

Kluwer Copyright Blog

Open Source Artificial Intelligence Definition 1.0 – A “take it or leave it” approach for open source AI systems?

Yaniv Benhamou, Michel Reymond (University of Geneva) · Tuesday, March 4th, 2025

Background

This blog post follows a [previous post](#) that discussed what constitutes “open source” AI, in particular in light of the EU AI Act. This post continues the discussion, in particular in light of the revision of the [Open Source AI Definition \(“OSAID”\)](#) released at the end of 2024, a welcome step in clarifying and unifying this definition.



Photo by [Luke Southern](#) from [Unsplash](#)

The definition of “open source” in AI development has been the subject of heavy debate and scrutiny. In the past two years, tech giants such as Meta and OpenAI released their AI models under self-certified “open” licenses. These releases have been criticized by open source experts as attempts at “openwashing”, namely because they are characterized by (1) the selective inclusion of model components and documentation, and (2) their restrictive license terms. It is worth looking at each point in turn.

Selective inclusion of model components and documentation. Unlike traditional software which may be deployed “out of the box” through access to source codes, AI models must be trained on enormous amounts of data so as to shape their behavior in fulfilling a specific task (e.g. a model being provided pictures of cat to recognize them as such). Due to this, open source specialists [have argued](#) that the release of a trained AI model alone fails to meet the definition of open source, as AI specialists cannot reproduce, analyze or modify its behavior without equal access to the other components shaping its performance. These components include the base AI model, the training and evaluation datasets, the model’s trained internal weights and parameters, and the documentation setting out the training process. Yet, [according to a recent report](#), full open releases of AI systems have not been the norm, as several self-described “open” releases do not provide access to training and evaluation datasets, and are furthermore lacking in documentation. Even the well-publicized recent release of DeepSeek, the Chinese AI large language model which disrupted both OpenAI and Meta, [is not exempt from this quandary since the downloadable model does include training datasets or training code](#).

In a [previous post](#), we argued that a more appropriate way of tackling what exactly constitutes “open source” would be to specify, on a case-by-case basis, the level of openness based on which components are open and which licensing terms apply.

Restrictive licensing terms. While described as “open“, AI models are often published under unilaterally drafted licensing terms which do not meet the openness standards set by even the less permissive open source licenses, such as the [GNU GPL](#) or the [BSD License](#). Famously, Meta’s release of the Llama 2 model in July 2023 contained a non-competition clause which prevented reuse for licensees having more than 700 million monthly active users. It also excluded reuse of the model for improving or fine-tuning other large language models. This led [commentators](#) to conclude that this release [was open source in name only](#). Another strand of restrictive licensing terms stems from Responsible AI Licenses (RAIL), which aim to prevent misuse of AI models and which [typically exclude unlawful uses such as discrimination or defamation](#). Regardless of their aims, RAIL terms are arguably at odds with currently accepted definitions of open source. For instance, Google’s 2024 release of its Gemma 2 model is advertised as an open model, yet, it comes with an expansive ‘[Prohibited Use Policy](#)’ which disallows use for several applications (e.g. spam, unlicensed practice of a profession, or sexually explicit content).

Against the above background – that of definitional uncertainty and of a risk of industry capture – stakeholders from the public and private sectors have pitched in with their own definition of open source for AI systems. [As reported previously](#), the EU crafted one such definition in its [AI Act](#) for the purpose of exempting open source AI projects from several of its documenting obligations. More recently, the Linux Foundation has also proposed the [Model Openness Framework](#) (“MOF”), which grades the openness of AI systems through a three-tiered system. On 28 October 2024, the Open Source Initiative finally published the first version of its [Open Source AI Definition](#) (“OSAI”).

The OSAID

Under the OSAID, an open source AI is a system which is made available under terms and in a way that grants the freedom to use, study, modify and share it. The definition applies to the base and trained models, to the weights and parameters and to “other structural elements”; it also requires the availability of the “complete source code used to train and run the system”. Finally, the definition does not require provision of training data but “sufficiently detail(ed) information about the data used to train the system so that a skilled person can build a substantially equivalent system”.

