GENERATIVE AI COMMONS

Generative AI Commons

Aug 21st, 2024

LF AI & DATA

Antitrust Policy Notice

- Linux Foundation meetings involve participation by industry competitors, and it is the intention of the Linux Foundation to conduct all of its activities in accordance with applicable antitrust and competition laws. It is therefore extremely important that attendees adhere to meeting agendas, and be aware of, and not participate in, any activities that are prohibited under applicable US state, federal or foreign antitrust and competition laws.

- Examples of types of actions that are prohibited at Linux Foundation meetings and in connection with Linux Foundation activities are described in the Linux Foundation Antitrust Policy available at http://www.linuxfoundation.org/antitrust-policy. If you have questions about these matters, please contact your company counsel, or if you are a member of the Linux Foundation, feel free to contact Andrew Updegrove of the firm of Gesmer Updegrove LLP, which provides legal counsel to the Linux Foundation.

Agenda

- Future Guest Speaker Update (1 minute)
- Follow-ups, updates, news (5 minutes)
- Mission statement (0 minute)
- Workstream Update (5 minutes each - 25 minutes)
  - Framework workstream
  - Model, Applications, and Data workstream
  - Education & Outreach workstream
  - Responsible AI workstream
- Open Discussion (remaining time)

Future Guest Speaker Update

(1 minute)

Future Guest Speaker Update

[Generative AI Commons Speaker Pipeline - Google Sheets]

No sales pitch; topics need to be Open Source/Open Science GenAI related.

Generative AI Commons Speaker Pipeline

File Edit View Insert Format Data Tools Extensions Help

A1 | Speaker Name

	A	B	C	D	E	F
1	**Speaker Name**	**Topic**	**Inviter**	**Has the speaker confirmed**	**Requested speaking date**	**GenAI Commons schedule status**
2	Nick Schifano	Fastcatalog.ai	Anni	Yes	July 24	confirmed
3	Yair Adato	Responsible Visual Generative AI	Ofer	Yes	October 2	confirmed
4	Frank Cappello	Evaluating a Generative AI Assistant for Science	Suparna	OK in principle, date TBD	Aug 21 (tentative)	
5	Assaf Elovic	GPT Research	Ofer			
6	Shayne Longpre	Responsible FM Cheatsheet and Data Provenance Initiative	Anni	Yes	Sep 4	Confirmed
7	Jaan Li, Kevin George	onefact.org	Anni	Yes	August 7	Confirmed
8	Michael Novac	Open Source AI policy update	Anni	Yes	August 7	Confirmed
9	R. Gab (Intel)	GenAI app landscape	Anni	Yes	September 4	Confirmed
10	Atai Barkai	Copilot Kit	Ofer			
11	Ariel Weinberger	Pezzo.ai	Ofer			
12	Christopher Burr	TEA Platform https://hackmd.io/@tea-platform/ryKcA7BP0	Ahmed	Yes	September 18	
13	Peter Madison	MLCommons	Anni			
14		SC42	Arnaud			
15	Chip Huyen	Voltron Data	Anni		Nov/Dec	

Follow-ups, Updates, News

Please Fill Out Survey

2024 Generative AI Survey

CLOUD NATIVE COMPUTING FOUNDATION | LF AI & DATA | THE LINUX FOUNDATION Research

This survey aims to understand the deployment, use, and challenges of generative AI technologies in organizations and the role of open source in this domain. It should take about 10 minutes to answer. Your participation is greatly appreciated.

https://www.research.net/r/QWTLLV7

Join Generative AI Commons!

Community Management (Mailing List, Calendar, Wiki, etc...);

Models, Applications, and Data: https://lists.lfaidata.foundation/g/gac-mad-workstream

Frameworks: https://lists.lfaidata.foundation/g/gac-frameworks-workstream

Education and Outreach: https://lists.lfaidata.foundation/g/gac-education-outreach

Responsible AI: https://lists.lfaidata.foundation/g/gac-responsible-ai-workstream

Slack Channels (LF AI & Data Slack): https://join.slack.com/t/lfaifoundation/shared_invite/zt-24yjiz9v7-lxkHxos4TBYf8WN_9E2R~g

Models, Applications, and Data: #gac-mad-workstream

Frameworks: #gac-frameworks-workstream

Education and Outreach: #gac-education-outreach-workstream

Responsible AI: #gac-responsible-ai-workstream

Joining Groups:

In order to join the groups.io group you will need to subscribe. Just send an email to any of the workstreams you intend to join.

