

# General-Purpose AI Code of Practice: Provide your feedback to the first draft

Fields marked with \* are mandatory.

## Provide your feedback to the first General-Purpose AI Code of Practice!

Thank you for participating in the drawing-up of the first General-Purpose AI Code of Practice.

Upon receiving the first draft, you are encouraged to express your comments on the content via this survey, **deadline Thursday 28 November 2024, 12:00 CET.**

Your feedback is essential in helping us understand how the Code of Practice can best serve and support stakeholders across diverse sectors, leading to a final Code of Practice which should reflect the different submissions as far as possible, while ensuring a convincing implementation of the legal framework. Please be aware that the survey does not cover Art. 53(1)(d) issues.

For each section/measure/sub-measure of the Code of Practice, participants will be asked to answer **two types of questions**:

1. **Opinion rating** (close-ended feedback): express the level of agreement with the content choosing among different options.
2. **Open-ended questions**: specific to each sub-section's measures and sub-measures, and additional questions cross-measures. This includes the opportunity to comment on each section and the overall draft.

In addition, you may upload supporting documents at the end of the survey.

**Each stakeholder can only submit one answer to this survey.** For organisations, please note, the Point of Contact is responsible for collecting the views of their organisations chosen working group representatives.

**You have the option to respond to all questions if you wish; however, you do not have to.** The survey is long to enable participants to provide detailed feedback on every aspect of the Code of Practice that interests them. We trust each respondent will provide valuable input into the sections that are most relevant to their area of expertise, which may include skipping sub-measures, measures, or sections not of interest to them. For the ease of engaging with the survey, please have the Code of Practice open for reference.

Please note, **it will not be possible to forward this survey.** Your responses will be kept confidential in

accordance with the Vademecum confidentiality agreements.

The AI Office is looking forward to this inclusive and transparent process!

**DISCLAIMER:** All content uploaded in the "Supporting Documents" section should also be included in the main free-text sections of the survey to ensure that the relevant Chairs engage with the content.

## Stakeholder information

---

Please provide your name, surname, email address, and the name of your organisation (if applicable). Please note that if your contact information does not correspond to an eligible participant or to the organisation's Point of Contact, your response will be discarded.

	Name	Surname	Email address	Organisation (if applicable)
Stakeholder	Teresa	Nobre	teresa@communia-association.org	COMMUNIA Association for the Public Domain

\* Which stakeholder category would you (or your organisation) consider yourself in?

- Academia (in a personal capacity)
- Civil society organisation
- Downstream provider of an AI system based on general-purpose AI models, or acting on behalf of such providers
- EU Member State representative
- European or international observer
- Other independent expert (in a personal capacity)
- Other industry organisation, or acting on behalf of such organisations
- Other organisation with relevant expertise
- Other stakeholder organisation
- Provider of a general-purpose AI model, or acting on behalf of such providers
- Rightsholder organisation

\* Please indicate all the working groups you participate in. Please note that if you are the Point of Contact of your organisation, you should select all the working groups of your representatives.

- Working Group 1: Transparency & copyright-related rules
- Working Group 2: Risk identification and assessment for systemic risk
- Working Group 3: Technical risk mitigation for systemic risk
- Working Group 4: Governance risk mitigation for systemic risk

\* Please indicate which section you wish to provide your feedback. If you wish to comment on all sections, please select all the options.

- Overall Code of Practice Draft
- Section II: [Working Group 1] Rules for providers of general-purpose AI models
- Section III: [Working Group 2] Taxonomy of systemic risks
- Section IV: [Working Groups 2/3/4] Rules for providers of general-purpose AI models with systemic risk

## Section II: [Working Group 1] Rules for Providers of General-Purpose AI Models

---

### ***Measures/Sub-measures Specific Feedback on Section II: [Working Group 1] Transparency***

In this section you are asked to provide your overall opinion on the measures and sub-measures included in the second section of the Code of Practice related to *Transparency*.