Save for training data (discussed below), the OSAID thus adopts a “take it or leave it” approach requiring all components of an AI system to be made available in compliance with the four open source freedoms (to use, study, modify and share). In this sense, it tracks rather closely with the definition of open source of Article 53 (2) of the AI Act, which also requires terms allowing for the “access, usage, modification, and distribution of the model”. By the same token, the OSAID is incompatible with RAIL terms and other restrictive licensing terms; it also departs from the more modular approach adopted by the Model Openness Framework. [This has led commentators to suggest that the OSAID perpetuates a conservative approach to open source.](#)

Training data and data sharing

The OSAID does not require the licensor to provide access to non-open or public training data, but only to “sufficiently detail(ed) information about the data used to train the system so that a skilled person can build a substantially equivalent system”. [As discussed in our previous post](#), requiring access to all training data would be too restrictive as this data is oftentimes limited by non-permissive terms (such as the deployment of technical tools to prevent scraping) or by other legal issues such as copyright and trade secrets. The approach chosen by the OSAID is thus in line with these considerations.

Beyond that, there remains the question of what exactly constitutes “sufficiently detailed information” about the training data. The OSAID states that this information “must include:

- (1) the complete description of all data used for training, including (if used) of unshareable data, disclosing the provenance of the data, its scope and characteristics, how the data was obtained and selected, the labeling procedures, and data processing and filtering methodologies;
- (2) a listing of all publicly available training data and where to obtain it; and
- (3) a listing of all training data obtainable from third parties and where to obtain it, including for fee”.

This requirement is, in our opinion, rather expansive, as it requires disclosure of how the data is used in the training process and does not provide an exception for trade secrets. This is consistent with the OSAID’s stringent “take it or leave it” approach.

The above disclosure obligation echoes an ongoing debate regarding Article 53 (1)(d) of the AI Act, under which providers of general-purpose AI must “draw up and make publicly available a sufficiently detailed summary about the content used for training of the general-purpose AI model”. On 17 January 2025, [the AI Office presented a template for this summary](#), which essentially requires disclosure of information about the AI model, about the training data (including its size, its characteristics, and its sources), and about measures implemented to respect copyright (and other rights) and the removal of unwanted content. The template [has received mixed reactions from the OSI community](#), notably due to its lack of granularity as to the obligation to disclose training data.

That the OSAID’s disclosure obligation is stricter than that foreseen under Article 53 (1)(d) AI Act is justified by the respective goals of the two instruments. While the AI Act generally governs access to the EU market, the OSAID, as discussed above, aims to impose a conservative watermark for AI system providers to certify their releases as “open source”. The coexistence of the two standards should thus not in principle lead to frictions or contradictions.

A table summarizing our comparison is provided below.

	OSAID	EU AI Act
Components of the AI model to be released	The model, the weights and parameters and its “other structural elements”, as well as “complete source code used to train and run the system”.	AI models released “under a free and open source licence that allows for the access, usage, modification, and distribution of the model, and whose parameters, including the weights, the information on the model architecture, and the information on model usage, are made publicly available” (Article 53 (2))
Training data to be released	No training data, but “[s]ufficiently detail(ed) information about the data used to train the system so that a skilled person can build a substantially equivalent system”, i.e. (1) the complete description of all data used for training, including (if used) of unshareable data; disclosing the provenance of the data, its scope and characteristics, how the data was obtained and selected, the labeling procedures, and data processing and filtering methodologies; (2) a listing of all publicly available training data and where to obtain it; and (3) a listing of all training data obtainable from third parties and where to obtain it, including for fee.	No training data, but providers general purpose AI models must provide a “[s]ufficiently detailed summary (art. 53 (d) AI Act).” Requires information about the training data (including its size, its characteristics, and its sources), and about measures implemented to respect copyright (and other rights) and the removal of unwanted content.

Final Thoughts

The OSAID is a welcome step in clarifying and unifying the definition of open source in the realm of AI models and systems. Its “take it or leave it” approach along with its obligation for AI providers to disclose detailed information on their training data will be useful to curb industry capture of the open source movement. Yet this approach might lead to only a subset of AI projects being compatible with its requirements. As we previously argued, a more flexible approach, including several tiers of “openness” may be preferred, as in the MOF proposal by the Linux Foundation. Nevertheless, with both the OSAID and the AI Act adopting similar approaches, it may very well be that the definition of open source in the AI space will remain narrowly framed.

To make sure you do not miss out on regular updates from the Kluwer Copyright Blog, please subscribe [here](#).

2024 Future Ready Lawyer Survey Report

Legal innovation: Seizing the future or falling behind?

Download your free copy →

 Wolters Kluwer



 Future
Ready
LAWYER

This entry was posted on Tuesday, March 4th, 2025 at 4:24 pm and is filed under [Artificial Intelligence \(AI\)](#), [European Union](#)

You can follow any responses to this entry through the [Comments \(RSS\)](#) feed. You can leave a response, or [trackback](#) from your own site.