Artificial Analysis Openness Index

Specification V1.0

Artificial Analysis is a leading independent AI benchmarking and insights provider

Artificial Analysis is an independent AI benchmarking company. Our website is a widely referred-to resource for independent benchmarks and is used every month by hundreds of thousands of people choosing which AI technologies to use. We operate across the entire AI value chain, including hardware and infrastructure, inference providers, models and applications, and users and enterprises adopting AI.

Our public platform, artificialanalysis.ai, is trusted by companies driving the frontier of AI as well as leading corporations, media, research and government institutions. We have been publicly referenced by:

Artificial Analysis has developed a framework to support the open source AI ecosystem by measuring openness across 2 dimensions: model availability and transparency

Objectives	• Provide a standardized framework that enables evaluating the openness of models • Provide useful synthesis metrics that can be used to compare the relative openness of models • Support AI labs in communicating the openness of their models	
Context	We see two primary dimensions of openness relevant to the open source AI ecosystem	
	1. Model availability	2. Model transparency
Focus	*Freedom to access and use the model without restriction*	*Degree to which the public could recreate the model or reuse its methodology and data*
Key questions of end-users	*"Can I use this model and access its outputs, and use it for commercial purposes?"* *"Can I run this model on my own hardware?"* *"Do I need to give attribution to the model creator if I use or fine-tune this model?"*	*"What optimizer and hyperparameters do I need to train this model?"* *"What data sources went into this model's training?"* *"Can I train a new model with the same techniques?"*
Why it matters	• Openness of usage is the 'normal' lens, representing whether the public can access model weights • License terms like commercial use limits or attribution also impact 'how open' model weights are	• Replication requires visibility of data recipes, allowing users to better understand factors like intellectual property, biases, and model capabilities • Disclosure of training methodologies and code enables model training advancement across the field

The Artificial Analysis Openness Index – a standardized framework to measure model openness

Dimension	Subcomponent	Scoring		
1. Model availability	**Weights**	Access	0	Closed weights, no API
			1	Closed weights, API limits token visibility
			2	Closed weights, API available
			3	Open weights
		License	0	Closed weights or no commercial use
			1	Commercial use, attribution required
			2	Commercial use, no attribution required
			3	Commercial use, no attribution required, no meaningful limitations
2. Model transparency	**Data: Pre & Post Training** *(score represents average across each)*	Access	0	No or limited disclosure
			1	Partial data source detail and categorization disclosed
			2	Full data mix disclosure, substantial data shared[1]
			3	Full data shared
		License (most restrictive)	0	No commercial use/no substantial data shared
			1	Commercial use, attribution required
			2	Commercial use, no attribution required
			3	Commercial use, no attribution required, no meaningful limitations
	Methodology	Disclosure	0	No or limited disclosure
			1	Model architecture disclosure
			2	Limited general technical disclosure
			3	Full technical details disclosed
		License/ Tooling	0	No code disclosed/released
			1	Frameworks disclosed, openly available for commercial use
			2	End-to-end training pipeline code or guide released
			3	End-to-end training pipeline code or guide released, and commercial use allowed

1. Partial data sharing requirement threshold varies for pre- and post-training components – see appendix for detail

Scoring methodology

Each subcomponent is scored on **a scale** based on the best-fitting openness 'archetype' with clear definitions

We synthesize these underlying factors into a unified metric, **the Artificial Analysis Openness Index:**

- Data elements are averaged between pre- and post-training
- Sum of all component scores (up to a maximum of 18/18)
- Normalized to a 0-100 scale

We will publish both the overall Openness Index and the subcomponent scores.

We will aim to visualize the data to make it most useful to users – potentially including a simple bar chart, a stacked bar chart and a scatterplot with usage and replication scores on the X and Y axes respectively.

Model availability: Model availability detailed scoring archetype definitions

Access

Score	High level	Details	Examples
0	**Closed weights, no API**	Closed models with no publicly accessible API (first-party or partner hosted)	Grok 4 Heavy, Windsurf SWE-1.5
1	**Closed weights, API limits token visibility**	Closed models with APIs showing only some output tokens (e.g., no raw reasoning tokens)	GPT-5.1, Gemini 3 Pro
2	**Closed weights, API available**	Closed models with publicly-available APIs showing all output tokens	Claude Sonnet 4.5, Qwen3 Max
3	**Open weights**	Model weights are publicly available for download	Meta Llama 4, DeepSeek v3.2, OLMo 3

License

Score	High level	Details	Examples
0	**Closed weights or no commercial use**	Weights not available, or license disallows commercial use (e.g., research-only)	GPT-5.1, Meta Llama 1
1	**Commercial use, attribution required**	Weights available with license allowing commercial use but requiring attribution	Meta Llama 2-4, Kimi K2
2	**Commercial use, no attribution**	Commercial use allowed without attribution, but may include other limitations	NVIDIA Nemotron Nano 9B v2
3	**No meaningful limitations**	Weights available without meaningful limitations (e.g., MIT/Apache 2.0)	gpt-oss-120b, GLM-4.6, OLMo 3

