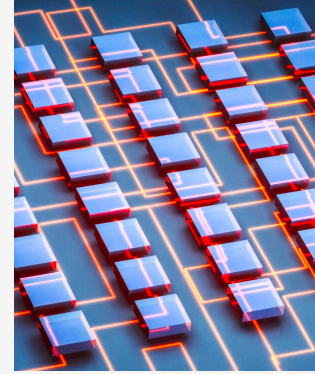


OSI releases first candidate definition of ‘open source AI’ for final comments

OCTOBER 16, 2024 5 MIN READ



Related Expertise

- [Artificial Intelligence](#)
- [Emerging and High Growth Companies](#)
- [Privacy and Data Management](#)
- [Technology](#)

Authors: [Sam Ip](#), [Joseph Ierullo](#), [Naomi Chernos](#)

As artificial intelligence (AI) continues to reshape industries, the role of open source AI systems is becoming increasingly relevant. Open source AI offers a collaborative and transparent foundation that can promote collaboration, transparency, foster innovation, lower barriers to entry and learning, and support ethical development. As open source AI gains traction, there is a growing need to clarify what “open source” means in the context of AI.

The [Open Source Initiative](#) (OSI), an influential and leading open source organization responsible for the “Open Source Definition” (OSD), has committed to addressing this. To ensure open source principles can be effectively applied to AI, OSI has been working since 2022, in consultation with global experts and the broader public, to create a new definition for open source AI. These efforts have resulted in a much anticipated [first candidate “Open Source AI” definition](#) (OSAID). Notably, the Osler team has been following this process, including through OSI’s various town halls.

The release of the first candidate OSAID emerges at a pivotal time when standards bodies, institutions and governments globally, including Canada, are introducing AI regulations that often contemplate the role of open source AI. To date, the Release Candidate 1 has received endorsements from several influential open source organizations, including Mozilla, SUSE, and the Eclipse Foundation.

Key elements of OSI’s proposed definition of open source AI

The OSI’s proposed open source AI definition is comprised of the following:

1. **What open source AI is:** An open source AI is an AI system that grants the freedoms to use, study, modify, and share the system, and its discrete elements. Exercising these freedoms requires access to the preferred form to make modifications to these systems.
2. **Preferred form to make modifications to machine learning systems.** The definition requires an open source AI to be made available in a preferred form that include (i) detailed information about the data used to train the system (ii) the complete source code used to train and run the system (iii) the model parameters, including weights and other

configuration settings.

3. **Open source models and open source weights.** The definition specifies that the requirements for an AI system to be open source will also apply to the AI model itself, the code used to run the model, the model architecture, model parameters, and the AI weights (defined as the set of learned parameters that overlay the model architecture to provide outputs).
4. **Definitions.** Definitions of “AI System” and “Machine Learning” are provided, relying on definitions from the OECD Artificial Intelligence guidance, clarifying the scope of these terms when used in the definition.

Reactions to this definition

While the definition is still open for comment, two of the more pressing concerns relate to data availability and interoperability.

The data dilemma

Data lies at the heart of machine learning, and ensuring the right level of disclosure required to constitute open source AI is crucial. The current draft does not mandate sharing the actual training data but only its metadata, which some argue could compromise the reproducibility of AI Systems. In particular, this means that a model trained on proprietary non-open data may still meet the terms of open source, which is somewhat contrary to the spirit and ethos of open source.

However, mandating the sharing of training data could limit adoption, as it could give rise to both competitive sensitivities or risk of noncompliance with regulation (e.g. copyright, data protection, etc.) OSI proposes that the amount of information about any given dataset required to meet the standard will depend on the sensitivity of the data. Whether this compromise is sufficiently clear or workable will require further discussion with stakeholders.

Interoperability with standards and regulations

Another significant issue impacting successful adoption of open source AI is ensuring consistency and coherence across a wide range of evolving global standards and regulations. OSAID broadly defines “AI System” in line with the [OECD AI Principles](#) and the EU AI Act. These definitions also map quite closely to the proposed definition in Canada's *Artificial Intelligence and Data Act (Part 3 of Bill C-27)*, although it has yet to be determined whether this definition will be adopted. It is worth noting that the EU AI Act contains exemptions not mirrored by the OSAID, such as exempting certain open source AI Systems, or AI Systems used for the sole purpose of scientific research and development. However, the absence of these exemptions likely will not affect the definition for the purposes of an open source license.

On the other hand, the International Organization for Standardization (ISO) takes a more tailored approach, defining “AI” in the context of an “engineered system” to center human involvement. This may give rise to interoperability challenges for organizations which base their AI governance strategy around the widely accepted ISO standards.

Looking forward

The Open Source Initiative's definition of open source AI lays the groundwork for a more inclusive, transparent, and ethical approach to AI development. As we navigate the evolving landscape of AI, embracing open source principles may be key to unlocking its full potential while safeguarding ethical standards.

In keeping with its ultimate goal of a definition that has been developed in collaboration with experts and stakeholders, OSI is seeking endorsements of the definition, as well as collecting further feedback from stakeholders, through forums, letters, and upcoming town halls, before releasing the official 1.0 definition on October 28. OSI is also focusing on further refining the accompanying documentation (including a checklist and FAQ). Given the likely impact of the Open Source AI Definition, as the dialogue around open source AI continues, it is essential for all stakeholders — developers, researchers, policymakers, and users — to engage with these principles and work together toward a future they wish to create.