation protocols and tools or otherwise of other evaluation methodologies. Evaluation strategies shall include evaluation criteria, metrics and the methodology on the identification of limitations;
- [ ] **Where applicable, a detailed description of the measures put in place for the purpose of conducting internal and/or external adversarial testing** (e. g., red teaming), model adaptations, including alignment and fine-tuning;
- [ ] **Where applicable, a detailed description of the system architecture** explaining how software components build or feed into each other and integrate into the overall processing;

Alternatively:

- [ ] No practices for any of the listed elements exist that achieve the above-mentioned purpose.
- [ ] I don't know

Links to relevant material

Related to model cards, this is a concept invented in 2018/2019. But it is NOT in daily practice.
https://huggingface.co/docs/hub/model-card-landscape-analysis

4. Beyond the minimal set of elements listed in the previous question, are there **oth er elements** that should, in your view, be included in **technical documentation** by general-purpose AI model providers **to the AI Office** and the national competent authorities?

- (●) Yes
- ( ) No
- ( ) I don't know

Please specify

700 character(s) maximum

Information on: a) model structure; b) version history, incl. updates, reasons for changes (e.g. improved accuracy, reduced bias, copyright notes…); c) of how the works and data was gathered (incl. commissioned entities e.g. corpora builder, data set developers), curated, including filtering, augmentation process to remove inappropriate context & data; d) of how the model was developed, incl. human interaction, the computational resources, any optimisations applied; e) how the model can be audited by third parties, incl. information on log files and other resources for examination; f) certification of legal access + on data set providers (private partnership with research institutions)

Links to relevant material

The status of the current technical reports by AI developers and its gaps, not meeting yet the requirements of the AI Act to find under https://rettighedsalliancen.com/wp-content/uploads/2024/09/Report-on-AI-model-providers-training-data-transparency-and-enforcement-of-copyrights.pdf - Report on AI model providers' training data transparency and enforcement of copyrights by the Danish Rights Alliance

C. Policy to respect Union copyright law

The AI Act requires providers of general-purpose AI 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.

5. What are, in your view, the main **elements that need to be included in the policy** that providers of general-purpose AI models have to put in place to **comply with Union law on copyright** and related rights, as required by the AI Act?

Please select all relevant options from the list of options suggested below. If selected, please elaborate further on the content of the measures and provide links to any good practices you are aware of.

- [x] Allocation of responsibility within the organisation for the implementation and monitoring of compliance with the policy and the measures therein;
- [x] Measures to identify and comply with the rights reservation from the text and data mining exception pursuant to Article 4(3) of Directive (EU) 2019/790;
- [x] Measures to obtain the authorisation from right holders, where applicable;
- [x] Measures to detect and remove collected copyright protected content for which rights reservation from the text and data mining exception has been expressed pursuant to Article 4(3) of Directive (EU) 2019/790;
- [x] Measures to prevent the generation, in the outputs of the model, of copyright infringing content;
- [x] Means for contact with rightsholders;
- [x] Measures for complaint handling from rightsholders;
- [x] Other
- [ ] I don't know

Please specify

700 character(s) maximum

In principle, the following aspects must be formally clarified prior to CoP drafting: a) the scope of TDM exceptions 2019/790; b) handling of works published before 6.7.2021; c) reliable definitions of „legally accessed“ and „publicy available“. Further, sufficient details for rightholders to provide evidence in a court procedure & to enforce their rights: d) time period of gathering of the work; e) sources & entities (corpora providers, incl. research institutions in private partnerships); f) information on compliance with "legally accessed"; g) licensing methods and remuneration schemes; h) confirmation that TDM opt out is accepted; i) Information on the current life cycle of the Model

Your comments

700 character(s) maximum

The premise, that 2019/790, Art 3 + 4 on TDM cover developing ‘generative AI’, is highly controversial. Accordingly, we respond with the proviso that the use of copyright-protected works for the development of general-purpose AI is a new right, which does not fall under the Art 4-regime. The use of works must only be subject to authorisation, remuneration and, in accordance with Art 19, 2019/790 (EU), right to information on title-specific works licensed + the extent and scope of use within the model, by deployers, and end users. Audits must be conducted. Providers must ensure that AI outputs do not replicate original (parts of) works. Dispute resolution must be installed.

Links to relevant material

a) Joint Statement of 13 Authors’, Artists’, and Performers’ Federations on the Scope of TDM Art 4: https://europeanwriterscouncil.eu/240425_cwos_jointstatement_ai-act/
b) Scientific Study on GenAI development and where the technical and legal scope of TDM ends: https://urheber.info/media/pages/diskurs/ki-training-ist-urheberrechtsverletzung/c943688809-1725462359/executive-summary_engl_final_29-08-2024.pdf

c) On contractual questions and the authors' right to Opt-out, in the AI Tool Kit for the Book Sector: https://europeanwriterscouncil.eu/ai-tool-kit2024/

6. How can, in your view, the policy to be put in place by providers of general-purpose AI models to comply with Union copyright law ensure that providers of those models comply with the **existing solutions for the expression of the text and data mining rights reservation**, pursuant to Article 4(3) of Directive (EU) 2019/790?

Please explain how this can be achieved and specify from the list below the state-of-the-art technologies you are aware of to identify and comply with the right reservations expressed by rightsholders, providing further information and examples.