#### ***Transparency, Measure 1: Documentation for the AI Office***

To what extent do you agree with this measure? (Measure 1)

- The measure should be removed in its entirety
- The measure should be substantially edited and/or further clarified
- The measure should be lightly edited and/or further clarified
- The measure is close to where it needs to be

Please explain your rating to this measure (Measure 1)

*2000 character(s) maximum*

Please provide suggestions on how to improve this measure (Measure 1)

*2000 character(s) maximum*

What KPI would you add for this measure? (Measure 1)

*2000 character(s) maximum*

**Transparency, Measure 2: Documentation for downstream providers**

To what extent do you agree with this measure? (Measure 2)

- The measure should be removed in its entirety
- The measure should be substantially edited and/or further clarified
- The measure should be lightly edited and/or further clarified
- The measure is close to where it needs to be

Please explain your rating to this measure (Measure 2)

*2000 character(s) maximum*

Please provide suggestions on how to improve this measure (Measure 2)

*2000 character(s) maximum*

What KPI would you add for this measure? (Measure 2)

*2000 character(s) maximum*

For the items listed in the table (at page 10), how should the Code of Practice provide greater detail? (Measure 2)

*2000 character(s) maximum*

**Additional Feedback on Section II: [Working Group 1] Transparency**

Please provide additional feedback on the content of the Code of Practice second section related to *Transparency* by answering the following questions.

What additional measure and/or sub-measures would you include in this sub-section? (Section II - Transparency)

*2000 character(s) maximum*

For which of the topics below is more clarification or specificity most needed? (*select all that apply*) (Section II - Transparency)

- General information
- Intended uses
- Acceptable use policies
- Methods of distribution
- Interaction with hardware and software
- Software versions
- Model architecture and parameters
- Input and output modalities
- License
- Technical means for downstream integration
- Training process
- Training, testing, validation data
- Computational resources
- Energy consumption
- Testing process

If you are a General-Purpose AI Model provider, which of the topics below related to the transparency obligations would be the costliest to comply with for your organization? (Section II - Transparency)

- General information
- Intended uses
- Acceptable use policies
- Methods of distribution
- Interaction with hardware and software
- Software versions
- Model architecture and parameters
- Input and output modalities
- License
- Technical means for downstream integration
- Training process
- Training, testing, validation data
- Computational resources
- Energy consumption
- Testing process

If you are a General-Purpose AI Model provider, for which of the topics below do you currently make information publicly available? (*select all that apply*) (Section II - Transparency)

- Not applicable
- General information
- Intended uses
- Acceptable use policies
- Methods of distribution
- Interaction with hardware and software
- Software versions
- Model architecture and parameters
- Input and output modalities
- License
- Technical means for downstream integration
- Training process
- Training, testing, validation data
- Computational resources
- Energy consumption
- Testing process

If you are a General-Purpose AI Model provider, for which of the topics below do you internally produce information for business purposes? (*select all that apply*) (Section II - Transparency)

- Not applicable
- General information
- Intended uses
- Acceptable use policies
- Methods of distribution
- Interaction with hardware and software
- Software versions
- Model architecture and parameters
- Input and output modalities
- License
- Technical means for downstream integration
- Training process
- Training, testing, validation data
- Computational resources
- Energy consumption
- Testing process

If you are a General-Purpose AI Model provider, how do you currently share information to downstream providers but not the public? (*select all that apply*) (Section II - Transparency)

- Not applicable
- We only share information publicly
- We provide contact information on our website
- We coordinate with our external distribution channels that distribute the model (GitHub, Hugging Face, Amazon Bedrock, Microsoft Azure, Google Vertex AI, Together)
- Other

If you are *not* a General-Purpose AI Model provider, for which of the topics below would you prefer information be encouraged to be made public? (*select all that apply*) (Section II - Transparency)

- Not applicable
- General information
- Intended uses
- Acceptable use policies
- Methods of distribution
- Interaction with hardware and software
- Software versions
- Model architecture and parameters
- Input and output modalities
- License
- Technical means for downstream integration
- Training process
- Training, testing, validation data
- Computational resources
- Energy consumption
- Testing process

### ***Measures/Sub-measures Specific Feedback on Section II: [Working Group 1] Copyright-related rules***

In this section you are asked to provide your overall opinion on the measures and sub-measures included in the second section of the Code of Practice related to *Copyright-related rules*.

