

G7 Vision on AI openness opportunities and shared language

SCOPE

This document is addressed to the broader AI ecosystem of G7 members and beyond, including governments, regulatory bodies, intergovernmental institutions, think tanks, non-governmental organizations (NGOs), model providers, foundations, Open Source communities and civil society. It may also serve as a reference for AI providers, companies, including Micro, Small and Medium Enterprises (MSMEs), innovation and digital promotion bodies, researchers, scientists, research funders, and public authorities. The objective is to call for greater clarity in the use of terminology describing AI openness and to encourage further G7 work on this matter, in cooperation with the community.

This document constitutes a non-binding reference. It is without prejudice to applicable national or international frameworks, including those related to AI governance, cybersecurity, data protection and intellectual property.

INTRODUCTION

1. We, the G7 Digital and Technology Ministers, recognize that Open Source Software has become fundamental for international research collaborations and discoveries, collaborative innovation and digital progress including by lowering barriers to technological access across economies. Similarly, in the field of AI, openness has delivered great benefits for G7 economies and beyond with further opportunities ahead. While recognising the significant benefits of AI openness for innovation, research and economic growth, G7 members also acknowledge that openness should be assessed in light of issues such as dual-use concerns, security vulnerabilities and supply-chain risks, among others.

Over the past three decades, software has transformed G7 economies providing benefits directly and indirectly and enabled unprecedented growth opportunities. During that time, Open Source software has become a cornerstone of our economies, substantially reducing software development costs, serving as the engine of discovery by enabling reproducibility, accelerating research, and lowering barriers to participation for researchers worldwide.

Artificial intelligence is set to transform G7 economies, and the openness of AI has the potential to play a similar role to openness of software and other dimensions for openness. However, AI poses new challenges in terms of definitional clarity: the meaning of Open-Weight or Open Source AI remains contested, and while the traditional definition of Open Source provides pointers, further clarification is needed.

This lack of clarity in the field of AI tends to cast doubt on the degree of openness of such technologies, thereby undermining their benefits. Therefore, We, the G7 Digital and Technology Ministers, encourage the broader ecosystem to be conscious of the nuances in the language used to describe the openness of AI. In order to support the ecosystem in moving towards a more shared understanding, we present this document for a G7 Vision on AI openness opportunities and shared language: a statement on the economic benefits related to AI openness and a set of principles including a more nuanced typology of openness, intended to contribute to the evolution of this emerging field. Clear distinctions can help users, policymakers and markets better understand the practical implications of different degrees of openness with regards to AI.

2. AI openness has been an essential contributor to our economies, fostering innovation and cooperation, and broadening access to technologies for companies and communities.

The benefits of AI openness span all layers of the AI lifecycle, from optimizing physical infrastructure to enabling advanced, AI-driven business solutions. As highlighted in the OECD paper on The Benefits of AI Openness, open-weight AI models and related open-source software can have positive micro- and macro-economic impacts, including on Gross Domestic Product. The benefits include strong value-for-money, particularly at scale, while supporting innovation, local value creation, sector-specific adaptation, and strategic digital autonomy. At the same time, future work would be needed to scope potential risks

presented by AI that is not fully open such as malicious use, security vulnerabilities, and challenges around accountability and oversight.

In this context, a clearer and shared vision of AI labelled as “open” would bring significant added value by reducing uncertainty, supporting innovation, contributing to economic growth and avoiding practices such as open washing.

PRINCIPLES

Principle 1 — AI Openness is community-driven

The development of Open Source norms, standards, and definitions has been shaped over time by the communities of developers, researchers, scientists, startups, and practitioners, among others, who build and use these technologies, whether contributing for profit or not-for-profit. G7 members acknowledge the central role of these communities in shaping what openness means in practice, they recognize existing efforts and encourage continued engagement between public authorities and Open Source communities in the evolution of shared standards for open AI.

Principle 2 — AI Openness is a spectrum

The openness of an AI is not binary. It exists on a spectrum, ranging from models that share only weights under restricted licenses, to models that make all elements fully available under open licenses. G7 members and stakeholders are encouraged to recognize and communicate which degree of openness applies to a given AI, rather than using the term “open” without further qualification.

Given the spectrum nature of openness, it is challenging to be exhaustive in developing precise definitions. However, an initial attempt to establish a technical typology can help G7 members to describe it more accurately. We further develop them below.

Principle 3 — AI has multiple elements

The openness of an AI is determined by the combination of the following elements (see Glossary), including but not limited to:

- Use restrictions
- Models’ weights
- Deployment code
- Training code
- Training data

Each additional component made available influences the degree of openness and, with it, the potential for innovation, reproducibility, and resilience.

Principle 4 — AI Openness is described through appropriate labels

In the absence of commonly accepted definitions of “open” terms as applied to AI, stakeholders are encouraged to use terminology that accurately and transparently reflects the degree of openness of a given model. As a minimum, any description of an AI as “open” should clearly state which components are made available, whether with or without restrictions, rather than using the term “open” as a blanket characterization or implying a greater degree of openness than what is actually offered.

TYPOLGY

The following tiered typology describes the different degrees of openness of AI, ordered from most open to less open. Given the spectrum nature of openness, such typology is not intended to be exhaustive, but rather to provide a common reference point moving beyond a purely binary vision. G7 members are encouraged to adopt this typology as a shared technical reference when describing AI. The G7 underscores the importance of transparency regarding access conditions, licensing terms, documentation and usage constraints when referring to AI as “open”.

- **Open Source AI with Open Data**

Open Source AI with Open Data is AI released free of charge under an Open Source license, including its models' weights, deployment code, training code, and full training data.

- **Open Source AI**

Open Source AI is AI released free of charge under an Open Source license, including its models' weights, deployment code, training code, and possibly full training data.

However, exceptions to full training data availability may apply where sharing is legally or technically impossible. In such cases, Data Information (see Glossary) about the unavailable data is to be provided in its place.

- **Open Weights AI**

Open Weights AI is AI released free of charge along with its weights and deployment code, under an Open Source License (see Glossary).

- **Weights Available AI**

Weights Available AI is AI released free of charge along with its weights and deployment code under a license which includes restrictions on how they can be used, such as commercial, geographical, or use-case restrictions.

Glossary

AI Model: In this document, an AI model refers to the component of an AI system that produces results through prediction and inference.

AI System: In this document, an AI system refers to a machine-based system designed to operate with varying levels of autonomy, which may exhibit adaptiveness after deployment, and which, for explicit or implicit objectives, infers from the input it receives how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments.

Training Data Information: Sufficiently detailed information about the data used to train the model. This must should include: a complete description of all training data, including any unshareable data, covering provenance, scope, characteristics, collection and selection methods, labeling procedures, and processing and filtering methodologies; a listing of all publicly available training data and where to obtain it; and a listing of all training data obtainable from third parties, including for a fee, and where to obtain it.

Deployment Code: The code required to run the AI or to ensure compatibility with standard inference frameworks.

Open Source License: A license that permits use, inspection, modification, and redistribution without restrictions. Lists of recognized Open Source licenses can be found at <https://opensource.org/licenses> and <https://www.gnu.org/licenses/license-list.html>.

Training Code: The code used to process training data and train the AI, including: data processing and filtering code; training code with all arguments and settings; validation and testing code; supporting libraries such as tokenizers and hyperparameter search code; and model architecture specifications.

Training Data: The data used to train the AI.

Use Restrictions: Conditions in the license that restrict how the model may be used, including commercial, geographical, or use-case limitations.

Weights: The learned numerical parameters of an AI model produced through training.