- [x] Technologies/tools that identify right reservations at the website/domain level
- [x] Technologies/tools that identify right reservations at work level
- [x] Technologies/tools that aggregate the expression of right reservations
- [x] Other
- [ ] I don't know

Please specify

700 character(s) maximum

Aspects to be formally clarified prior to CoP drafting: a) scope of TDM exceptions 2019/790 not covering GenAI; b) handling of works published before 6.7.2021; c) reliable definitions of „legally accessed/publicly available“; d) clarification of “machine-readable” or other sufficient means (Recital 18) incl. for physical works when digitised (Art 8 2019/790)); e) traceable control mechanism to ensure that opt-out is respected by AI developers. Rightholders are obliged to “apply measures to ensure that their reservations are respected” - but not as one-way street! f) clarification of not allowed sublicensing by libraries; g) measures to ensure that GenAI opt out not harms SEO visibility

Your comments

700 character(s) maximum

The premise, that Art. 3+4 2019/790 on TDM cover ‘GenAI’, is highly controversial. Accordingly, we respond with the proviso that the use of copyright-protected works for the development of general-purpose AI is a new right, which does not fall under the Art. 4 regime. The CoP must therefore secure that right reservations are traceable respected by AI developers & providers, e.g. through real-time access + documentation, e.g. via a data directory (e.g. with ISCC+rights declaration methods) hosted by an independent authority, to verify works and title-specific opt-outs or licensing information. AI developers shall contribute significantly to rightsholders' expenses for this.

Links to relevant material

a) Joint Statement of 13 Authors, Artists and Performers Federations on the Scope of TDM Art 4: https://europeanwriterscouncil.eu/240425_cwos_jointstatement_ai-act/
b) Scientific Study on GenAI development and where the technical and legal scope of TDM ends: https://urheber.info/media/pages/diskurs/ki-training-ist-urheberrechtsverletzung/c943688809-1725462359/executive-summary_engl_final_29-08-2024.pdf
c) On contractual questions and the moral right to Opt-out, in the AI Tool Kit for the Book Sector: https://europeanwriterscouncil.eu/ai-tool-kit2024/
d) On the ISO 24138 Standard ISCC plus rights declaration to build up an EU wide / international data directory (NOT a registry like in the US!) to aggregate licensing or rights reservation information and to respect and enforce the rights of authors as primary rightsholders to opt-out: https://iscc.codes ; https://iscc.io/ https://www.youtube.com/watch?v=S1vK8LMK0f4 ;
e) W3C recommendations on TDMRep with Onix: https://docs.tdmai.org/

D. Summary about content used for the training of general-purpose AI models

The AI Act requires providers 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. While due account should be taken of the need to protect trade secrets and confidential business information, the summary is to be generally comprehensive in its scope instead of technically detailed to facilitate parties with legitimate interests, including copyright holders, to exercise and enforce their rights under Union law. The template that should be drafted by the AI Office for the sufficiently detailed summary should be simple, effective, and allow providers to provide the required summary in narrative form.

7. What are in your view the **categories of information** sources that should be presented in the summary to ensure that it comprehensively describes the main sources of data used for the training of the general-purpose AI model?

From the list below, please select all options that you consider relevant.

- [x] Public/ open data repositories
- [x] Content/data publicly available online (e.g. scraped from the internet)
- [x] Proprietary data generated by the provider
- [x] User-generated data obtained through the services or products provided by the provider
- [x] Copyright protected content licensed by rightsholders
- [x]

Other data/content or data sets acquired from third parties (e.g. licensed proprietary databases, data acquired from datahubs, public interest institutions such as libraries etc.)

- [x] Synthetically generated data
- [x] Other
- [ ] I don't know

Please specify

700 character(s) maximum

Clarifications prior to defining this list: a) scope of TDM exceptions not covering GenAI; b) handling of works published before 6.7.21; c) reliable definitions of „legally accessed/publicly available"; d) clarification of "machine-readable" or other sufficient means (Recital 18). The categories above (non-compliant with legal conception) are no basis for a CoP template: incorrect legal assumptions, e.g. special regulation for PII (= non-legal compliant term) or libraries, esp. not under Art 8. 2019/790. ‚Content/data publicly available online' may include protected & public domain works. ‚Public/open data repositories' and ‚Copyright protected content licensed by rightsholders' overlap

If selected, please **specify the level of granularity/detail for each of the selected options**, keeping in mind that AI Act requires the summary to be comprehensive instead of technically detailed and provided in a narrative form to facilitate parties with legitimate interests, including rightsholders, to exercise and enforce their rights under Union law, while taking due account of the need to protect providers' trade secrets and confidential business information. If additional categories should be considered, please specify them and the level of granularity /detail. You can motivate your choice and provide links to any good practices.

700 character(s) maximum

Controversial premise, Art 3+4 2019/790 on TDM cover 'GenAI'. Accordingly, proviso that use of ©-protected works for development of GPAI is a new right, which does not fall under Art 4. Authors as primary rights owners of declaring opt-out & claiming transparency need to be able to 1) discover unauthorised copying, 2) license for AI development, 3) initiate enforcement activities. As authors are obliged to declare opt-out for EACH work, it is essential to disclose always work-specific information on: time period of collection; sources & entities (corpora developers); compliance with "legally accessed"; licensing & remuneration schemes; accepted TDM opt out; life cycle status of the Model.