#### ***Copyright-related rules, Measure 3: Put in place copyright policy***

To what extent do you agree with this measure? (Measure 3)

- The measure should be removed in its entirety
- The measure should be substantially edited and/or further clarified
- The measure should be lightly edited and/or further clarified
- The measure is close to where it needs to be

Please explain your rating to this measure (Measure 3)

*2000 character(s) maximum*

Please provide suggestions on how to improve this measure (Measure 3)

*2000 character(s) maximum*

What KPI would you add for this measure? (Measure 3)

*2000 character(s) maximum*



---

### **Copyright-related rules, Sub-Measure 3.1: Draw up and implement a copyright policy**

To what extent do you agree with this sub-measure? (Sub-Measure 3.1)

- The sub-measure should be removed in its entirety
- The sub-measure should be substantially edited and/or further clarified
- The sub-measure should be lightly edited and/or further clarified
- The sub-measure is close to where it needs to be

Please explain your rating to this sub-measure (Sub-Measure 3.1)

*2000 character(s) maximum*

Please provide suggestions on how to improve this sub-measure (Sub-Measure 3.1)

*2000 character(s) maximum*

What KPI would you add for this sub-measure? (Sub-Measure 3.1)

*2000 character(s) maximum*

### **Copyright-related rules, Sub-Measure 3.2: Upstream copyright compliance**

To what extent do you agree with this sub-measure? (Sub-Measure 3.2)

- The sub-measure should be removed in its entirety
- The sub-measure should be substantially edited and/or further clarified
- The sub-measure should be lightly edited and/or further clarified
- The sub-measure is close to where it needs to be

Please explain your rating to this sub-measure (Sub-Measure 3.2)

*2000 character(s) maximum*

This sub-measure does not appear to take into account that model providers will not be able to pass onto the upstream providers of open data sets any copyright compliance obligations.

Please provide suggestions on how to improve this sub-measure (Sub-Measure 3.2)

*2000 character(s) maximum*

It should be clarified that the sub-measure does not apply to data sets made available under non-exclusive free licenses for the benefit of any users, unless the signatories enter into an individually negotiated contract with the third party providers about the use of those data sets for the development of a general-purpose AI model. In the absence of such a contract, the signatories remain responsible for identifying and complying with, including through state-of-the-art technologies, rights reservations expressed pursuant to Article 4(3) of Directive (EU) 2019/790.

What KPI would you add for this sub-measure? (Sub-Measure 3.2)

*2000 character(s) maximum*

### **Copyright-related rules, Sub-Measure 3.3: Downstream copyright compliance**

To what extent do you agree with this sub-measure? (Sub-Measure 3.3)

- The sub-measure should be removed in its entirety
- The sub-measure should be substantially edited and/or further clarified
- The sub-measure should be lightly edited and/or further clarified
- The sub-measure is close to where it needs to be

Please explain your rating to this sub-measure (Sub-Measure 3.3)

*2000 character(s) maximum*

This sub-measure should be removed for the following reasons:

1) Its scope goes beyond the scope of protection of Article 53(1)(c) of the AI Act. Only model providers are required to put in place a copyright compliance policy; downstream systems and applications are not targeted by this provision.

2) It assumes that a similar output is necessarily an infringing output, which is incorrect. A similar output can only be qualified as an infringing output if no copyright exception or limitation applies, or if it does not qualify as an independent similar creation.

3) It raises fundamental rights concerns. System-level measures to prevent output similarity cannot be proposed in the context of a code of practice, without legislative intervention, due to the risks to fundamental rights that they entail. Contrary to model-level measures, system-level measures (e.g. input filters and output filters) are triggered by an interaction with an end-user. Users rights considerations come into play when output similarity is the result of an intentional act of "extraction" (e.g. specific instructions) by the end-user to cause a downstream system to generate outputs similar to copyrighted works. Depending on the purpose of the use, this act of "extraction" of similar expressions in the output can be considered a legitimate use of a copyrighted work. For example, the use can be covered by the quotation right, by the exception for caricature, parody and pastiche, by the incidental inclusion exception, or even by the education or research exceptions if the output is aimed at serving an education or research purpose. Any measure to prevent output similarity has therefore the potential to affect legitimate uses.

4) It assumes that model providers are able to police downstream copyright compliance. Open source model providers cannot pass onto downstream users copyright compliance obligations.

