Code of Practice for General-Purpose AI Models

Transparency Chapter

Nuria Oliver
Working Group 1 Co-Chair

Rishi Bommasani
Working Group 1 Vice-Chair

Introductory note by the Chair and Vice-Chair of the Transparency Chapter.

The Transparency Chapter of the Code of Practice describes three Measures which Signatories commit to implementing to comply with their transparency obligations under Article 53(1), points (a) and (b), and the corresponding Annexes XI and XII of the AI Act.

To facilitate compliance and fulfilment of the commitments contained in Measure 1.1, we include a user-friendly Model Documentation Form which allows Signatories to easily compile the information required by the aforementioned provisions of the AI Act in a single place.

The Model Documentation Form indicates for each item whether the information is intended for downstream providers, the AI Office or national competent authorities. Information intended for the AI Office or national competent authorities is only to be made available following a request from the AI Office, either *ex officio* or based on a request to the AI Office from national competent authorities. Such requests will state the legal basis and purpose of the request and will concern only items from the Form that are strictly necessary for the AI Office to fulfil its tasks under the AI Act at the time of the request, or for national competent authorities to exercise their supervisory tasks under the AI Act at the time of the request, in particular to assess compliance of providers high-risk AI systems built on general-purpose AI models where the provider of the system is different from the provider of the model.

In accordance with Article 78 AI Act, the recipients of any of the information contained in the Model Documentation Form are obliged to respect the confidentiality of the information obtained, in particular intellectual property rights and confidential business information or trade secrets, and to put in place adequate and effective cybersecurity measures to protect the security and confidentiality of the information obtained.

Objectives

The overarching objective of this Code of Practice ("Code") is to improve the functioning of the internal market, to promote the uptake of human-centric and trustworthy artificial intelligence ("AI"), while ensuring a high level of protection of health, safety, and fundamental rights enshrined in the Charter, including democracy, the rule of law, and environmental protection, against harmful effects of AI in the Union, and to support innovation pursuant to Article 1(1) AI Act.

To achieve this overarching objective, the specific objectives of this Code are:

A. To serve as a guiding document for demonstrating compliance with the obligations provided for in Articles 53 and 55 AI Act, while recognising that adherence to the Code does not constitute conclusive evidence of compliance with these obligations under the AI Act.

B. To ensure providers of general-purpose AI models comply with their obligations under the AI Act and to enable the AI Office to assess compliance of providers of general-purpose AI models who choose to rely on the Code to demonstrate compliance with their obligations under the AI Act.

Recitals

Whereas:

(a) The Signatories recognise the particular role and responsibility of providers of general-purpose AI models along the AI value chain, as the models they provide may form the basis for a range of downstream AI systems, often provided by downstream providers that need 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 (see recital 101 AI Act).

(b) The Signatories recognise that in the case of a fine-tuning or other modification of a general-purpose AI model, where the natural or legal person, public authority, agency or other body that modifies the model becomes the provider of the modified model subject to the obligations for providers of general purpose AI models, their Commitments under the Transparency Chapter of the Code should be limited to that modification or fine-tuning, to comply with the principle of proportionality (see recital 109 AI Act). In this context, Signatories should take into account relevant guidelines by the European Commission.

(c) The Signatories recognise that, without exceeding the Commitments under the Transparency Chapter of this Code, when providing information to the AI Office or to downstream providers they may need to take into account market and technological developments, so that the information continues to serve its purpose of allowing the AI Office and national competent authorities to fulfil their tasks under the AI Act, and downstream providers to integrate the Signatories' models into AI systems and to comply with their obligations under the AI Act (see Article 56(2), point (a), AI Act).

This Chapter of the Code focuses on the documentation obligations from Article 53(1), points (a) and (b), AI Act that are applicable to all providers of general-purpose AI models (without prejudice to the exception laid down in Article 53(2) AI Act), namely those concerning Annex XI, Section 1, and Annex XII AI Act. The documentation obligations concerning Annex XI, Section 2, AI Act, applicable only to providers of general-purpose AI models with systemic risk are covered by Measure 10.1 of the Safety and Security Chapter of this Code.