Links to relevant material

a) Joint Statement of 13 Authors, Artists and Performers Federations on the Scope of TDM Art 4: https://europeanwriterscouncil.eu/240425_cwos_jointstatement_ai-act/
b) Scientific Study on GenAI development and where the technical and legal scope of TDM ends: https://urheber.info/media/pages/diskurs/ki-training-ist-urheberrechtsverletzung/c943688809-1725462359/executive-summary_engl_final_29-08-2024.pdf
c) On contractual questions and the moral right to Opt-out, in the AI Tool Kit for the Book Sector: https://europeanwriterscouncil.eu/ai-tool-kit2024/

d) The status of the current legal copyright transparency by AI developers and its gaps, nor meeting yet the requirements of the CDSM 2019/290 and neither the AI Act to find under https://rettighedsalliancen.com/wp-content/uploads/2024/09/Report-on-AI-model-providers-training-data-transparency-and-enforcement-of-copyrights.pdf (Report on AI model providers' training data transparency and enforcement of copyrights by the Danish Rights Alliance)

8. In your view, should the summary include one or more of the following **characteristics/information about the data used for the training**/of the general-purpose AI model in order to facilitate parties with legitimate interests, including copyright holders, to enforce their rights under Union law?

Please select all relevant options from the list of options suggested below. If selected, please explain your choice and provide links to any good practices.

- [x] Modalities / type of data (text, images, videos, music, etc);
- [x] Nature of the data (personal, non-personal or mixed);
- [x] Time of acquisition/collection of the data;
- [x] Data range of the data (e.g. time span), including date cutoffs
- [x] In case of data scraped from the internet, information about the crawlers used;
- [x] Information about diversity of the data (for example linguistic, geographical, demographic diversity);
- [x] Percentage of each of the main data sources to the overall training/fine-tuning;
- [x] Legal basis for the processing under Union copyright law and data protection law, as applicable;
- [x] Measures taken to address risks to parties with legitimate interests (e.g. measures to identify and respect opt-out from the text and data mining exception, respect data protection and address privacy risks, bias, generation of illegal or harmful content;
- [x] Other
- [ ] I don't know

Please specify

700 character(s) maximum

a) Title/work specific list of works and all creators & rightsholders; b) methods of acquiring and data delivering entities incl. data set curators, corpora builders, research institutions with private partnerships; c) specification of licensing or legal ground, e.g. but not limited to information on CC licenses; d) work specific identification standards, e.g. but not limited to: ISCC+rights declaration identifiers; ISNI Codes; ISBN and DOI Codes; e) provenance of data sets; f) confirmation of liability of corpora builders or further data set providers. Info: https://iscc.codes ; https://iscc.io/ https://www.youtube.com/watch?v=S1vK8LMK0f4

Your comments

700 character(s) maximum

The premise, that Art 3+4 2019/790 on TDM cover ‘GenAI’, is far from settled. Accordingly, we respond with the proviso that the use of ©-protected works for the development of GPAI is a new right, which does not fall under the Art.4 regime. However, GPAI developers & GenAI providers must provide a title-specific work and domain specific summary, incl. detailed disclosure of data sources, legal basis of the scraping (Arts. 3 or 4 2019/790), contractual licensing, permissive licences, e.g. sort of Creative Commons, and respect for personal rights to be sufficient. If models are placed in the EU markets, they must implement right reservations mechanisms for authors and further rightsholders.

Link to relevant material

The status of the current by AI developers and its gaps on used data and works, not meeting yet the requirements of the CDSM Directive and enable authors to enforce their rights is to find under https://rettighedsalliancen.com/wp-content/uploads/2024/09/Report-on-AI-model-providers-training-data-transparency-and-enforcement-of-copyrights.pdf - Report on AI model providers’ training data transparency and enforcement of copyrights by the Danish Rights Alliance

9. Considering the purpose of the summary to provide **meaningful information to facilitate the exercise of the rights** of parties with legitimate interests under Union law, while taking due account of the need to respect **business confidentiality and trade secrets** of providers, what **types of information** in your view are **justified not to be disclosed** in the summary as being not necessary or disproportionate for its purpose described above?

700 character(s) maximum

None are justified to be not disclosed, as this prevents the legally guaranteed exercise and, above all, enforcement of the rights of authors / rightsholders and would constitute a breach of EU Union law. Compromise: Trade secrets excluded from public disclosure in the development data sets must be flagged for confidential inspection by regulatory authorities. In principle, however, alleged economic concerns do not justify the undermining of copyright, especially as the economic impact resulting from the use of GenAI has proven to be destructive for the sources of every GenAI system, the authors, artists, performers. Transparency will be a USP of European AI developers.

Section 2. General-purpose AI models with systemic risk: risk taxonomy, assessment and mitigation

A. Risk taxonomy

Some general-purpose AI models could pose systemic risks, which should be understood to increase with model capabilities and model reach and can arise along the entire lifecycle of the model.

‘Systemic risks’ refer to risks that are specific to the high-impact capabilities of general-purpose AI models (matching or exceeding the capabilities of the most advanced general-purpose AI models); have a significant impact on the Union market due to their reach; or are due to actual or reasonably foreseeable negative effects on public health, safety, public security, fundamental rights, or society as a whole, that can be propagated at scale across the value chain (AI Act Article 3(65)).

Systemic risks are influenced by conditions of misuse, model reliability, model fairness and model security, the level of autonomy of the model, its access to tools, novel or combined modalities, release and distribution strategies, the potential to remove guardrails and other factors.

The Code of Practice should help to establish a risk taxonomy of the type and nature of the systemic risks at Union level, including their sources. The Code should take into account international approaches.

10. Do you consider the following list of **systemic risks** based on AI Act Recital 110 and international approaches to be comprehensive to inform a taxonomy of systemic risks from general-purpose AI models? If additional risks should be considered in your view, please specify.

