GENERATIVE AI COMMONS

MODEL OPENNESS FRAMEWORK

Open Licenses in the Age of AI

May 2024

LF AI & DATA

GENERATIVE AI COMMONS

🇪🇺🤓 Researcher at Linux Foundation

📜🤖 PhD in Social Data Science at University of Oxford on Political Economy of Open Source AI

👩‍💻🌍 Contributor to scikit-learn (community) & OSI “Defining Open Source AI” Working Group

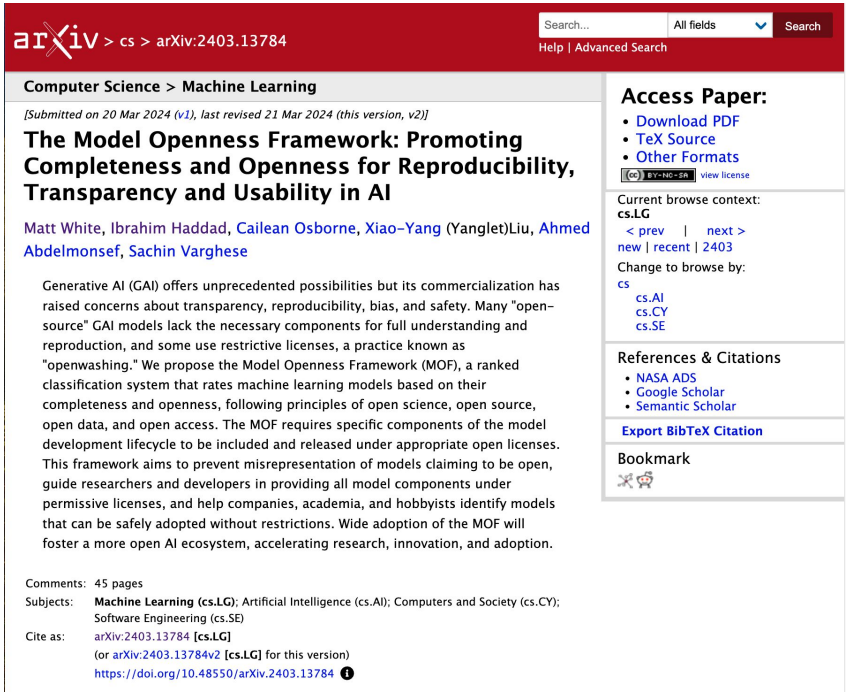
𝕏 @cailean_osborne

📧 cailean.osborne@contractor.linuxfoundation.org

LF AI & DATA

Introducing The Model Openness Framework

https://arxiv.org/abs/2403.13784



arXiv > cs > arXiv:2403.13784

Search... All fields Search

Help | Advanced Search

Computer Science > Machine Learning

[Submitted on 20 Mar 2024 (v1), last revised 21 Mar 2024 (this version, v2)]

The Model Openness Framework: Promoting Completeness and Openness for Reproducibility, Transparency and Usability in AI

Matt White, Ibrahim Haddad, Cailean Osborne, Xiao-Yang (Yanglet)Liu, Ahmed Abdelmonsef, Sachin Varghese

Generative AI (GAI) offers unprecedented possibilities but its commercialization has raised concerns about transparency, reproducibility, bias, and safety. Many "open-source" GAI models lack the necessary components for full understanding and reproduction, and some use restrictive licenses, a practice known as "openwashing." We propose the Model Openness Framework (MOF), a ranked classification system that rates machine learning models based on their completeness and openness, following principles of open science, open source, open data, and open access. The MOF requires specific components of the model development lifecycle to be included and released under appropriate open licenses. This framework aims to prevent misrepresentation of models claiming to be open, guide researchers and developers in providing all model components under permissive licenses, and help companies, academia, and hobbyists identify models that can be safely adopted without restrictions. Wide adoption of the MOF will foster a more open AI ecosystem, accelerating research, innovation, and adoption.

Comments: 45 pages

Subjects: **Machine Learning (cs.LG)**; Artificial Intelligence (cs.AI); Computers and Society (cs.CY); Software Engineering (cs.SE)

Cite as: arXiv:2403.13784 **[cs.LG]**

(or arXiv:2403.13784v2 **[cs.LG]** for this version)

https://doi.org/10.48550/arXiv.2403.13784

Access Paper:

- Download PDF
- TeX Source
- Other Formats

CC BY-NC-SA view license

Current browse context: cs.LG

< prev | next >

new | recent | 2403

Change to browse by:

cs

cs.AI

cs.CY

cs.SE

References & Citations

- NASA ADS
- Google Scholar
- Semantic Scholar

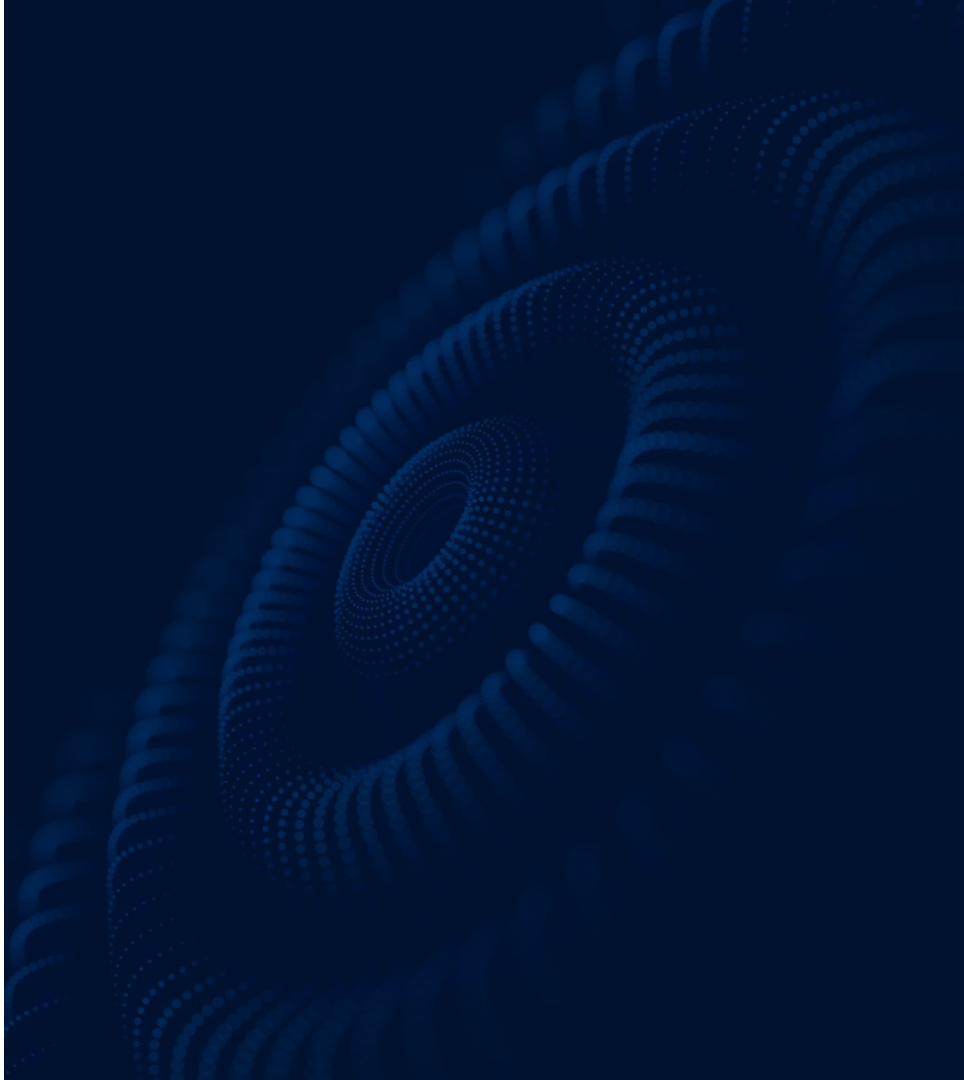
Export BibTeX Citation

Bookmark

Agenda

- Background
- Concepts of Openness in Science & Technology
- Introducing the MOF (Components, Licenses, Classes, Implementation)
- Benefits & Limitations of the MOF
- Relevance to EU AI Act
- Conclusion
- Questions

Background



MODEL OPENNESS FRAMEWORK

BACKGROUND

The rise in the **commercialization of General Artificial Intelligence (GAI)** has introduced numerous challenges.

These include growing concerns about **reproducibility, transparency, safety, ethics, and appropriate usage**, etc..

A significant number of AI systems function as **black boxes**, posing challenges in comprehending their operations and ensuring fairness.

INDUSTRY NEEDS

"Open" models lack essential components for full understanding and reproduction. Some use confusing licenses and engage in **"openwashing."**

To address this, we propose the **Model Openness Framework**, evaluating models based on completeness and openness with open science principles.

ADDED VALUE

The **Model Openness Framework offers guidance** to researchers and developers seeking to enhance model transparency and reproducibility while allowing permissive usage.