Commitment 1 Documentation

LEGAL TEXT: Articles 53(1)(a), 53(1)(b), 53(2), 53(7), and Annexes XI and XII AI Act

In order to fulfil the obligations in Article 53(1), points (a) and (b), AI Act, Signatories commit to drawing up and keeping up-to-date model documentation in accordance with Measure 1.1, providing relevant information to providers of AI systems who intend to integrate the general-purpose AI model into their AI systems ('downstream providers' hereafter), and to the AI Office upon request (possibly on behalf of national competent authorities upon request to the AI Office when this is strictly necessary for the exercise of their supervisory tasks under the AI Act, in particular to assess the compliance of a high-risk AI system built on a general-purpose AI model where the provider of the system is different from the provider of the model[1]), in accordance with Measure 1.2, and ensuring quality, security, and integrity of the documented information in accordance with Measure 1.3. In accordance with Article 53(2) AI Act, these Measures do not apply to providers of general-purpose AI models released under a free and open-source license that satisfy the conditions specified in that provision, unless the model is a general-purpose AI model with systemic risk.

Measure 1.1 Drawing up and keeping up-to-date model documentation

Signatories, when placing a general-purpose AI model on the market, will have documented at least all the information referred to in the Model Documentation Form below (hereafter this information is referred to as the 'Model Documentation'). Signatories may choose to complete the Model Documentation Form provided in the Appendix to comply with this commitment.

Signatories will update the Model Documentation to reflect relevant changes in the information contained in the Model Documentation, including in relation to updated versions of the same model, while keeping previous versions of the Model Documentation for a period ending 10 years after the model has been placed on the market.

Measure 1.2 Providing relevant information

Signatories, when placing a general-purpose AI model on the market, will publicly disclose via their website, or via other appropriate means if they do not have a website, contact information for the AI Office and downstream providers to request access to the relevant information contained in the Model Documentation, or other necessary information.

Signatories will provide, upon a request from the AI Office pursuant to Articles 91 or 75(3) AI Act for one or more elements of the Model Documentation, or any additional information, that are necessary for the AI Office to fulfil its tasks under the AI Act or for national competent authorities to exercise their supervisory tasks under the AI Act, in particular to assess compliance of high-risk AI systems built on general-purpose AI models where the provider of the system is different from the provider of the model,[2] the requested information in its most up-to-date form, within the period specified in the AI Office's request in accordance with Article 91(4) AI Act.

[1] See Article 75(1) and (3) AI Act and Article 88(2) AI Act.
[2] See Article 75(1) and (3) and Article 88(2) AI Act

Signatories will provide to downstream providers the information contained in the most up-to-date Model Documentation that is intended for downstream providers, subject to the confidentiality safeguards and conditions provided for under Articles 53(7) and 78 AI Act. Furthermore, without prejudice to the need to observe and protect intellectual property rights and confidential business information or trade secrets in accordance with Union and national law, Signatories will provide additional information upon a request from downstream providers insofar as such information is necessary to enable them to have a good understanding of the capabilities and limitations of the general-purpose AI model relevant for its integration into the downstream providers' AI system and to enable those downstream providers to comply with their obligations pursuant to the AI Act. Signatories will provide such information within a reasonable timeframe, and no later than 14 days of receiving the request save for exceptional circumstances.

Signatories are encouraged to consider whether the documented information can be disclosed, in whole or in part, to the public to promote public transparency. Some of this information may also be required in a summarised form as part of the training content summary that providers must make publicly available under Article 53(1), point (d), AI Act, according to a template to be provided by the AI Office.

Measure 1.3 Ensuring quality, integrity, and security of information

Signatories will ensure that the documented information is controlled for quality and integrity, retained as evidence of compliance with obligations in the AI Act, and protected from unintended alterations. In the context of drawing-up, updating, and controlling the quality and security of the information and records, Signatories are encouraged to follow the established protocols and technical standards.

Model documentation form

Below is a static, non-editable version of the Model Documentation Form. In this version, the input fields cannot be filled in. An interactive and fillable version of this form is separately available.