Systemic risk from model malfunctions

- **Harmful bias and discrimination:** The ways in which models can give rise to harmful bias and discrimination with risks to individuals, communities or societies.
- **Misinformation and harming privacy:** The dissemination of illegal or false content and facilitation of harming privacy with threats to democratic values and human rights.
- **Major accidents:** Risks in relation to major accidents and disruptions of critical sectors, that a particular event could lead to a chain reaction with considerable negative effects that could affect up to an entire city, an entire domain activity or an entire community.
- **Loss of control:** Unintended issues of control relating to alignment with human intent, the effects of interaction and tool use, including for example the capacity to control physical systems, ‘self-replicating’ or training other models.

Systemic risk from malicious use

- **Disinformation:** The facilitation of disinformation and manipulation of public opinion with threats to democratic values and human rights.
- **Chemical, biological, radiological, and nuclear risks:** Dual-use science risks related to ways in which barriers to entry can be lowered, including for weapons development, design acquisition, or use.
- **Cyber offence:** Risks related to offensive cyber capabilities such as the ways in which vulnerability discovery, exploitation, or operational use can be enabled.

Other systemic risks, with reasonably foreseeable negative effects on

- **public health**
- **safety**
- **democratic processes**
- **public and economic security**
- **fundamental rights**
- **the society as a whole.**

☐ Yes, this list of systemic risks is comprehensive.
☑ Further or more specific systemic risks should be considered.
☐ I don't know

Please specify

700 character(s) maximum

Premise, that Art 4 2019/790 covers the use for developing 'GenAI', is legally and technically controversial and a systemic risk within. Plus: damaging "AI business models" have cropped up in the book sector with fake authors/books/translations. Without output labelling & input transparency respecting authors' rights, GenAI enables copyright infringement, disinformation, royalty fraud. Public funding for prizes + grants must be sure that it is honouring human authorship. AI products shall not benefit from the reduced VAT & other benefits legally dedicated to cultural assets. Using GenAI instead human labour will result in lower payments for taxes, into social security & pension schemes.

11. What are in your view **sources of systemic risks** that may stem from the development, the placing on the market, or the use of general-purpose AI models? Systemic risks should be understood to increase with model capabilities and model reach.

Please select all relevant elements from the list. If additional sources should be considered, please specify. You can also provide details on any of the sources or other considerations.

- [x] **Level of autonomy of the model**: The degree to which a general-purpose AI model has the capability to autonomously interact with the world, plan ahead, and pursue goals.
- [x] **Adaptability to learn new, distinct tasks:** The capability of a model to independently acquire skills for different types of tasks.
- [x] **Access to tools:** A model gaining access to tools, such as databases or web browsers, and other affordances in its environment.
- [x] **Novel or combined modalities:** Modalities a model can process as input and generate as output, such as text, images, video, audio or robotic actions.
- [x] **Release and distribution strategies:** The way a model is released, such as under free and open-source license, or otherwise made available on the market.
- [x] **Potential to remove guardrails:** The ability to bypass or disable pre-defined safety constraints or boundaries set up to ensure a model operates within desired parameters and avoids unintended or harmful outcomes.
- [ ] **Amount of computation used for training the model:** Cumulative amount of computation ('compute') used for model training measured in floating point operations as one of the relevant approximations for model capabilities.
- [x] **Data set used for training the model:** Quality or size of the data set used for training the model as a factor influencing model capabilities.
- [x] **Other**
- [ ] **I don't know**

Please specify

700 character(s) maximum

Add "job loss". Due to GenAI, writers, translators, illustrators & audio book narrators, have lost jobs & income: 30% of translators lost commissions, illustrators are reporting 60%. Audio Book narrators are replaced by synthetic voices, writers are forced to use GPT, against lower or even no royalties. To apply opt-out rights reservation is an approx. cost-factor for the EU book sector of ca. 390 million Euros, bare minimum, for the ca. 14 Mio. prior to 2021 published works. To destroy the sources of a value chain of 22.3 billion Euros/year, the authors, is a high risk for the future of the EU: losing innovation and investment. https://europeanwriterscouncil.eu/gai-is-based-on-theft/

Your comments

700 character(s) maximum

A critical source of systemic risk is the development of novel or combined modalities, where models can process and generate various forms of output, such as text, images, video, audio, voices. If not clearly labelled or flagged, the spread of disinformation rise and pose serious threats to democratic processes. To mitigate these risks, it is essential that outputs, particularly images or audiovisual works with synthetic voices, are flagged using state of the art technology. Likewise, “human-readable” labelling is essential to allow democratic opinion forming and to reduce the risk of royalty fraud incl. but not limited to remuneration by CMOs and public funding.

B. Risk identification and assessment measures

In light of potential systemic risks, the AI Act puts in place effective rules and oversight. Providers of general-purpose AI models with systemic risks should continuously assess and mitigate systemic risks.

The Code of Practice should be focused on specific risk assessment measures for general-purpose AI models with systemic risk. Following the risk taxonomy, **appropriate measures could be applied to assess different systemic risks, tailored to each specific type and nature of risk,** including their sources.

In addition to further risk assessment measures which will be detailed out in the Code of Practice, the AI Act requires providers to perform the necessary model evaluations, in particular prior to its first placing on the market, including conducting and documenting adversarial testing of the model, also, as appropriate, through internal or independent external testing.

The following concerns technical risk assessment measures, including model evaluation and adversarial testing. This is in line with the focus of the Code of Practice Working Group 2 “Risk identification and assessment measures for systemic risks”.

12. How can the effective implementation of **risk assessment measures reflect differences in size and capacity** between various providers such as SMEs and start-ups?