For enterprise and industry, it provides a **clarity about which models are suitable for commercial use without restrictions.**

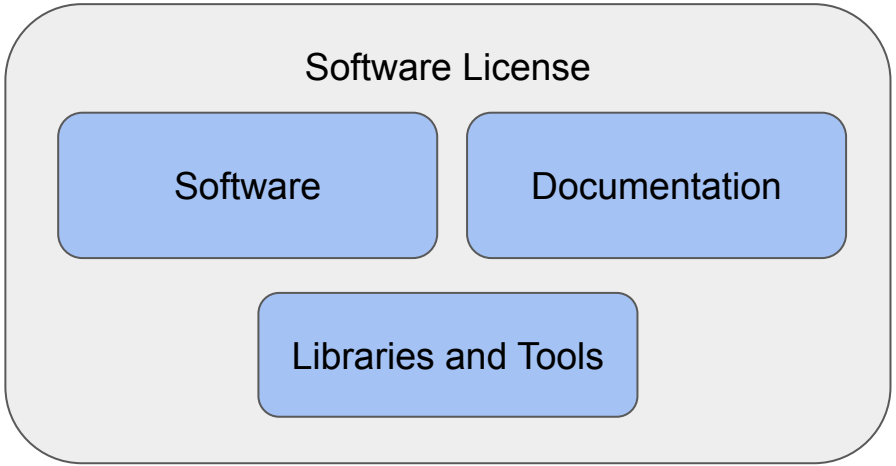
Challenges in Open AI

- Rapidly growing # of available models
- A lot of “open-washing”
- Many licenses with restrictions (AUPs)
- No consistent definition of “open” in AI
- Lack of understanding of license implications
- Many components not released (e.g. datasets)
- Illegally converted licenses
- Open source licenses used on non-software
- AUPs and user agreements used on model weights

Open vs. Proprietary/Restrictive Licenses

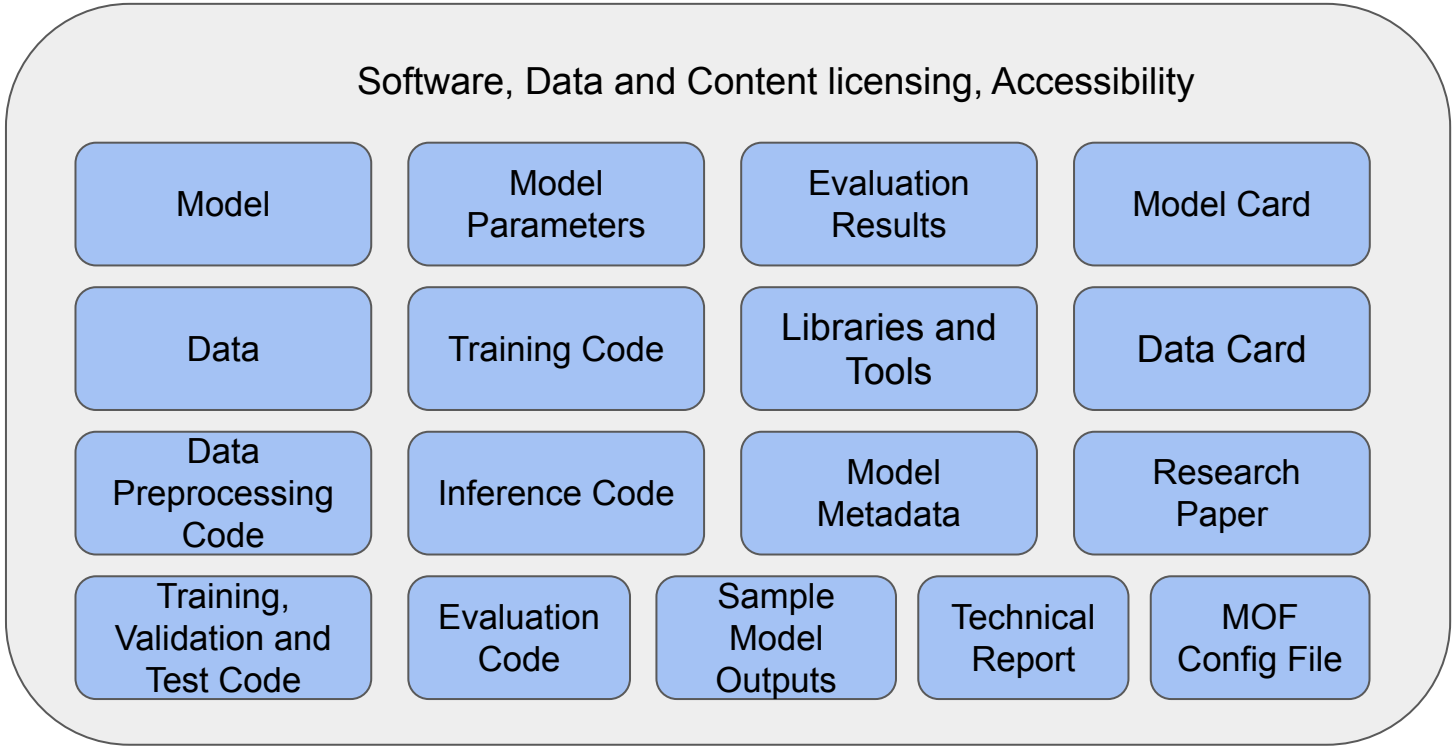
Open	Proprietary/Restrictive
Apache 2.0 (Pythia, MPT, Mis/xtral, RWKV, Gemma)	Open RAIL (BLOOM/Stable Diffusion)
MIT (GPT-2, Dolly)	Llama 2 (Llama 2)
BSD-3-Clause	DBRX Model License (DBRX)
CDLA-Permissive-2.0	TII License (Falcon)
CC-BY-SA-4.0 (StableLM-Alpha)*	Google Gemma AUP (Gemma weights)
	Ai2 ImpACT (Tulu 2)