Please provide suggestions on how to improve this sub-measure (Sub-Measure 3.3)

*2000 character(s) maximum*

No downstream copyright compliance obligations should be included in this code of practice.

What KPI would you add for this sub-measure? (Sub-Measure 3.3)

*2000 character(s) maximum*

### **Copyright-related rules, Measure 4: Compliance with the limits of the TDM exception**

To what extent do you agree with this measure? (Measure 4)

- The measure should be removed in its entirety
- The measure should be substantially edited and/or further clarified
- The measure should be lightly edited and/or further clarified
- The measure is close to where it needs to be

Please explain your rating to this measure (Measure 4)

*2000 character(s) maximum*

Please provide suggestions on how to improve this measure (Measure 4)

*2000 character(s) maximum*

What KPI would you add for this measure? (Measure 4)

*2000 character(s) maximum*

### **Copyright-related rules, Sub-Measure 4.1: Respect Robots.txt**

To what extent do you agree with this sub-measure? (Sub-Measure 4.1)

- The sub-measure should be removed in its entirety
- The sub-measure should be substantially edited and/or further clarified
- The sub-measure should be lightly edited and/or further clarified
- The sub-measure is close to where it needs to be

Please explain your rating to this sub-measure (Sub-Measure 4.1)

*2000 character(s) maximum*

A sub-measure that targets one specific implementation of the concept of machine readable rights reservations is insufficient, especially since robots.txt has a number of conceptual shortcomings that make it unsuitable as an expression of rights reservations, including the following:

- 1) Robots.txt policies can only be set by entities that control websites/online publishing platforms. In many cases, these entities will not be the rightholders themselves.
- 2) Robots.txt has limited usefulness for types of content (such as music or AV content) that are not predominantly distributed via the open internet.
- 3) Robot.txt is unable to deal with embedded media files.

4) Last but not the least, Robot.txt does not allow opting out of TDM or specific applications of TDM (e.g. AI training). Currently, robots.txt does not allow a web publisher to indicate a horizontal opt-out that applies to all crawlers crawling for a specific type of use. Instead, web publishers are required to target each individual crawler individually, often lacking information about the crawlers in use and the purposes of individual crawlers (this can only be partially addressed by the transparency obligations under measure 5).

By giving robots.txt with all the above shortcomings an elevated position the CoP risks further entrenching a solution with known limitations.

Please provide suggestions on how to improve this sub-measure (Sub-Measure 4.1)

*2000 character(s) maximum*

Any sub-measure aimed to ensure compliance with the limits of the TDM exception must require a respect for multiple forms of machine readable rights reservations. Ideally, the sub-measure should require signatories to only use data ingestion mechanisms that read and follow machine readable opt-outs from either TDM or opt-outs specifically aimed at AI training. In any case, sub-measures 4.1 and 4.3 should be combined, applying the same standard of commitment to all appropriate machine-readable means to express a rights reservation.

What KPI would you add for this sub-measure? (Sub-Measure 4.1)

*2000 character(s) maximum*

### **Copyright-related rules, Sub-Measure 4.2: No effect on findability**

To what extent do you agree with this sub-measure? (Sub-Measure 4.2)

- The sub-measure should be removed in its entirety
- The sub-measure should be substantially edited and/or further clarified
- The sub-measure should be lightly edited and/or further clarified
- The sub-measure is close to where it needs to be

Please explain your rating to this sub-measure (Sub-Measure 4.2)

*2000 character(s) maximum*

This sub-measure does not sufficiently take into account that crawlers can have multiple purposes and that the boundaries between search and AI are not very well defined.

Please provide suggestions on how to improve this sub-measure (Sub-Measure 4.2)

*2000 character(s) maximum*

See our comments to sub-measure 4.1. above. Requiring signatories to only use data collection mechanisms that read and follow granular machine-readable opt-outs (in this case allowing opting out of AI training without opting out of indexing for search) would address the underlying concern in a more structured way.