700 character(s) maximum

The implementation of risk assessment measures should not differentiate between SMEs, start-ups, and larger providers, as models developed by smaller entities can potentially create high-risk systems just as those by larger providers. This equity approach is consistent with other EU regulatory frameworks, such as product liability and pharmaceutical regulations, regardless of the size or profit margins of the enterprise. Similarly, in AI, the obligation to thoroughly document and assess risks must be upheld uniformly to ensure public safety and trust, without compromising on standards for the sake of business size or capacity.

13. In the **current state of the art**, which specific **risk assessment measures** should, in your view, general-purpose AI model providers take to effectively assess systemic risks along the entire model lifecycle, *in addition* to evaluation and testing?

Please **indicate to what extent you agree** that providers should take the risk assessment measures from the list. You can add additional measures and provide details on any of the measures, such as what is required for measures to be effective in practice.

Potential risk assessment measures	Strongly agree	Somewhat agree	Neither agree nor disagree	Disagree	I don't know
Determining risk thresholds and risk tolerance, incl. acceptable levels of risks and capabilities for model development and deployment, and respective quantification of risk severity and probability	○	◉	○	○	○
Forecasting model capabilities and risks before and during model development	◉	○	○	○	○
Continuous monitoring for emergence of risks, including data from users, relevant stakeholders, incident databases or similar	◉	○	○	○	○
Determining effectiveness of risk mitigation measures	○	◉	○	○	○
Safety cases to demonstrate that the model does not exceed maximum risk thresholds	○	◉	○	○	○
Aggregate risk assessment before model development	◉	○	○	○	○
Aggregate risk assessment before model deployment	○	◉	○	○	○
Aggregate risk assessment along the entire model lifecycle	○	◉	○	○	○
Third-party involvement in risk assessment, for example, related to inspections of training data, models or internal governance	◉	○	○	○	○

And/or:

- (●) **Other**

Please specify, including the extent you agree that providers should take the measures from the list

700 character(s) maximum

What is missing are the definitions of risks and their scope on economy, tax payments, social security, infringement in output, liability in cases of disinformation, in short: defining the parameters and the instruments of long-term monitoring of the impacts, and the parameters of measuring the impact in a holistic, not only technical way, on societies and CCI value chains.

If table is not submitted

- ( ) I don't know

Your comments

700 character(s) maximum

Third-party involvement in risk assessment is a MUST to ensure accountability of GPAI models incl. GenAI. Independent audits of used works and data, of models, and internal governance and compliance guidelines are essential, to safeguard the interests of authors and rightsholders. By involving experts from affected stakeholder groups and individuals, developers & providers can identify and mitigate systemic risks that may not be addressed through internal evaluation alone. This approach with the involvement of authors /rightsholders ensures that AI models respect IP rights, comply with EU standards, and uphold fair and trustworthy principles, ultimately fostering trust in AI deployment.

14. Please provide **links to relevant material** on state-of-the-art risk assessment measures, such as model cards, data sheets, templates or other publications.

(to our knowledge, no AI developer or provider is complying with all of these measures yet)
The status of the current measures by AI developers and its gaps, not meeting yet the requirements of the AI Act to find under https://rettighedsalliancen.com/wp-content/uploads/2024/09/Report-on-AI-model-providers-training-data-transparency-and-enforcement-of-copyrights.pdf - Report on AI model providers' training data transparency and enforcement of copyrights by the Danish Rights Alliance

15. In the **current state of the art**, which specific practices related to **model evaluations** should, in your view, general-purpose AI model providers take with a view to identifying and mitigating systemic risks?

Model evaluations can include various techniques, such as benchmarks and automated tests, red teaming and adversarial testing including stress testing and boundary testing, white-box evaluations with model explanation and interpretability techniques, and sociotechnical evaluations like field testing, user studies or uplift

studies.

Please **indicate to what extent you agree** that providers should implement the practice from the list. You can add additional practices and provide details on any of the practices. You can also indicate which model evaluation techniques listed above or which other techniques can reliably assess which specific systemic risks.

Potential evaluation practices	Strongly agree	Somewhat agree	Neither agree nor disagree	Disagree	I don't know
Performing evaluations at several checkpoints throughout the model lifecycle, in particular **during development and prior to internal deployment**	◉	○	○	○	○
Performing evaluations at several checkpoints throughout the model lifecycle, in particular **when the model risk profile changes** such as with access to tools or with different release strategies	◉	○	○	○	○
Ensuring evaluations inform model deployment in real-world conditions	○	○	○	◉	○
Ensuring evaluations provide insights into the degree to which a model introduces or exacerbates risks	○	◉	○	○	○
Using non-public model evaluations, as appropriate	○	○	○	◉	○
Involve independent external evaluators, including with appropriate levels of access to the model and related information	◉	○	○	○	○
Involve affected persons, to understand effects of human interactions with a particular model over time	◉	○	○	○	○
Documenting evaluation strategies and results	◉	○	○	○	○
Reporting evaluation strategies and results **publicly**, as appropriate	◉	○	○	○	○

Reporting evaluation strategies and results to **selected authorities and administrative bodies**, as appropriate, including sensitive evaluation results	○	◉	○	○	○
Continuously evaluate and improve evaluation strategies based on information from risk assessment and mitigation measures, including from incidents and near-misses	◉	○	○	○	○

And/or:

◉ **Other**

Please specify, including the extent you agree that providers should implement the practice from the list

700 character(s) maximum

The premise that CDSM 2019/790, Art. 4 also cover the use for developing 'GenAI', is legally and technically highly controversial and a general systemic risk. AI providers should only evaluate risks within controlled, sandbox (non-live) test environments and not in real world conditions. This secured practice allows for testing of models in a safe, isolated setting where potential further systemic risks can be identified and mitigated without impacting real-world users or value chains. This is essential for maintaining safety, reliability, and trust in AI systems, including GenAI, when using IP protected works, or data with the potential of violation of GDPR or personal rights.