Pre-Open AI Era - Software Licensing

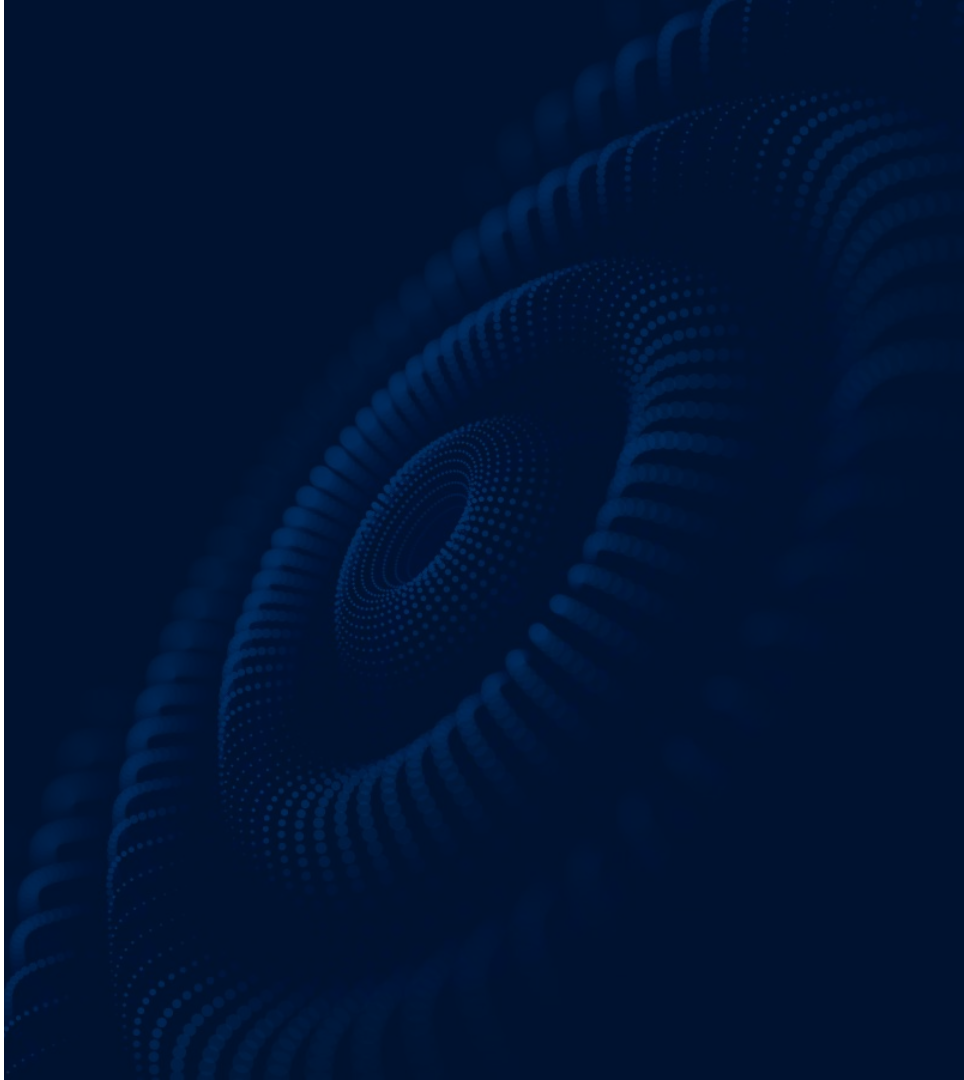


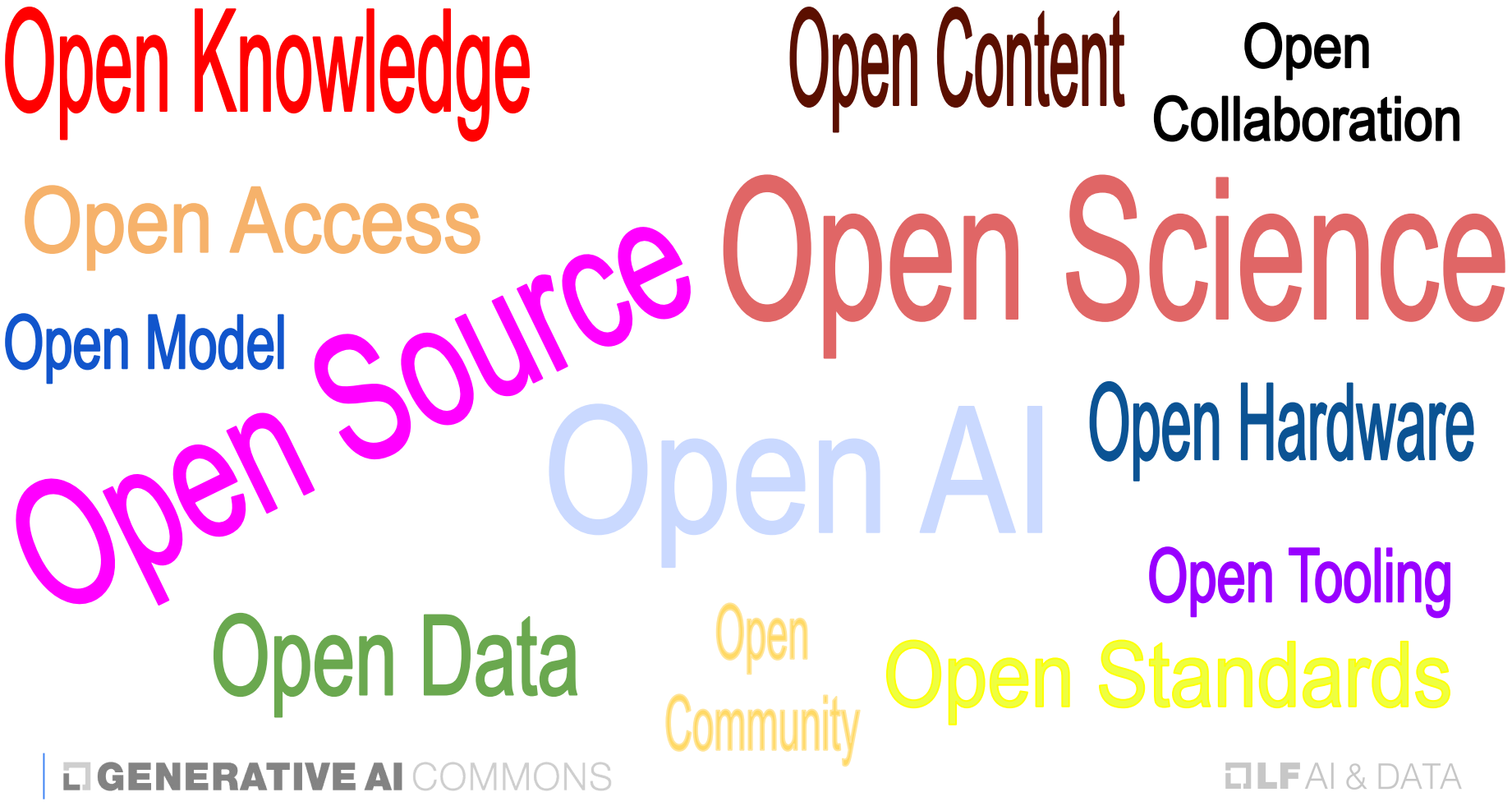
Source: Wikipedia

Post-Open AI Era - Everything Licensing



Concepts of Openness




Open Knowledge
Open Content
Open Collaboration
Open Access
Open Science
Open Model
Open Source
Open AI
Open Hardware
Open Tooling
Open Data
Open Community
Open Standards

Open Science

Open science is a movement that advocates for making scientific research, data, and dissemination accessible to all levels of society. It encompasses practices such as:

- **Open access:** Making research publications freely available online without paywalls or subscription fees.
- **Open data:** Sharing research data openly, allowing others to access, reuse, and redistribute it.
- **Open source:** Sharing code used in research to others without restrictions.
- Open methodology: Documenting and sharing the methods, tools, and software used in research for transparency and reproducibility.
- **Open peer review:** Conducting the peer review process transparently, with the identities of authors and reviewers disclosed.
- **Open collaboration:** Fostering collaboration among researchers, institutions, and disciplines to accelerate scientific progress.

The goal of open science is to increase transparency, reproducibility, and accessibility in research, leading to more efficient and impactful scientific advancements that benefit society as a whole.

Open Source

Open source refers to a model of software development and distribution in which the source code of a program is made publicly accessible, allowing anyone to view, modify, and distribute the code. The key principles of open source include:

- **Transparency:** The source code is openly shared, enabling users to understand how the software works and to verify its functionality and security.
- **Collaboration:** Open source projects encourage collaboration among developers, who can contribute improvements, bug fixes, and new features to the codebase.
- **Free distribution:** Open source software can be freely used, modified, and distributed by anyone, without the need for licensing fees or restrictive agreements.
- **Community-driven:** Open source projects are often driven by a community of developers and users who work together to maintain and improve the software.
- **Licensing:** Open source software is typically released under specific licenses, such as the GNU General Public License (GPL) or the MIT License, which grant users the rights to use, modify, and distribute the code while ensuring that these freedoms are preserved.

The open source model has been widely adopted in software development, leading to the creation of numerous successful projects, such as the Linux operating system, the Apache web server, and the Mozilla Firefox browser. The principles of open source have also been applied to other domains, such as open source hardware and open source scientific research.

**OSI working on a definition of "open-source AI". Our working term is "open AI".*

**Most of today's "open-source models" are actually "source-available models". Availability and open are not one in the same.*

Open Data

Open data refers to data that is freely available for anyone to access, use, modify, and share without restrictions. The key principles of open data include:

- **Availability:** Open data is made publicly available and can be easily accessed by anyone, typically through online platforms or repositories.
- **Machine-readability:** Open data is provided in formats that can be easily processed by computers, such as CSV, JSON, or XML, enabling automated analysis and integration with other systems.
- **Reusability:** Open data is licensed in a way that allows for its reuse and redistribution, including for commercial purposes, with minimal or no restrictions.
- **Completeness:** Open data should be as complete as possible, providing all relevant information needed for its interpretation and use.
- **Timeliness:** Open data should be made available as quickly as possible after its creation or collection to maximize its value and relevance.

Open data initiatives are often driven by governments, public institutions, and research organizations to promote transparency, accountability, and innovation.

**Currently the biggest challenge in open AI is the lack of shared datasets and methodologies used to treat and train on datasets.*

Open Content

Open content refers to creative works that are made freely available for anyone to access, use, modify, and share without restrictions. The key principles of open content include:

- **Availability:** Open content is made publicly available and can be easily accessed by anyone, typically through online platforms or repositories.
- **Reusability:** Open content is licensed in a way that allows for its reuse and redistribution, including for commercial purposes, with minimal or no restrictions.
- **Modifiability:** Open content can be modified, adapted, or built upon by others, enabling derivative works and collaborative improvements.
- **Completeness:** Open content should be as complete as possible, providing all relevant information or components needed for its interpretation and use.
- **Transparency:** Open content should be accompanied by clear licensing terms that specify the permissions and conditions under which it can be used and shared.

Open content initiatives are often driven by individuals, organizations, and communities that value the benefits of sharing knowledge and creative works freely, such as promoting education, innovation, and cultural exchange. Examples of open content include open educational resources, open source software documentation, and creative works published under Creative Commons licenses.

**Currently the biggest challenge with open content is that there is very little open content, most is copyrighted. However open content licenses are appropriate for documentation.*

Open Access

Open access (OA) refers to the practice of making research outputs, such as scholarly articles and books, freely available online without any paywall, subscription fees, or other access restrictions. The main characteristics of open access include:

- **Free availability:** Open access content can be accessed, read, downloaded, and distributed by anyone with an internet connection.
- **Permanent access:** Open access materials remain accessible indefinitely, ensuring long-term preservation and availability.
- **Immediate access:** Open access content is made available as soon as it is published, without any embargo period.
- **Reuse rights:** Open access materials are typically published under licenses that allow for reuse, such as Creative Commons licenses, enabling others to build upon and share the work with proper attribution.

The goal of open access is to remove barriers to accessing scholarly research, making it more widely available to researchers, students, policymakers, and the general public. This increased accessibility can lead to greater visibility, impact, and progress in various fields of study.