Model Documentation Form

This Form includes all the information to be documented as part of Measure 1.1 of the Transparency Chapter of the Code of Practice. Crosses on the right indicate whether the information documented is intended for the AI Office (AIO), national competent authorities (NCAs) or downstream providers (DPs), namely providers of AI systems who intend to integrate the general-purpose AI model into their AI systems. Whilst information intended for DPs should be made available to them proactively, information intended for the AIO or NCAs is only to be made available following a request from the AIO, either ex officio or based on a request to the AIO from NCAs. Such requests will state the legal basis and purpose of the request and will concern only items from the Form strictly necessary for the AIO to fulfil its tasks under the AI Act at the time of the request, or for NCAs to exercise their supervisory tasks under the AI Act at the time of the request, in particular to assess compliance of high-risk AI systems built on general-purpose AI models where the provider of the system is different from the provider of the model.

Any elements of information from the Model Documentation Form shared with the AIO and NCAs shall be treated in accordance with the confidentiality obligations and trade secret protections set out in Article 78 AI Act.

Date this document was last updated: Click or tap to enter a date.

Document version number: Click or tap here to enter text.

General information		AIO	NCAs	DPs
Legal name for the model provider:	Click here to add text.	☒	☒	☒
Model name:	The unique identifier for the model (e.g. Llama 3.1-405B), including the identifier for the collection of models where applicable, and a list of the names of the publicly available versions of the concerned model covered by the Model Documentation.	☒	☒	☒
Model authenticity:	Evidence that establishes the provenance and authenticity of the model (e.g. a secure hash if binaries are distributed, or the URL endpoint in the case of a service), where available.	☒	☒	☐
Release date:	Click or tap to enter a date. Date when the model was first released through any distribution channel.	☒	☒	☒
Union market release date:	Click or tap to enter a date. Date when the model was placed on the Union market.	☒	☒	☒
Model dependencies:	If the model is the result of a modification or fine-tuning of one or more general-purpose AI models previously placed on the market, list the model name(s) (and relevant version(s) if more than one version has been placed on the market) of those model(s). Otherwise write 'N/A'.	☒	☒	☒

Model properties		AIO	NCAs	DPs
Model architecture:	A general description of the model architecture, e.g. a transformer architecture. *[Recommended 20 words]*	☒	☒	☒
Design specifications of the model:	A general description of the key design specifications of the model, including rationale and assumptions made, to provide basic insight into how the model was designed. *[Recommended 100 words]* If any other please specify:	☒	☒	☐
Input modalities:	☐Text ☐Images ☐Audio ☐Video If any other please specify:	☒	☒	☒
For each selected modality please include maximum input size or write 'N/A' if not defined	Maximum size: … Maximum size: … Maximum size: … Maximum size: … Maximum size: …	☒	☐	☒
Output modalities:	☐Text ☐Images ☐Audio ☐Video If any other please specify:	☒	☒	☒
For each selected modality please include maximum output size or write 'N/A' if not defined	Maximum size: … Maximum size: … Maximum size: … Maximum size: … Maximum size: …	☐	☐	☒
Total model size:	The total number of parameters of the model, recorded with at least two significant figures, e.g. 7.3*10^10 parameters.	☒	☐	☐
The range within which the total number of parameters falls.	☐1—500M ☐500M—5B ☐ 5B—15B ☐15B—50B ☐50B—100B ☐100B—500B ☐500B—1T ☐>1T	☐	☒	☒

Methods of distribution and licenses		AIO	NCAs	DPs
Distribution channels:	A list of the methods of distribution (e.g. enterprise or subscription-based access through existing software suites or enterprise-specific solutions; public or subscription-based access through an API; public or proprietary access through integrated development environments, device-specific applications or firmware, open-source repositories) through which the model has been made available for distribution or use in the Union market. For each listed method of distribution, please include a link to information about how the model can be accessed, where available, and the level of model access (e.g. weights-level access, black-box access).	☒	☒	☐

	A list of the methods of distribution (e.g. enterprise or subscription-based access through existing software suites or enterprise-specific solutions; public or subscription-based access through an API; public or proprietary access through integrated development environments, device-specific applications or firmware, open-source repositories) through which the model can be made available to downstream providers.	☐	☐	☒
License:	A link to model license(s) (otherwise provide a copy of the license(s) upon a request from the AIO pursuant to Article 91) or indicate that no model license exists.	☒	☒	☐
	The type or category of licence(s) under which the model can be made available to downstream providers such as free and open source licences where models can be openly shared and providers can freely access, use, modify and redistribute them or modified versions thereof; less permissive licenses that impose certain restrictions on the use (e.g. to ensure ethical use), or proprietary licences that restrict access to the model's source code and impose limitations on usage, distribution, and modification. In the absence of a license, describe how access to the model is provided for downstream use, such as through terms of service.	☐	☐	☒
	A list of additional assets (e.g. training data, data processing code, model training code, model inference code, model evaluation code), if any, that are made available with a description of how each can be accessed and what licenses, if any, relate to their use.	☒	☐	☒