What KPI would you add for this sub-measure? (Sub-Measure 4.2)

*2000 character(s) maximum*

**Copyright-related rules, Sub-Measure 4.3: Best efforts regarding other appropriate means**

To what extent do you agree with this sub-measure? (Sub-Measure 4.3)

- The sub-measure should be removed in its entirety
- The sub-measure should be substantially edited and/or further clarified
- The sub-measure should be lightly edited and/or further clarified
- The sub-measure is close to where it needs to be

Please explain your rating to this sub-measure (Sub-Measure 4.3)

*2000 character(s) maximum*

See our comments to sub-measure 4.1 above. It is also unclear what it means “expressions of rights reservations at aggregate level”.

Please provide suggestions on how to improve this sub-measure (Sub-Measure 4.3)

*2000 character(s) maximum*

What KPI would you add for this sub-measure? (Sub-Measure 4.3)

*2000 character(s) maximum*

**Copyright-related rules, Sub-Measure 4.4: Commitment to collaborative development of rights reservations’ standards**

To what extent do you agree with this sub-measure? (Sub-Measure 4.4)

- The sub-measure should be removed in its entirety
- The sub-measure should be substantially edited and/or further clarified
- The sub-measure should be lightly edited and/or further clarified
- The sub-measure is close to where it needs to be

Please explain your rating to this sub-measure (Sub-Measure 4.4)

*2000 character(s) maximum*

Please provide suggestions on how to improve this sub-measure (Sub-Measure 4.4)

*2000 character(s) maximum*

What KPI would you add for this sub-measure? (Sub-Measure 4.4)

2000 character(s) maximum

**Copyright-related rules, Sub-Measure 4.5: No crawling of piracy websites**

To what extent do you agree with this sub-measure? (Sub-Measure 4.5)

- The sub-measure should be removed in its entirety
- The sub-measure should be substantially edited and/or further clarified
- The sub-measure should be lightly edited and/or further clarified
- The sub-measure is close to where it needs to be

Please explain your rating to this sub-measure (Sub-Measure 4.5)

2000 character(s) maximum

The sub-measure encourages signatories to exclude websites listed in the Commission Counterfeit and Piracy Watch List (the “Watch List”) and in analogous exclusion lists published by relevant public authorities in the jurisdictions where they are established. However, the Watch List is just a Commission Staff Working Document – it is not an official statement of the College and does not have any legal effect.

The Watch List is a selection of physical marketplaces and online service providers reported by stakeholders. The Commission’ services only reflect the results of stakeholder consultations. The listed online service providers reportedly engage in, facilitate or benefit from counterfeiting or piracy, but there is no procedure in place to confirm the alleged counterfeit or piracy activity. While the Commission’ services try to ensure that the information contained in the Watch List is accurate, they do not necessarily rely on court decisions or other authoritative documents. The Commission does not take any position on the content of the stakeholders’ allegations and the List does not contain findings of legal violations.

In sum, the listed websites cannot be qualified as piracy websites just because they made it to the Watch List. Therefore, encouraging the signatories to exclude them from their crawling activities has no legal basis.

The aim of the Watch List is to encourage the operators to take the necessary actions and measures to reduce the availability of IPR infringing goods or services on their websites. Repurposing it in the context of this code of practice is strongly inadvisable.

Please provide suggestions on how to improve this sub-measure (Sub-Measure 4.5)

2000 character(s) maximum

The reference to the Commission Counterfeit and Piracy Watch List and other analogous exclusion lists should be deleted.

What KPI would you add for this sub-measure? (Sub-Measure 4.5)

2000 character(s) maximum

### **Copyright-related rules, Measure 5: Transparency**

To what extent do you agree with this measure? (Measure 5)

- The measure should be removed in its entirety
- The measure should be substantially edited and/or further clarified
- The measure should be lightly edited and/or further clarified
- The measure is close to where it needs to be

Please explain your rating to this measure (Measure 5)

*2000 character(s) maximum*

This section currently lacks any connection to the training data transparency obligation in Article 53(1)(d) of the Act and the training data transparency template to be provided by the AI Office.