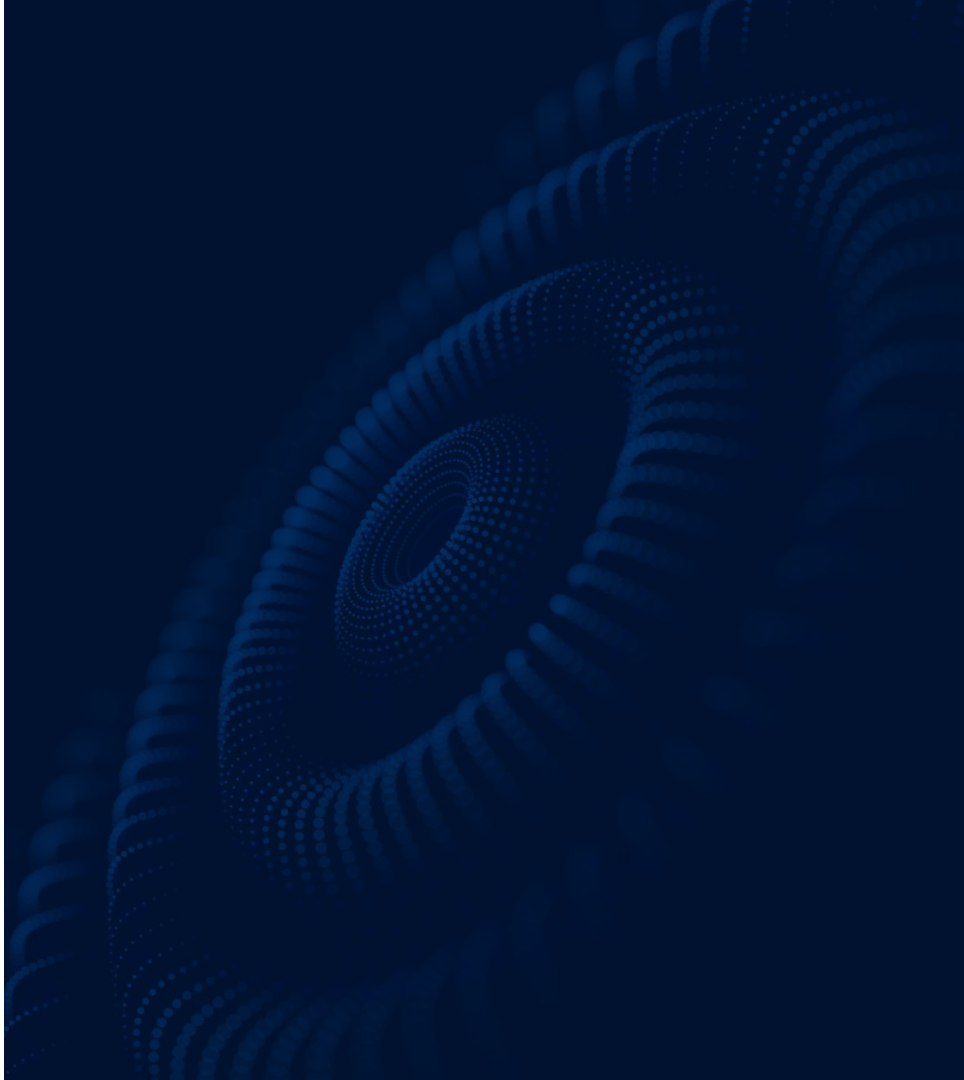
**Please do not say "Open access models" or "open access weights". More appropriate term is "widely available"*

Completeness vs. Openness

- **Completeness:** In open science, completeness refers to providing comprehensive and well-documented information for all components, ensuring that each element is thorough, self-contained, and meaningfully usable without requiring additional context or resources.

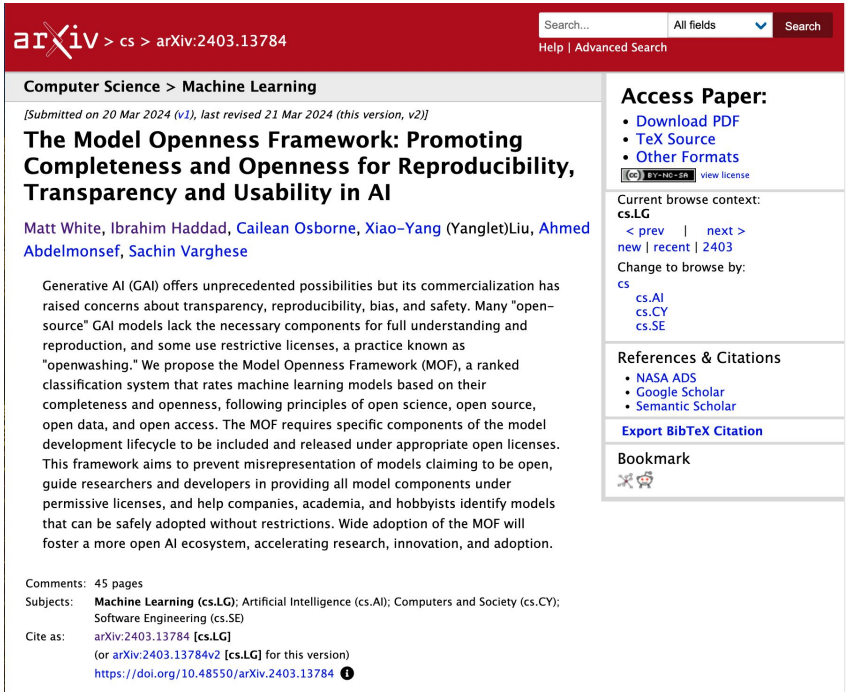
- **Openness**: Openness is a binary property that indicates whether a particular component is licensed under an open license or not. A component is considered "open" if and only if it is distributed under a license that grants users the rights to freely access, use, modify, and share the component. If a component is not licensed under such an open license, it is considered "not open." There is no intermediate state between "open" and "not open"; a component either meets the criteria for openness or it does not.

The Solution: Model Openness Framework



Introducing the Model Openness Framework

https://arxiv.org/abs/2403.13784



arXiv > cs > arXiv:2403.13784

Search... All fields Search

Help | Advanced Search

Computer Science > Machine Learning

[Submitted on 20 Mar 2024 (v1), last revised 21 Mar 2024 (this version, v2)]

The Model Openness Framework: Promoting Completeness and Openness for Reproducibility, Transparency and Usability in AI

Matt White, Ibrahim Haddad, Cailean Osborne, Xiao-Yang (Yanglet)Liu, Ahmed Abdelmonsef, Sachin Varghese

Generative AI (GAI) offers unprecedented possibilities but its commercialization has raised concerns about transparency, reproducibility, bias, and safety. Many "open-source" GAI models lack the necessary components for full understanding and reproduction, and some use restrictive licenses, a practice known as "openwashing." We propose the Model Openness Framework (MOF), a ranked classification system that rates machine learning models based on their completeness and openness, following principles of open science, open source, open data, and open access. The MOF requires specific components of the model development lifecycle to be included and released under appropriate open licenses. This framework aims to prevent misrepresentation of models claiming to be open, guide researchers and developers in providing all model components under permissive licenses, and help companies, academia, and hobbyists identify models that can be safely adopted without restrictions. Wide adoption of the MOF will foster a more open AI ecosystem, accelerating research, innovation, and adoption.

Comments: 45 pages

Subjects: **Machine Learning (cs.LG)**; Artificial Intelligence (cs.AI); Computers and Society (cs.CY); Software Engineering (cs.SE)

Cite as: arXiv:2403.13784 **[cs.LG]**

(or arXiv:2403.13784v2 **[cs.LG]** for this version)

https://doi.org/10.48550/arXiv.2403.13784

Access Paper:

- Download PDF
- TeX Source
- Other Formats

CC BY-NC-SA view license

Current browse context: **cs.LG**

< prev | next >

new | recent | 2403

Change to browse by:

cs

cs.AI

cs.CY

cs.SE

References & Citations

- NASA ADS
- Google Scholar
- Semantic Scholar

Export BibTeX Citation

Bookmark

What is Openness in AI?

- **Freedom:** The ability to view, use, modify, and distribute for any purpose including commercial usage.
- **Accessibility:** Freely available for anyone.
- **Collaboration:** Encourages collective input and contributions.
- **Transparency:** The development process, research, and artifacts are open to public scrutiny.

Achieved through open licenses!

MOF Components

CODE


Evaluation Code


Preprocessing Code


Model Architecture


Libraries & Tools


Training Code

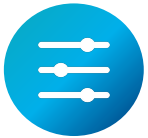

Inference Code

DATA


Datasets


Evaluation Data


Sample Model Outputs*


Model Weights & Parameters


Model Metadata


Configuration File

DOCUMENTATION


Data Card


Research Paper


Evaluation Results


Model Card


Technical Report

*Can be code or data

MOF Acceptable Licenses

COMPONENT	DOMAIN	CONTENT TYPE	ACCEPTED OPEN LICENSE
Datasets	Data	Data	Preferred: CDLA-Permissive-2.0, CC-BY-4.0 Acceptable: Any including unlicensed
Data Preprocessing Code	Data	Code	Acceptable: OSI-approved
Model Architecture	Model	Code	Acceptable: OSI-approved
Model Parameters	Model	Data	Preferred: CDLA-Permissive-2.0 Acceptable: OSI-Approved, Permissive Open Data Licenses
Model Metadata	Model	Data	Preferred: CDLA-Permissive-2.0 Acceptable: CC-BY-4.0, Permissive Open Data Licenses
Training Code	Model	Code	Acceptable: OSI-approved
Inference Code	Model	Code	Acceptable: OSI-approved
Evaluation Code	Model	Code	Acceptable: OSI-approved
Evaluation Data	Model	Data	Preferred: CDLA-Permissive-2.0 Acceptable: CC-BY-4.0, Permissive Open Data Licenses
Evaluation Results	Model	Documentation	Preferred: CC-BY-4.0 Acceptable: Permissive Open Content Licenses
Supporting libraries and Tools	Model	Code	Acceptable: OSI-approved
Model Card	Model	Documentation	Preferred: CC-BY-4.0 Acceptable: Permissive Open Content Licenses
Data Card	Data	Documentation	Preferred: CC-BY-4.0 Acceptable: Permissive Open Content Licenses
Technical Report	Model & Data	Documentation	Preferred: CC-BY-4.0 Acceptable: Permissive Open Content Licenses
Research Paper	Model & Data	Documentation	Preferred: CC-BY-4.0 Acceptable: Permissive Open Content Licenses
Sample Model Outputs	Model	Data or Code	Unlicensed

MOF Classes

MOF Class	Components Included
Class III – Open Model	• Model Architecture • Model Parameters (final checkpoint) • Technical Report • Evaluation Results • Model Card • Data Card • Sample Model Outputs (optional)