It table is not submitted

○ I don't know

Your comments

700 character(s) maximum

We reiterate: Third-party involvement in risk assessment is a MUST to ensure accountability of GP AI models incl. GenAI esp. on IP rights and handling of licensing including the option to have a mediation in unclear cases. Independent audits of used works and data, of models, and internal governance and compliance guidelines are essential, to safeguard the interests of authors and rightsholders. By involving experts from affected stakeholder groups and individuals, developers & providers can identify and mitigate systemic risks that may not be addressed through internal evaluation alone but need to be met within EU law.

16. Please provide **links to relevant material** on state-of-the-art model evaluation practices, such as model cards, data sheets, templates or other publications.

--- (to our knowledge, no AI developer or provider is complying with all of these measures yet)

17. What are the **greatest challenges** that a general-purpose AI model provider could face in implementing risk assessment measures, including model evaluations?

700 character(s) maximum

To admit they had been infringing the copyright and moral rights incl. right to integrity of authors all the time, as it is very likely that Directive 2019/790, Arts. 3+4 on TDM do NOT cover the use for developing 'generative AI'. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4946214

C. Technical risk mitigation

Codes of Practice should also be focused on specific risk mitigation measures for general-purpose AI models with systemic risk. Following the risk taxonomy, **appropriate measures could be applied to mitigate different systemic risks, tailored to each specific type and nature of risk**, including their sources.

The following concerns technical risk mitigation measures, including cybersecurity protection for the general-purpose AI model and the physical infrastructure of the model. Measures can relate to model design, development or deployment.

This is in line with the focus of the Code of Practice Working Group 3 "Risk mitigation measures for systemic risks".

18. How can the effective implementation of **technical risk mitigation measures reflect differences in size and capacity** between various providers such as SMEs and start-ups?

700 character(s) maximum

The implementation of technical risk mitigation measures should not differentiate between SMEs, start-ups, and larger providers, as models developed by smaller entities can potentially create high-risk systems just as those by larger providers. This equity approach is consistent with other EU regulatory frameworks, regardless of the size or profit margins of the enterprise. Similarly, in AI, the obligation to thoroughly mitigate risks must be upheld uniformly to ensure safety and trust, esp. on technical aspects around the processing, storing, reproduction, making available and further exploitation of IP protected works and data.

19. In the **current state of the art**, which specific **technical risk mitigation measures** should, in your view, general-purpose AI model providers take to effectively mitigate systemic risks along the entire model lifecycle, *in addition* to cybersecurity protection?

Please **indicate to what extent you agree** that providers should take the measures from the list. You can add additional measures and provide details on any of the measures, such as what is required for measures to be effective in practice.

Potential technical risk assessment measures	Strongly agree	Somewhat agree	Neither agree nor disagree	Disagree	I don't know
Data governance such as data selection, cleaning, quality control	◉	○	○	○	○
Model design and development to achieve an appropriate level of trustworthiness characteristics such as model reliability, fairness or security	◉	○	○	○	○
Fine-tuning for trustworthiness and alignment such as through Reinforcement Learning from Human Feedback (RLHF) or Constitutional AI	○	○	◉	○	○
Unlearning techniques such as to remove specific harmful capabilities from models	◉	○	○	○	○
Technical deployment guardrails, such as content and other filters, capability restrictions, fine-tuning restrictions or monitoring-based restrictions in case of misuse by users	○	○	○	◉	○
Mitigation measures relating to the model architecture, components, access to tools or model autonomy	○	◉	○	○	○
Detection, labelling and other measures related to AI-generated or manipulated content	◉	○	○	○	○
Regular model updates, including security updates	○	◉	○	○	○
Measuring model performance on an ongoing basis	○	◉	○	○	○
Identification and mitigation of model misuse	○	◉	○	○	○
Access control to tools and levels of model autonomy	○	○	○	○	◉

And/or:

- (●) **Other**

Please specify, including the extent you agree that providers should take the measures from the list

700 character(s) maximum

Data governance and unlearning should be given highest priority. (Gen)AI developers + providers must establish governance frameworks that include a) the respect of IP rights for authors and rightsholders, as well as rights provided under the GDPR; b) install mechanisms to trace the scope of any usage, incl. but not limited to rights to: access to understand how their works are used, rectification to correct inaccuracies, to letting remove their works from the model, to restrict further processing, to object to the use of their works, and not to depend on decisions solely based on automated processing without human interaction within liability and compliance mechanisms.

If table is not submitted

- ( ) I don't know

Your comments

700 character(s) maximum

In addition to the need for data governance under IP rights and GDPR aspects, as well as unlearning techniques to be subject of the rights of authors and rightsholders, we see a particular need for a strict labelling process and to enable the tracking of misuse. Machine-generated products that simulate cultural works, media information or educational works / scientific results have a high-risk potential for remuneration fraud, violation of personal rights, disinformation, educational gaps, misuse of public funding, abuse of privileges for cultural assets including reduced VAT or the awarding of prizes and scholarships. Public funds must not be misused by non-labelled machine products.

20. Please provide **links to relevant material** on state-of-the-art technical risk mitigation practices, such as model cards, data sheets, templates or other publications.