Use		**AIO**	**NCAs**	**DPs**
Acceptable Use Policy:	Provide a link to the acceptable use policy applicable (or attach a copy to this document) or indicate that none exists.	☒	☒	☒
Intended uses:	A description of either (i) the uses that are intended by the provider (e.g. productivity enhancement, translation, creative content generation, data analysis, data visualisation, programming assistance, scheduling, customer support, variety of natural language tasks, etc..) or (ii) the uses that are restricted and/or prohibited by the provider (beyond those prohibited by EU or international law, including Article 5 AI Act), in both cases as specified in the information supplied by the provider in the instructions for use, terms and conditions, promotional or sales materials and statements, as well as in the technical documentation. If specifying (i) or (ii) is incompatible with the nature of the license under which the model is provided, then 'N/A' can be entered. *[Recommended 200 words]*	☒	☒	☒
Type and nature of AI systems in which the general-purpose AI model can be integrated:	A list or description of either (i) the type and nature of AI systems into which the general-purpose AI model can be integrated or (ii) the type and nature of AI systems into which the general-purpose AI model should not be integrated. Examples may include autonomous systems, conversational assistants, decision support systems, creative AI systems, predictive systems, cybersecurity, surveillance, or human-AI collaboration. *[Recommended up to 300 words]*	☒	☒	☒
Technical means for model integration:	A general description of the technical means (e.g. instructions for use, infrastructure, tools) required for the general-purpose AI model to be integrated into AI systems. *[Recommended 100 words]*	☐	☐	☒
Required hardware:	A description of any hardware, including the version, required to use the model, where applicable. If not applicable (e.g. model offered via an API), 'N/A' should be entered. *[Recommended 100 words]*	☐	☐	☒
Required software:	A description of any software, including the version, required to use the model where applicable. If not applicable, 'N/A' should be entered. *[Recommended 100 words]*	☐	☐	☒

Training process		**AIO**	**NCAs**	**DPs**
Design specifications of the training process:	A general description of the main steps or stages involved in the training process, including training methodologies and techniques, the key design choices, assumptions made and what the model is designed to optimise for, and the relevance of different parameters, as applicable. For example, "the model is initialized with randomly selected weights and optimised using gradient-based optimization via the Adam optimizer in two steps. First, the model is trained to predict the next word on a large pretraining corpus using the cross-entropy loss, passing over the data for a single epoch. Second, the model is post-trained on a dataset of human preferences for 10 epochs to align the model with human values and make it more useful in responding to user prompts". *[Recommended 400 words]*	☒	☒	☐
Decision rationale:	A description of how and why key design choices were made in model training. *[Recommended 200 words]*	☒	☒	☐

Information on the data used for training, testing, and validation					**AIO**	**NCAs**	**DPs**
Data type/modality: *Select all that apply.*	☐Text	☐Images	☐Audio	☐Video	If any other please specify:	☒	☒	☒
Data provenance: *Select all that apply*	☐Web crawling		☐Private non-publicly available datasets obtained from third parties	☐User data		☒	☒	☒
For definitions of each listed category, see the Template for the Public Summary of the Training Content of General-Purpose AI models provided by the AI Office	☐Publicly available datasets			☐Data collected through other means				
	☐Synthetic data that is not publicly accessible (when created directly by or on behalf of the provider)			If any other please specify:				