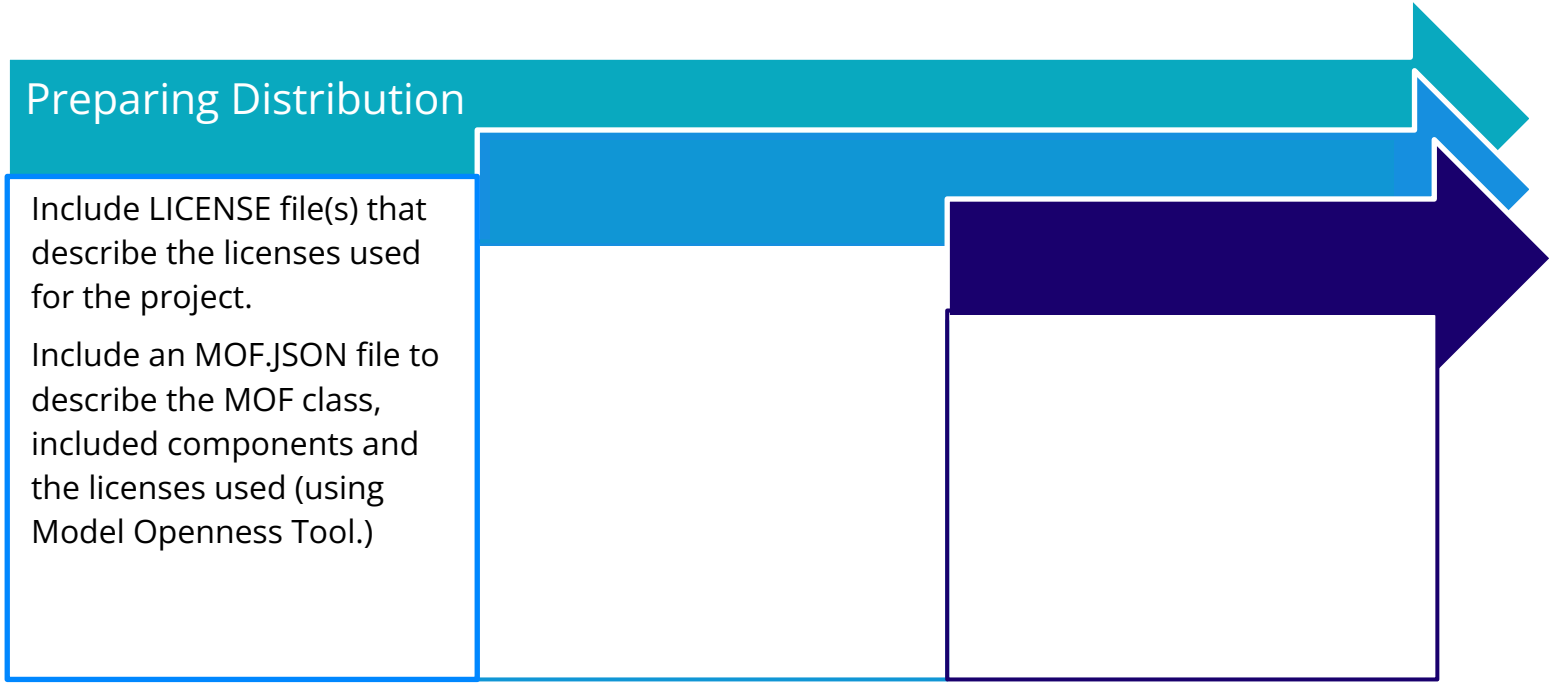
MOF Classes

MOF Class	Components Included
Class II – Open Tooling	• Training Code • Inference Code • Evaluation Code • Evaluation Data • Supporting Libraries & Tools (optional) • And all Class III Components
Class III – Open Model	• Model Architecture • Model Parameters (final checkpoint) • Technical Report • Evaluation Results • Model Card • Data Card • Sample Model Outputs (optional)

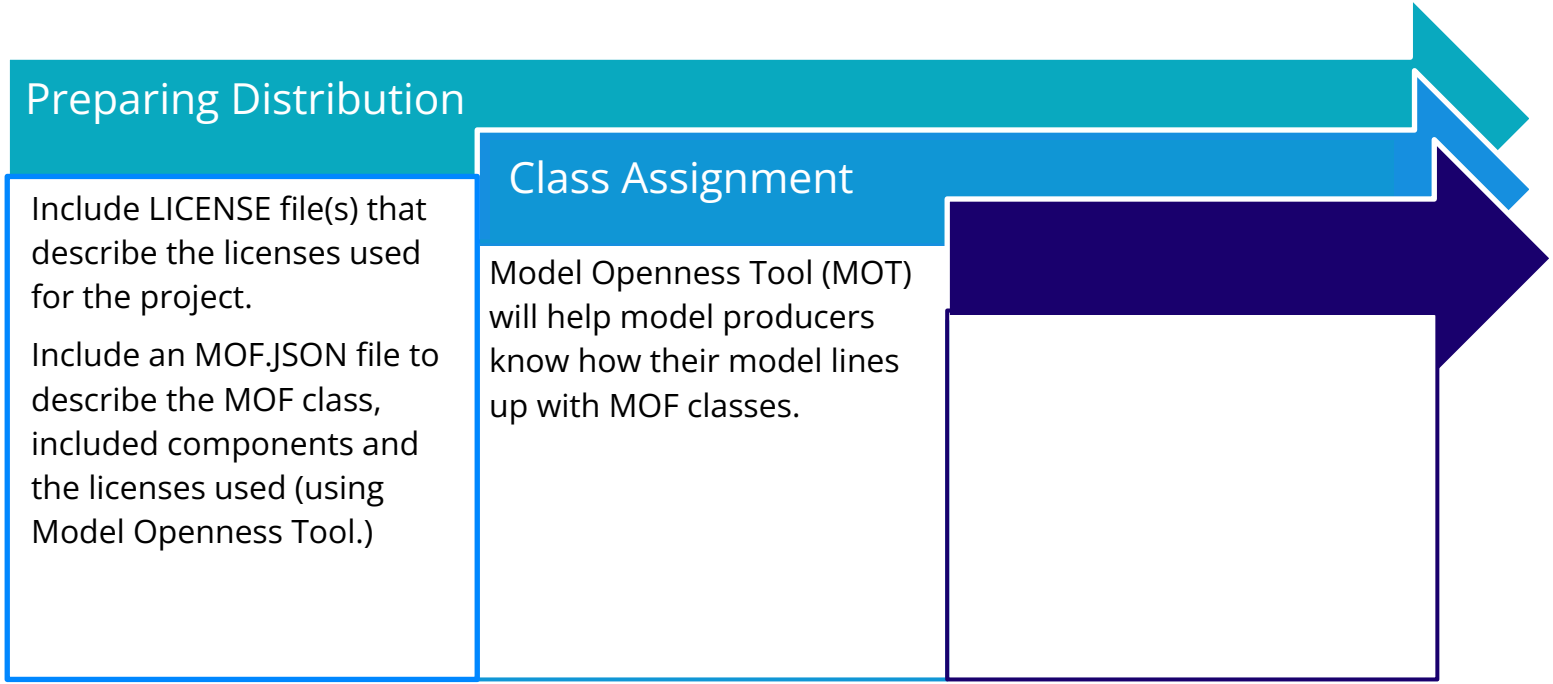
MOF Classes

MOF Class	Components Included
Class I – Open Science	• Research Paper • Datasets (any license or unlicensed) • Data Preprocessing Code • Model Parameters (intermediate checkpoints) • Model Metadata (optional) • And all Class II Components
Class II – Open Tooling	• Training Code • Inference Code • Evaluation Code • Evaluation Data • Supporting Libraries & Tools (optional) • And all Class III Components
Class III – Open Model	• Model Architecture • Model Parameters (final checkpoint) • Technical Report • Evaluation Results • Model Card • Data Card • Sample Model Outputs (optional)