Please provide suggestions on how to improve this measure (Measure 5)

*2000 character(s) maximum*

What KPI would you add for this measure? (Measure 5)

*2000 character(s) maximum*

### **Copyright-related rules, Sub-Measure 5.1: Public information about rights reservation compliance**

To what extent do you agree with this sub-measure? (Sub-Measure 5.1)

- The sub-measure should be removed in its entirety
- The sub-measure should be substantially edited and/or further clarified
- The sub-measure should be lightly edited and/or further clarified
- The sub-measure is close to where it needs to be

Please explain your rating to this sub-measure (Sub-Measure 5.1)

*2000 character(s) maximum*

Please provide suggestions on how to improve this sub-measure (Sub-Measure 5.1)

*2000 character(s) maximum*

What KPI would you add for this sub-measure? (Sub-Measure 5.1)

*2000 character(s) maximum*

**Copyright-related rules, Sub-Measure 5.2: Crawler name and robots.txt features**

To what extent do you agree with this sub-measure? (Sub-Measure 5.2)

- The sub-measure should be removed in its entirety
- The sub-measure should be substantially edited and/or further clarified
- The sub-measure should be lightly edited and/or further clarified
- The sub-measure is close to where it needs to be

Please explain your rating to this sub-measure (Sub-Measure 5.2)

*2000 character(s) maximum*

The wording "and their relevant robots.txt features, including at the time of crawling" is not clear.

Please provide suggestions on how to improve this sub-measure (Sub-Measure 5.2)

*2000 character(s) maximum*

What KPI would you add for this sub-measure? (Sub-Measure 5.2)

*2000 character(s) maximum*

**Copyright-related rules, Sub-Measure 5.3: Single point of contact and complaint handling**

To what extent do you agree with this sub-measure? (Sub-Measure 5.3)

- The sub-measure should be removed in its entirety
- The sub-measure should be substantially edited and/or further clarified
- The sub-measure should be lightly edited and/or further clarified
- The sub-measure is close to where it needs to be

Please explain your rating to this sub-measure (Sub-Measure 5.3)

*2000 character(s) maximum*

Please provide suggestions on how to improve this sub-measure (Sub-Measure 5.3)

*2000 character(s) maximum*

What KPI would you add for this sub-measure? (Sub-Measure 5.3)

*2000 character(s) maximum*



## **Copyright-related rules, Sub-Measure 5.4: Documentation of data sources and authorisations**

To what extent do you agree with this sub-measure? (Sub-Measure 5.4)

- The sub-measure should be removed in its entirety
- The sub-measure should be substantially edited and/or further clarified
- The sub-measure should be lightly edited and/or further clarified
- The sub-measure is close to where it needs to be

Please explain your rating to this sub-measure (Sub-Measure 5.4)

*2000 character(s) maximum*

In the absence of the training data transparency template, this measure cannot be assessed.

Please provide suggestions on how to improve this sub-measure (Sub-Measure 5.4)

*2000 character(s) maximum*

What KPI would you add for this sub-measure? (Sub-Measure 5.4)

*2000 character(s) maximum*

## **Additional Feedback on Section II: [Working Group 1] Rules for providers of General-Purpose AI Models**

Please provide additional feedback on the content of the Code of Practice second section related to *Copyright-related rules* by answering the following questions.

What additional measure and/or sub-measures would you include in this sub-section? (Section II - Copyright)

*2000 character(s) maximum*

Section II lacks a restatement of the public transparency obligation arising from Art 53(1)d, similar to the restatement of Art 53(1)c in measure 3. However, it should be kept in mind that Art 53(1)d goes beyond the legitimate interests of copyright holders.

Section II also needs to define the effects that must be given by AI developers to opt-outs received from rights holders. It is clear that the responsibility is not limited to simply removing (or not including) a particular work, but that AI developers must also use their best efforts to ensure that all instances of the opted-out work are removed from the data used to train models. Conversely, it is also clear that the responsibility cannot include measures that are not technologically feasible (such as removing opted-out works from already trained models). In this context, it will be important for AI developers to clearly communicate deadlines for training models. Such considerations must be an essential part of the code of practice for AI model developers that is being developed.

Are there measures and/or sub-measures that are not technically feasible? (Section II - Copyright)

*2000 character(s) maximum*

## Supporting Documents

---

Please provide supporting documents, excluding the listing of links to additional websites. If you provide links, please be aware that these will be disregarded. Supporting documents should have a maximum length of n. 5 pages.

### Contact

[aiofficesupport@intelleraconsulting.com](mailto:aiofficesupport@intelleraconsulting.com)