How data was obtained and selected:	A description of the methods used to obtain and select training, testing, and validation data, including methods and resources used to annotate data, and models and methods used to generate synthetic data where applicable. For data previously obtained from third parties, a description of how the provider obtained the rights to the data if not already disclosed in the public summary of training data published in accordance with Article 53(1), point (d). *[Recommended 300 words]*	☒	☒	☐
Number of data points:	The size (in number of data points) of the training, testing, and validation data respectively, together with the definition of the unit of data points (e.g. tokens or documents, images, hours of video or frames), recorded with at least one significant figure (e.g. $3x10^{13}$ tokens).	☐	☒	☐
	The size (in number of data points) of the training, testing, and validation data respectively, together with the definition of the unit of data points (e.g. tokens or documents, images, hours of video or frames), recorded with at least two significant figures (e.g. $1.5x10^{13}$ tokens).	☒	☐	☐
Scope and main characteristics:	A general description of the scope and main characteristics of the training, testing and validation data, such as domain (e.g. healthcare, science, law,...), geography (e.g. global, restricted to a certain region,...), language, modality coverage, where applicable. *[Recommended 200 words]*	☒	☒	☐
Data curation methodologies:	General description of the data processing involved in transforming the acquired data into training, testing, and validation data for the model, such as cleaning (e.g. filtering out irrelevant content such as advertisements), normalisation (e.g. tokenizing), augmentation (e.g. back-translation). *[Recommended 300 words]*	☒	☒	☒
Measures to detect unsuitability of data sources:	A description of any methods implemented in data acquisition or processing, if any, to detect the presence of unsuitable data sources considering the model's intended uses, including but not limited to illegal content, child sexual abuse material (CSAM), non-consensual intimate imagery (NCII), and personal data leading to its unlawful processing. *[Recommended 400 words]*	☒	☒	☐
Measures to detect identifiable biases:	A description of any methods implemented in data acquisition or processing, if any, to address the prevalence of identifiable biases in the training data. *[Recommended 200 words]*	☒	☒	☐
Computational resources (during training)		**AIO**	**NCAs**	**DPs**
Training time:	A description of what period is being measured along with the range that its duration falls under, within the following ranges: less than 1 month, 1—3 months, 3—6 months, more than 6 months.	☐	☒	☐
	A description of what period is being measured along with the duration in wall clock days (e.g. $9x10^{1}$ days) and in hardware days (e.g. $4x10^{5}$ Nvidia A100 days and $2x10^{5}$ Nvidia H100 days), both recorded with at least one significant figure.	☒	☐	☐
Amount of computation used for training:	Measured or estimated amount of computation used for training, reported in floating point operations and recorded up to its order of magnitude (e.g. 10^{24} floating point operations).	☐	☒	☐
	Measured or estimated amount of computation used for training, reported in computational operations and recorded with at least two significant figures (e.g. $2.4x10^{25}$ floating point operations).	☒	☐	☐
Measurement methodology:	In the absence of a delegated act adopted in accordance with Article 53(5) AI Act to detail measurement and calculation methodologies, describe the methodology used to measure or estimate the amount of computation used for training.	☒	☒	☐
Energy consumption (during training and inference)		**AIO**	**NCAs**	**DPs**
Amount of energy used for training:	Measured or estimated amount of energy used for training, reported in Megawatt-hours and recorded with at least two significant figures (e.g. $1.0x10^{2}$ MWh). If the amount of energy used for training cannot be estimated due to the lack of critical information from a compute or hardware provider, enter 'N/A'.	☒	☒	☐
Measurement methodology:	In the absence of a delegated act adopted in accordance with Article 53(5) AI Act to detail measurement and calculation methodologies, describe the methodology used to measure or estimate the amount of energy used for training. Where the energy consumption of the model is unknown, the energy consumption may be estimated based on information about computational resources used. If the amount of energy used for training cannot be estimated due to a lack of critical information from a compute or hardware provider, the provider should disclose the type of information they lack. *[Recommended 100 words]*	☒	☒	☐
Benchmarked amount of computation used for inference[1]:	Benchmarked amount of computation used for inference, reported in floating point operations, recorded with at least two significant figures (e.g. $5.1x10^{17}$ floating point operations).	☒	☒	☐
Measurement methodology:	In the absence of a delegated act adopted in accordance with Article 53(5) AI Act to detail measurement and calculation methodologies, provide a description of a computational task (e.g. generating 100000 tokens) and the hardware (e.g. 64 Nvidia A100s) used to measure or estimate the amount of computation used for inference.	☒	☒	☐

[1] This item relates to energy consumption during inference, which makes up the "energy consumption of the model" (Annex XI, 2(e), AI Act) together with energy consumption during training. Since energy consumption during inference depends on more than just the model itself, the information required for this item is limited to relevant information depending only on the model, namely computational resources used for inference.