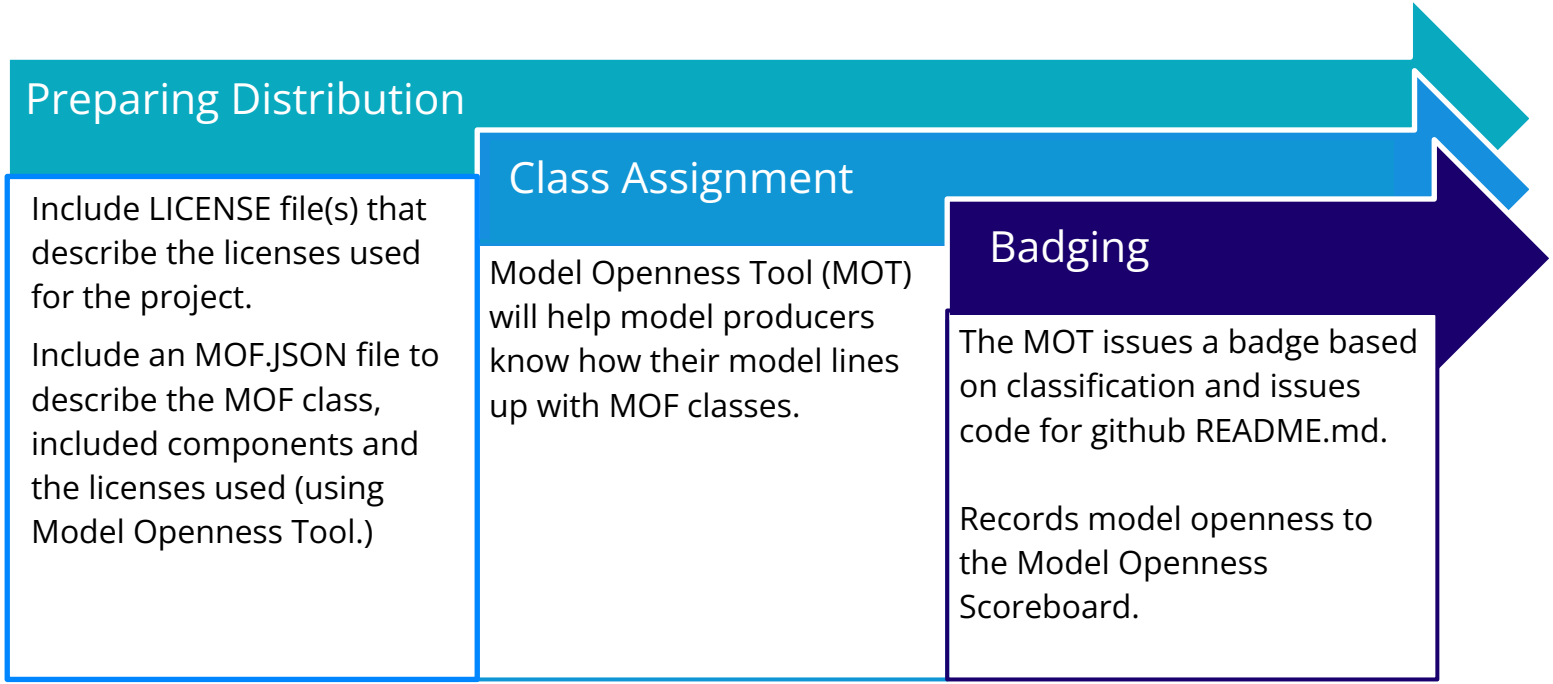
MOF Implementation



MOF Implementation



MOF Implementation



MOF Benefits

Clarity: The framework distinctly outlines the included components and the corresponding licenses, facilitating a clear understanding of acceptable use and indicating whether a project is genuinely open and complete.	**Accountability & Fairness**: Publicly available data and models can be audited for unwanted biases and harms, allowing the community to notify model producers of discovered issues.
Openness: By categorizing models and their artifacts into increasing levels of openness, the MOF encourages model producers to create comprehensive and open models. This advancement supports open science, benefiting both academic and commercial applications.	**Continuous Improvement:** Open models facilitate building upon existing work, accelerating innovation and progress in AI by enabling model producers and consumers to collaborate and evolve collectively.
Reproducibility: Providing comprehensive access to data, code, and models enables independent replication of results, helping identify errors, biases, or disparities and enhancing scientific rigor.	**Collaboration**: Sharing open resources encourages collaboration among model producers and consumers across different fields and organizations, pooling knowledge and capabilities.
Transparency & Explainability: Opening model architectures, weights, training code, and documentation illuminates how models operate, fostering trust and enabling scrutiny for potential issues.	**Education & Learning:** Open access to data, code, and models supports teaching and learning about AI, making it easier for students, new researchers, and developers to enter the field.
Data Provenance: Releasing details about data origins and attribution allows for tracing bias in models and identifying sources of Personally Identifiable Information (PII) leakage.	**Regulation**: Openness in models enhances their suitability for oversight and governance, unlocking potential public policy options.

Benefit to Model Producers

- Build a vibrant ecosystem around your models
- Spur innovation and improve upon your work
- Improve your models and datasets through feedback
- Appease regulators through greater transparency and reproducibility
- Improve safety and security of models

MOF Effect: Move model producers towards releasing more components using open licenses

Benefit to Model Consumers

- Clarity on what models can be used for what purposes
- Research, education and innovation
- Build products on top of open models
- Enhance models for one's own purposes
- Collaboration with a broader community
- Access to more components of the model development lifecycle

MOF Effect: Make it clear which models are actually open and what is included.

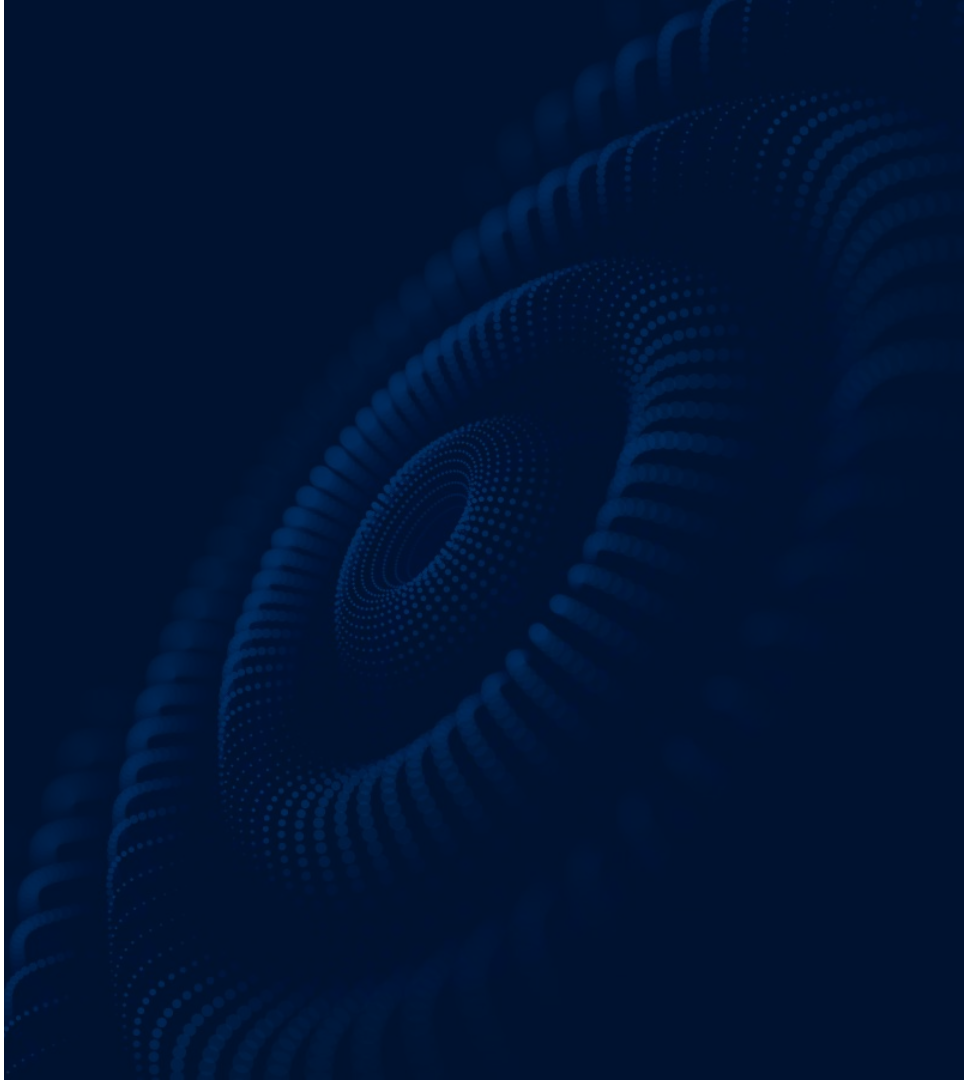
MOF Limitations

- The MOF is **tailored for deep learning artifacts** and may not directly apply to other learning forms, although a similar approach can be adopted for statistical machine learning or reinforcement learning.
- Model producers are urged to be transparent about the availability of released components, the openness of licenses for each, and the overall completeness of their models. **Convincing producers to share their work openly, especially without initial restrictions, is a challenge**.
- Openness objectives for components (e.g. datasets) need to be balanced against considerations like p**rivacy, data protection, intellectual property (IP), institutional policies, and commercialization pressures**.
- **Classifying models may not fully capture their functionality**, and concerns about bias, safety, and other harms persist. However, openness facilitates external audits for quality and completeness.
- Simplicity in classification **may overlook nuances**, but continuous improvement of the rubric is possible.
- **Does not address the use of copyrighted materials** in training data, an issue currently navigated through legal and legislative processes. Using an open license for data is sufficient for qualification as "open," but researchers and developers are encouraged to respect copyrights and use authorized data in model training.

Out Of Scope

- Bias and Fairness
- AI Safety
- Model and Code Review
- Trustworthiness
- Performance Testing
- Red-Teaming
- Security and Privacy

Relevance to EU AI Act



Exceptions in the EU AI Act

Source: Orrick, 20 May 2024, "The EU AI Act: Open-Source Exceptions and Considerations for Your AI Strategy", https://www.orrick.com/en/Insights/2024/05/The-EU-AI-Act-Open-Source-Exceptions-and-Considerations-for-Your-AI-Strategy

AI systems in general

- Article 2(12): **"Does not apply to AI systems released under free and open-source licenses, unless they are placed on the market or put into service as high-risk AI systems or as an AI system that falls under Article 5 or 50",** i.e. AI systems that are prohibited or interact with natural persons

- Whether a AI system falls into a category of prohibited or high-risk AI systems identified in Arts 5-6 (Annexes I and III) will be determined by Commission.

