

ANNEX XI

Technical documentation referred to in Article 53(1), point (a) - technical documentation
for providers of general-purpose AI models

Section 1

Information to be provided by all providers of general-purpose AI models

The technical documentation referred to in Article 53(1), point (a) shall contain at least the following information as appropriate to the size and risk profile of the model:

1. A general description of the general-purpose AI model including:
 - (a) the tasks that the model is intended to perform and the type and nature of AI systems in which it can be integrated;
 - (b) the acceptable use policies applicable;
 - (c) the date of release and methods of distribution;
 - (d) the architecture and number of parameters;
 - (e) the modality (e.g. text, image) and format of inputs and outputs;
 - (f) the licence.

2. A detailed description of the elements of the model referred to in point 1, and relevant information of the process for the development, including the following elements:
- (a) the technical means (e.g. instructions of use, infrastructure, tools) required for the general-purpose AI model to be integrated in AI systems;
 - (b) 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;
 - (c) 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;
 - (d) 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;
 - (e) known or estimated energy consumption of the model.

With regard to point (e), where the energy consumption of the model is unknown, the energy consumption may be based on information about computational resources used.

Section 2

Additional information to be provided

by providers of general-purpose AI models with systemic risk

1. 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.
2. 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.
3. 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.