21. What are the **greatest challenges** that a general-purpose AI provider could face in implementing technical risk mitigation measures?

700 character(s) maximum

They must have the potential to grow with the model right from the start. In addition to the assessment of technical risks and their mitigation, this also requires addressing issues of legal development, which, as can be seen from recent studies such as https://urheber.info/media/pages/diskurs/ki-training-ist-urheberrechtsverletzung/c943688809-1725462359/executive-summary_engl_final_29-08-2024.pdf , are still in the evolutionary phase. In addition, arbitration bodies must be established, such as, but not limited to, the question of who has the right, for example, to remove a work from the learning body or to gain insight into the scope of use.

D. Internal risk management and governance for general-purpose AI model providers

The following concerns policies and procedures to operationalise risk management in internal governance of general-purpose AI model providers, including keeping track of, documenting, and reporting serious incidents and possible corrective measures.

This is in line with the focus of the Code of Practice Working Group 4 “Internal risk management and governance for general-purpose AI model providers”.

22. How can the effective implementation of **internal risk management and governance measures reflect differences in size and capacity** between various providers such as SMEs and start-ups?

700 character(s) maximum

To comply with the provisions of the AI Act and related EU law, the implementation of risk assessment measures should not differentiate between SMEs, start-ups, and larger providers, as models developed by smaller entities can potentially develop high-risk systems just as those from larger providers. This approach is consistent with other regulatory frameworks, such as product liability and pharmaceutical regulations.

Links to relevant material

23. In the **current state of the art**, which specific **internal risk management and governance measures** should, in your view, general-purpose AI model providers take to effectively mitigate systemic risks along the entire model lifecycle, *in addition* to serious incident reporting?

Please **indicate to what extent you agree** that providers should take the measures from the list. You can add additional measures and provide details on any of the measures, such as what is required for measures to be effective in practice.

Potential internal risk management and governance measures	Strongly agree	Somewhat agree	Neither agree nor disagree	Disagree	I don’t know
Risk management framework across the model lifecycle	○	◉	○	○	○

Internal independent oversight functions in a transparent governance structure, such as related to risks and ethics	◉	○	○	○	○
Traceability in relation to datasets, processes, and decisions made during model development	◉	○	○	○	○
Ensuring that staff are familiar with their duties and the organisation's risk management practices	○	◉	○	○	○
Responsible scaling policies	○	◉	○	○	○
Acceptable use policies	○	◉	○	○	○
Whistleblower protections	○	◉	○	○	○
Internal resource allocation towards risk assessment and mitigation measures as well as research to mitigate systemic risks	○	◉	○	○	○
Robust security controls including physical security, cyber security and information security	○	◉	○	○	○
External accountability measures such as third-party audits, model or aggregated data access for researchers	◉	○	○	○	○
Other collaborations and involvements of a diverse set of stakeholders, including impacted communities	◉	○	○	○	○
Responsible release practices including staged release, particularly before open-sourcing a model with systemic risk	○	◉	○	○	○
Transparency reports such as model cards, system cards or data sheets	◉	○	○	○	○
Human oversight mechanisms	◉	○	○	○	○
Know-your-customer practices	○	○	○	○	◉
Logging, reporting and follow-up of near-misses along the lifecycle	○	◉	○	○	○
Measures to mitigate and remediate deployment issues and vulnerabilities	○	◉	○	○	○

Complaints handling and redress mechanisms, such as bug bounty programs	○	○	○	○	◉
Mandatory model updating policies and limit on maximum model availability	○	○	○	○	◉
Third-party and user discovery mechanisms and reporting related to deployment issues and vulnerabilities	◉	○	○	○	○

And/or:

◉ **Other**

Please specify, including the extent you agree that providers should take the measures from the list

700 character(s) maximum

Continuous internal monitoring of compliance with legal practice must be installed. The TDM exceptions cannot be relied upon, particularly in the stages of the process of developing so-called generative AI, as the copyright-relevant stages of development go beyond the TDM process, including but not limited to continued storage, repeated copying, memorisation, communication to the public, making available to the public, semantic plagiarism, etc. All of this requires licensing and transparency practices in accordance with Union law, which also requires technical monitoring.

If table is not submitted

○ I don't know

Your comments

700 character(s) maximum

To mitigate further systemic risks besides better not to rely only on the TDM exception(s), general-purpose AI model providers must prioritize independent oversight and external accountability measures. These are crucial, particularly in relation to copyright infringements. Internal risk management should not rely on self-regulation; it should include mechanisms similar to those required for social media networks under Art. 17 Directive 2019/790. This should involve mandatory external audits, transparent reporting processes, and the implementation of effective content identification and removal tools to ensure that authors and rightsholders rights are enforceable.

24. Please provide **links to relevant material** on state-of-the-art governance risk mitigation practices, such as model cards, data sheets, templates or other publications.

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4946214

25. What are the **greatest challenges** that a general-purpose AI provider could face in implementing governance risk mitigation measures?

700 character(s) maximum

To confirm they had been infringing the copyright and moral rights incl. right to integrity of authors, as it is very likely that: a) Directive 2019/790, Art. 4 on TDM does NOT cover the use for developing ‘generative AI’; b) all utilisation also before 7.6.2021 are not covered by any law and fall under the Berne Convention. This may result in a liquidation of I. existing models; II. existing foundation models and data sets; III. legal procedures against those developers, providers and operators. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4946214

Section 3. Reviewing and monitoring of the General-Purpose AI Code of Practice

The process of drawing-up the first Code of Practice will start immediately after the AI Act enters into force and will last for 9 months, in view of enabling providers of general-purpose AI models to demonstrate compliance on time. The AI Office shall aim to ensure that the Code of Practice clearly sets out their specific objectives and contains commitments or measures, including key performance indicators as appropriate, to ensure the achievement of those objectives.