MOF Acceptable Licenses

COMPONENT	DOMAIN	CONTENT TYPE	ACCEPTED OPEN LICENSE
Datasets	Data	Data	Preferred: CDLA-Permissive-2.0, CC-BY-4.0 Acceptable: Any including unlicensed
Data Preprocessing Code	Data	Code	Acceptable: OSI-approved
Model Architecture	Model	Code	Acceptable: OSI-approved
Model Parameters	Model	Data	Preferred: CDLA-Permissive-2.0 Acceptable: OSI-Approved, Permissive Open Data Licenses
Model Metadata	Model	Data	Preferred: CDLA-Permissive-2.0 Acceptable: CC-BY-4.0, Permissive Open Data Licenses
Training Code	Model	Code	Acceptable: OSI-approved
Inference Code	Model	Code	Acceptable: OSI-approved
Evaluation Code	Model	Code	Acceptable: OSI-approved
Evaluation Data	Model	Data	Preferred: CDLA-Permissive-2.0 Acceptable: CC-BY-4.0, Permissive Open Data Licenses
Evaluation Results	Model	Documentation	Preferred: CC-BY-4.0 Acceptable: Permissive Open Content Licenses
Supporting libraries and Tools	Model	Code	Acceptable: OSI-approved
Model Card	Model	Documentation	Preferred: CC-BY-4.0 Acceptable: Permissive Open Content Licenses
Data Card	Data	Documentation	Preferred: CC-BY-4.0 Acceptable: Permissive Open Content Licenses
Technical Report	Model & Data	Documentation	Preferred: CC-BY-4.0 Acceptable: Permissive Open Content Licenses
Research Paper	Model & Data	Documentation	Preferred: CC-BY-4.0 Acceptable: Permissive Open Content Licenses
Sample Model Outputs	Model	Data or Code	Unlicensed

Exceptions in the EU AI Act

Source: Orrick, 20 May 2024, "The EU AI Act: Open-Source Exceptions and Considerations for Your AI Strategy", https://www.orrick.com/en/Insights/2024/05/The-EU-AI-Act-Open-Source-Exceptions-and-Considerations-for-Your-AI-Strategy

Limited exceptions for General-Purpose AI Models (GPAIM)

- **GPAIM Definition**: AI model that displays significant generality and is capable to competently perform a wide range of distinct tasks regardless of the way the model is placed on the market and that can be integrated into a variety of downstream systems or applications. This does not cover models that are used before release on the market for research, development & prototyping activities.

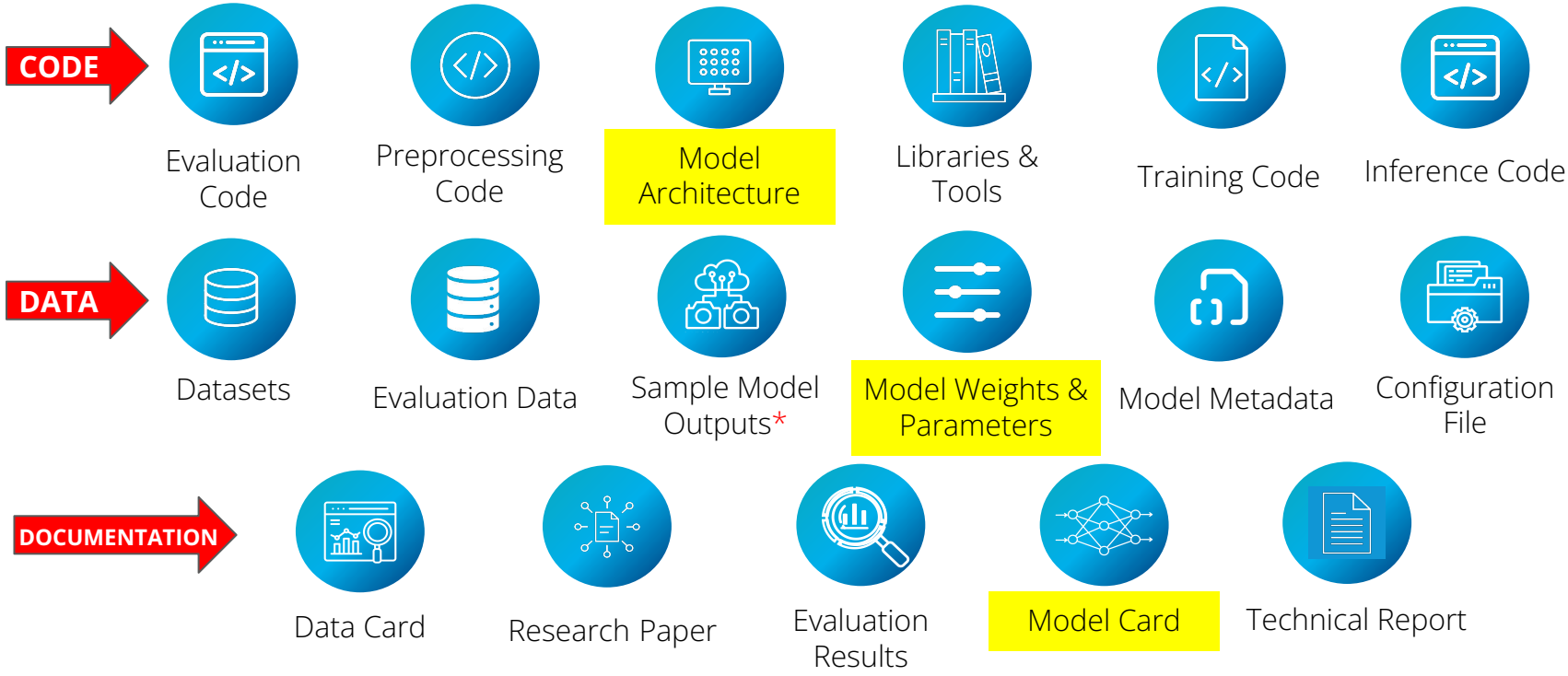
Exceptions in the EU AI Act

Source: Orrick, 20 May 2024, "The EU AI Act: Open-Source Exceptions and Considerations for Your AI Strategy", https://www.orrick.com/en/Insights/2024/05/The-EU-AI-Act-Open-Source-Exceptions-and-Considerations-for-Your-AI-Strategy

Limited exceptions for General-Purpose AI Models (GPAIM)

- To qualify, providers must enable **"access, usage, modification & distribution of the model…" where the model's parameters "…including weights, information on the model architecture, and information on model usage, are made publicly available."**
- **However, GPAIM won't qualify for the exception if they "present systemic risks",** i.e. if they have "high-impact capabilities" evaluated on the basis of technical tools and methodologies or otherwise designated as such by the Commission

MOF Components



*Can be code or data

Exceptions in the EU AI Act

Source: Orrick, 20 May 2024, "The EU AI Act: Open-Source Exceptions and Considerations for Your AI Strategy", https://www.orrick.com/en/Insights/2024/05/The-EU-AI-Act-Open-Source-Exceptions-and-Considerations-for-Your-AI-Strategy

Criteria for GPAIM exception

- Make the GPAIM available on conditions that satisfy the req. of Article 53(2)
- Account for systemic risks

Then, the **provider is exempt from transparency obligations in Article 53(1)(a) and (b),** including exempt from the obligations to create and maintain (i) up-to-date technical documentation and (ii) information intended for downstream providers seeking to integrate the GPAIM in their systems.

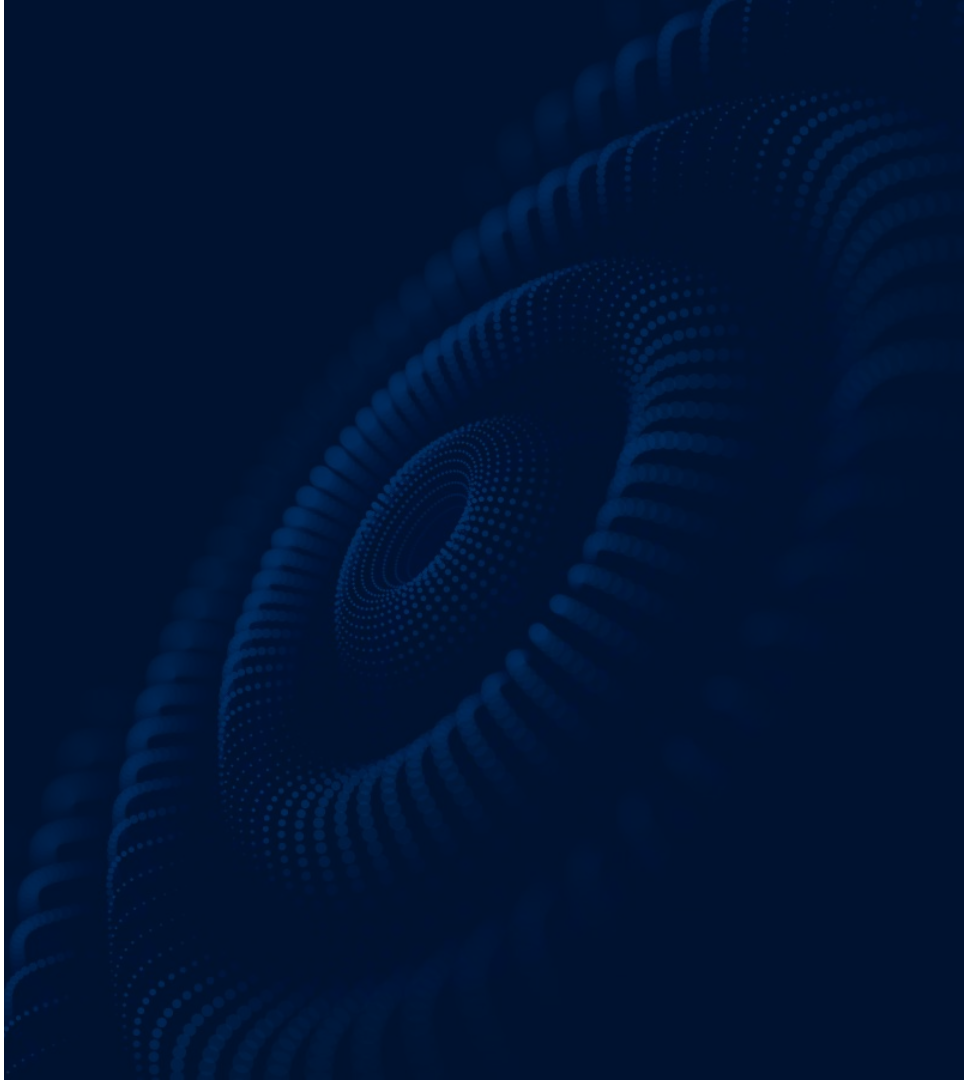
Exception in Practice

Source: Orrick, 20 May 2024, "The EU AI Act: Open-Source Exceptions and Considerations for Your AI Strategy", https://www.orrick.com/en/Insights/2024/05/The-EU-AI-Act-Open-Source-Exceptions-and-Considerations-for-Your-AI-Strategy

Nonetheless, the model provider of the open-source GPAIM must:

- Share, at distribution, a detailed summary of the content used for model training (with a level of specificity regulators will determine).
- Establish a policy with respect to EU copyright law, including for identifying and respecting a rightsholders reservation of rights though "the use of state of the art technologies" and pursuant to Article 4(3) of Directive (EU) 2019/790.