Model transparency: Pre-training data detailed scoring archetype definitions

Access

Score	High level	Details	Examples
0	**No or limited disclosure**	No disclosure or general references to data usage without specific sources	Grok 4, GPT-5, Magistral Small 1.2, DeepSeek v3.2, gpt-oss-120b
1	**Partial data source detail and categorization disclosed**	Details of training data sources are disclosed, such as specific source types/dataset names, filtering approach, and detailed token distribution	Llama 3.3 70B, NVIDIA Llama Nemotron Super 49B v1.5
2	**Full data mix disclosure, >1T tokens of pretraining data shared**	Disclosures of both: - Full details of pre-training data sources (including both shared and withheld sources), and - Open release of more than 1 trillion tokens of pre-training data	NVIDIA Nemotron Nano 9B v2
3	**Full data shared**	Full disclosure of all training data needed to produce the final checkpoint if access were available to the original training code	OLMo 3

License (most restrictive[1])

Score	High level	Details	Examples
0	**No commercial use**	Data not available for commercial use, or no substantial data released	Stanford Alpaca 7B, NVIDIA Llama Nemotron Super 49B v1.5
1	**Commercial use, attribution required**	Data released with license allowing commercial use but requiring attribution	OLMo 3, NVIDIA Nemotron Nano 9B v2
2	**Commercial use, no attribution**	Data released allowing commercial use without attribution requirement	—
3	**No meaningful limitations**	Data available for commercial use without meaningful limitations	—

1. License (most restrictive) reflects the most restrictive license applied to shared data components

Model transparency: Post-training data detailed scoring archetype definitions

Access

Score	High level	Details	Examples
0	**No or limited disclosure**	No disclosure or general references to data usage without specific sources	Grok 4, GPT-5, Gemini 2.5 Pro, Claude Sonnet 4.5, DeepSeek v3.2, gpt-oss-120b
1	**Partial data source detail and categorization disclosed**	Details of training data sources are disclosed, such as specific source types/dataset names, generation methodology, and filtering approach	Llama 3.3 70B, Magistral Small 1.2, NVIDIA Llama Nemotron Super 49B v1.5
2	**Full data mix disclosure, some usable training data shared**	Both: - Disclosure of full details of post-training data sources (including both shared and withheld sources), and - Significant partial release of post-training data that enables independent reuse for training	NVIDIA Nemotron Nano 9B v2
3	**Full data shared**	Full disclosure of all training data needed to produce the final checkpoint if access were available to the original training code	OLMo 3

License (most restrictive[1])

Score	High level	Details	Examples
0	**No commercial use**	Data not available for commercial use, or no data released	—
1	**Commercial use, attribution required**	Data released with license allowing commercial use but requiring attribution	NVIDIA Nemotron Nano 9B v2, OLMo 3
2	**Commercial use, no attribution**	Data released allowing commercial use without attribution requirement	—
3	**No meaningful limitations**	Data available for commercial use without meaningful limitations	—

1. License (most restrictive) reflects the most restrictive license applied to shared data components

Model transparency: Methodology detailed scoring archetype definitions

Disclosure

Score	High level	Details	Examples
0	**No or limited disclosure**	No disclosure or general references to methodology without meaningful detail	Grok 4, GPT-5, Gemini 2.5, Claude 4.5
1	**Model architecture disclosure**	Basic architecture details disclosed (transformer variant, parameter count, layer config)	Qwen3 family, GLM-4.6
2	**Limited general technical disclosure**	Qualitative technical details (e.g., post-training approach) or partial hyperparameters	DeepSeek v3.2, Llama 3.3 70B
3	**Full technical details disclosed**	All key technical details disclosed: optimization, hyperparameters, RL algorithms	NVIDIA Nemotron Nano 9B v2, GLM-4.5, OLMo 3

License/Tooling (most restrictive[1])

Score	High level	Details	Examples
0	**No code disclosed/released**	No training code, frameworks, or implementation details released	GPT-5, Claude 4.5, Gemini 2.5
1	**Frameworks disclosed, openly available**	Frameworks/dependencies disclosed and openly available for commercial use	NVIDIA Nemotron Nano 9B v2
2	**End-to-end training pipeline code or guide released**	Full training code with end-to-end pipeline scripts or guides to reproduce checkpoint	—
3	**End-to-end training pipeline code or guide released, and commercial use allowed**	Full training code with E2E pipeline AND license permitting commercial use	OLMo 3

1. License/Tooling (most restrictive) reflects the most restrictive license or tooling availability applied to shared methodological components

Artificial Analysis

hello@artificialanalysis.ai

https://artificialanalysis.ai/