The AI Office shall aim to ensure that participants to the Code of Practice report regularly to the AI Office on the implementation of the commitments and the measures taken and their outcomes, including as measured against the key performance indicators as appropriate. Key performance indicators and reporting commitments shall reflect differences in size and capacity between various participants. The AI Office and the Board shall regularly monitor and evaluate the achievement of the objectives of the Code of Practice by the participants and their contribution to the proper application of this Regulation.

The AI Office shall, as appropriate, encourage and facilitate the review and adaptation of the Code of Practice.

26. What are examples of **key performance indicators** which are, in your view, effective to measure the compliance of participants with the objectives and measures which will be established by the Code of Practice?

700 character(s) maximum

- Number and rate of IP, GDPR, personal Rights Infringement Reduction: incidents and percentage decrease in infringement incidents reported

- Speed of resolution: average time taken to resolve and remove infringing works and data after reporting
- Transparency in used or planned to be used works, data and other content identification: availability and accuracy of reporting mechanisms and transparency reports
- Feedback and Dispute Resolution Efficiency: rightsholder satisfaction with feedback and dispute resolution processes incl. mediation bodies
- Continued collaboration between AI Office and Authors', Artists', Performers' Federations in relevant fora and targeted consultations

Links to relevant material

27. How can **key performance indicators and reporting commitments** for providers **reflect differences in size and capacity** between various providers such as SMEs and start-ups?

700 character(s) maximum

The effective implementation of key performance indicators should not differentiate between SMEs, start-ups, and larger providers, as models developed by smaller entities can potentially create high-risk systems just as those from larger providers. In GPAI and esp. generative informatics, the obligation to document and assess key performer indicators must be upheld uniformly to ensure public safety and trust, without compromising on standards for the sake of business size.

Links to relevant material

28. Which aspects should inform the timing of **review and adaptation of the content of the Code of Practice** for general-purpose AI models in order to ensure that the **state of the art** is reflected? This does not necessarily imply a complete review, but can only involve pertinent parts.

Please rank all relevant aspects from the following list from 1 to 4 (1 being the most important). You can add additional aspects and provide details on any of the aspects or other considerations under "Specify".

	Rank 1	Rank 2	Rank 3	Rank 4
Pre-planned intervals to assess the need for revision: Assessments of whether the content of the Code of Practice for general-purpose AI models needs to be revised or adapted should be pre-planned for specific time intervals.	○	○	○	◉
Alerts by independent experts or other stakeholders: Alerts by selected independent experts, such as by the Scientific Panel which will be set up in the AI Act governance structure, or				

by other stakeholders such as downstream providers, academia or civil society should inform a revision of the content of the Code of Practice.	○	◉	○	○
Monitoring and foresight: Independent monitoring and foresight related to the AI ecosystem, technological and market developments, emergence of systemic risks and any other relevant trends, such as related to sources of risks like model autonomy, should inform a revision of the content of the Code of Practice	○	○	◉	○
Other	◉	○	○	○

Specify for "**Other**"

The CoP shall prior to implementation clarify the legal + technical scope of the exceptions for TDM (Art. 3+4, 2019/790 CDSM Directive), as it is to expect they are NOT applicable to cover any developing of (Gen)AI. The statutory language and text of the provision, its conception, the ratio of the exception indicate that it must not be applied to the training of GenAI models. Hence, the training of GenAI models without the authors and rightsholders opt-in can be classified as both a copyright infringement and a violation of duties in the AI Act. We call on the AI Office to get the scope of the TDM Art 4 exception formally clarified before the Code of Practice is drafted.

If ranking is not submitted

○ **I don't know**

Your comments

700 character(s) maximum

Affected stakeholders, in the case of GenAi primarily authors, performers and artists, must be consistently involved. Monitoring should include: a. Evaluations of the damage done before 2021 and since (economic, environmental, social & labour policy); b. Evaluation of the promised profits through AI; c. Legal policy responsiveness; d. Increase/reduction of litigation risks through the application and enforceability of IP rules; e. Predetermined breaking points e.g. unresolved contractual issues and the right to licence and remuneration; f. Future feasibility of a data directory (NOT A registry!) beyond TDM opt-outs, but for the application of opt-in; g. Defining tasks of arbitration bodies.

Links to relevant material

https://europeanwriterscouncil.eu/240425_cwos_jointstatement_ai-act/
https://urheber.info/media/pages/diskurs/ki-training-ist-urheberrechtsverletzung/c943688809-1725462359/executive-summary_engl_final_29-08-2024.pdf

Option to upload a document for additional information

You have the option to upload one document to share further information with the AI Office. Please download the template that is structured along the topics covered by the Code of Practice Working Groups. Based on the submissions and answers to the targeted questions, a first draft of the Code of Practice will be developed.

Please upload your document in a doc or docx format, instead of pdf or similar.

Template_for_free-text_submissions.docx

Please upload your file(s)

Only files of the type doc,docx are allowed

4a24a066-0746-44c3-ab76-020ddfb5e194/EWC_Ceatl_JOINT_SUBMISSION_FINAL_240916.docx

Thank you

Thank you for participating in the consultation. Please don't forget to click on submit.

The AI Office will publish a summary of the results of the consultation. Results will be based on aggregated data and respondents will not be directly quoted.

All contributions to this consultation may be made publicly available.

Contact

Contact Form