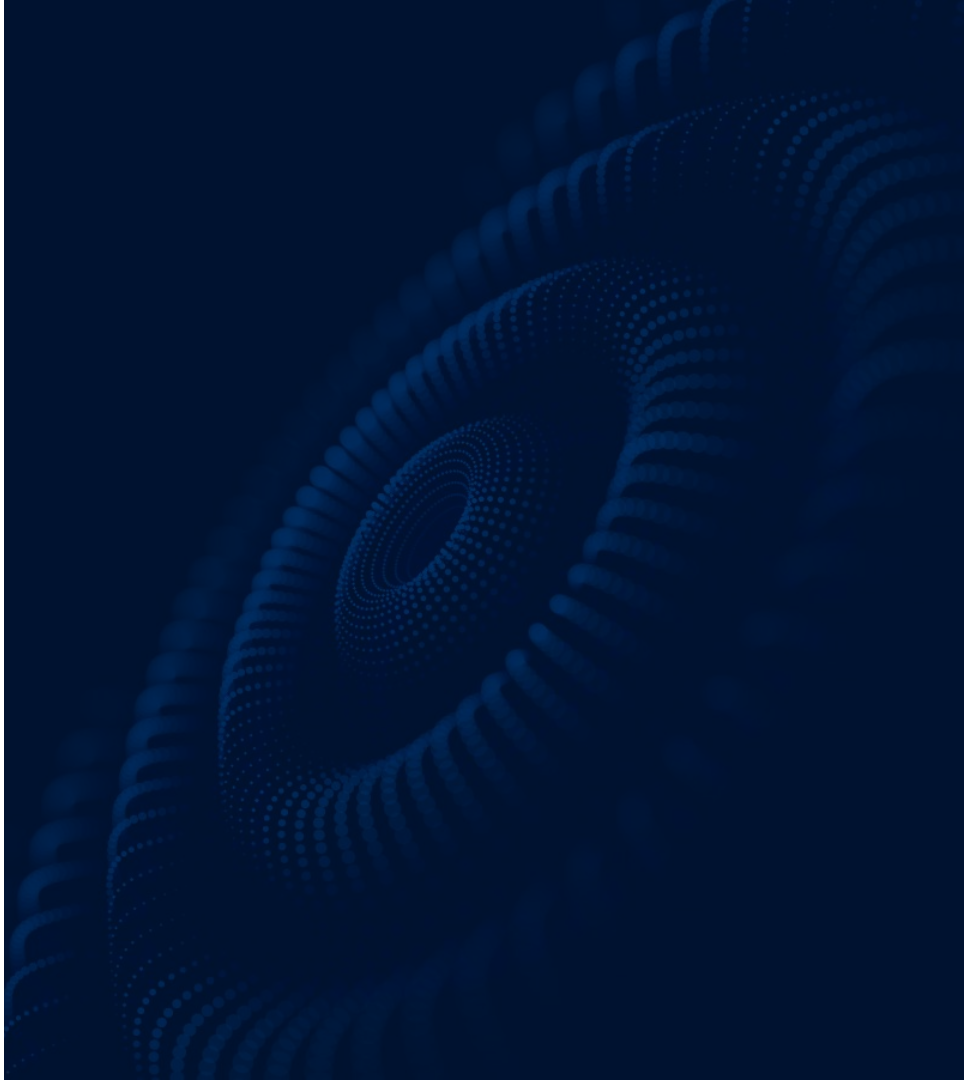
Conclusion



Conclusion

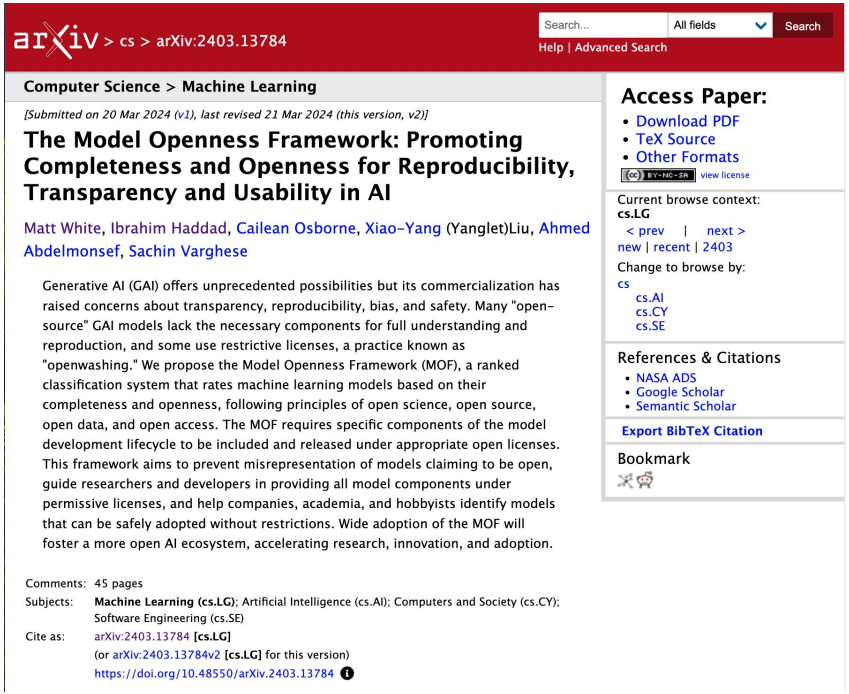
- **A clear and actionable methodology** to assess and enhance transparency in machine learning models. It provides a roadmap for model producers by specifying components, such as training data, code, models, and documentation, that should be openly released, promoting reproducible and ethical AI development.
- **Adopting open licenses**, as recommended by the framework, fosters collaboration, community engagement, and the freedom to use, modify, and distribute models and components within the license terms. The tiered classification system encourages the release of models with increasing levels of openness, promoting collective innovation and ensuring fairness, safety, and public oversight.
- Achieving this vision requires a **collaborative effort from all AI stakeholders**, including researchers, developers, institutions, companies, and governments, to embrace both completeness and openness as fundamental principles. The substantial benefits for science, business, and society justify the challenges of pursuing model transparency.
- With well-designed incentives, policies, and community norms, the ideals of open source and open science can become the standard in AI, rather than the exception. Through cross-domain collaboration, we can shape the progress of AI to be as open, ethical, and empowering as possible. The Model Openness Framework provides practical guidance for this journey towards **trustworthy and democratized artificial intelligence**.

Call to Action



Request for Comments

https://arxiv.org/abs/2403.13784



arXiv > cs > arXiv:2403.13784

Search... All fields Search

Help | Advanced Search

Computer Science > Machine Learning

[Submitted on 20 Mar 2024 (v1), last revised 21 Mar 2024 (this version, v2)]

The Model Openness Framework: Promoting Completeness and Openness for Reproducibility, Transparency and Usability in AI

Matt White, Ibrahim Haddad, Cailean Osborne, Xiao-Yang (Yanglet)Liu, Ahmed Abdelmonsef, Sachin Varghese

Generative AI (GAI) offers unprecedented possibilities but its commercialization has raised concerns about transparency, reproducibility, bias, and safety. Many "open-source" GAI models lack the necessary components for full understanding and reproduction, and some use restrictive licenses, a practice known as "openwashing." We propose the Model Openness Framework (MOF), a ranked classification system that rates machine learning models based on their completeness and openness, following principles of open science, open source, open data, and open access. The MOF requires specific components of the model development lifecycle to be included and released under appropriate open licenses. This framework aims to prevent misrepresentation of models claiming to be open, guide researchers and developers in providing all model components under permissive licenses, and help companies, academia, and hobbyists identify models that can be safely adopted without restrictions. Wide adoption of the MOF will foster a more open AI ecosystem, accelerating research, innovation, and adoption.

Comments: 45 pages

Subjects: **Machine Learning (cs.LG)**; Artificial Intelligence (cs.AI); Computers and Society (cs.CY); Software Engineering (cs.SE)

Cite as: arXiv:2403.13784 **[cs.LG]**

(or arXiv:2403.13784v2 **[cs.LG]** for this version)

https://doi.org/10.48550/arXiv.2403.13784

Access Paper:

- Download PDF
- TeX Source
- Other Formats

view license

Current browse context: cs.LG

< prev | next >

new | recent | 2403

Change to browse by:

cs

cs.AI

cs.CY

cs.SE

References & Citations

- NASA ADS
- Google Scholar
- Semantic Scholar

Export BibTeX Citation

Bookmark

Get Involved

Model Openness Tool (coming soon): isitopen.ai

Join the Generative AI Commons open community: genaicommons.org

Questions? Thank you! :)

𝕏 @cailean_osborne