Models, Applications, and Data: gac-mad-workstream+subscribe@lists.lfaidata.foundation

Frameworks: gac-frameworks-workstream+subscribe@lists.lfaidata.foundation

Education and Outreach: gac-education-outreach-workstream+subscribe@lists.lfaidata.foundation

Responsible AI: gac-responsible-ai-workstream+subscribe@lists.lfaidata.foundation

Meeting Time Summary

- Generative AI Commons bi-weekly meetings: Every other Wednesday 7 am PST
- Frameworks Workstream - Every other Tuesday 9:00 am PST
- Model, Applications, & Data Workstream - Every other Thursday 8:30 am PST (same as former Model & Data)
- ~~Applications Workstream - Every other Wednesday 8 am PST~~
- Edu & Outreach Workstream - Every other Wednesday 7 am PST
- Responsible AI Workstream - Every other Thursday 7 am PST

Note - A meeting reminder will be sent to all the workstream participants one day prior to the meeting

Marketing Updates

Content Submission Form

August

- **Blogs**
  - **Building Code Search Engine Using Open Source**
  - Combating Cyber Threats: The Power of Open Source
  - LF AI & Data TAC Talks: How OPEA Unlocks GenAI for Enterprises
  - Launching TAC Talks: Your Guide to LF AI & Data Projects
  - LF AI & Data Foundation Mid-Year Review: Significant Growth in the First Half of 2024
  - Introducing the MAD Workstream: Uniting Models, Applications & Data Under Generative AI Commons to Drive Innovation
  - **LF AI & Data Announces Monocle as its Latest Sandbox Project**
  - Part III of Addressing the Challenges of Open AI: Balancing Innovation with Regulation through the Model Openness Framework
- **White Paper**
  - The Alchemy of Intelligence: How Generative AI can revolutionize Business Intelligence and Analytics in Modern Enterprises
- **Social Content**
  - Promote GenAI 2024 Survey
- **Events**
  - Unity Catalog Fireside Chat - August 21
  - OPEA Demo-palooza - August 21
  - Unity Catalog Community Meeting - August 22, occur biweekly
  - AI for Developers Meeting: LF Edge and LF AI & Data - September 11-13
  - OpenLineage Meetup - September 12
  - The Role of Data in GenAI - September 12
  - OSS EU - Sept 16-18 in Vienna, Austria

AI_dev Conference

Mission Statement

GENERATIVE AI COMMONS

LF AI & DATA

Mission

We are dedicated to fostering the democratization, advancement and adoption of efficient, secure, reliable, and ethical Generative AI open source innovations through neutral governance, open and transparent collaboration and education.

Workstreams and Updates

Workstreams: How We Are Organized

** Needs Update:
Models, Data & Application Workstreams combined!!

Models, Data

- Hosting models and weights
- Hosting ethical datasets
- Data processing tools.
- Benchmarks and Reports
- Training and inference code

Applications

- Hosting AI Application frameworks
- Databases (vector stores)
- Agent frameworks
- Interface and metadata standards

Workstream Tracking Sheet: here

Frameworks

- Generative AI adoption
- Reference architectures
- Model Openness Framework
- Best practices and guidelines
- Compliance in code

Education & Outreach

- Education and training
- Thought leadership
- Educational outreach
- Legislative representation

Responsible AI

- Responsible AI
- Security, Privacy and Safety
- Informing Policy
- Copyright Issues
- Model and Data Lineage

MOF & MOT Status

- Initial MOF Draft Spec up for review
  - https://docs.google.com/document/d/1ZLRVO0C1ZUSAAHvuXysJz_I1bjO7C_cGmZrNa75vMqs/edit
  - Need agreement on whether to pursue or not

- MOT github repo available
  - https://github.com/lfai/model_openness_tool
  - Arnaud repo admin for now, can give others write access (i.e., Maintainer privileges)
  - Many issues, lots of work needed

- Professor Liu from the Columbia University working with his students on updating the MOT data
  - https://docs.google.com/spreadsheets/d/1VjHp4Jsueged3taI6sU-O2beNJwyv32Hfw9_2gc_b84/edit
  - Will probably just replace existing content with the updated spreadsheet

Frameworks Update

Ahmed, Alexy

(5 mins)

Frameworks – Update

- Meetings every other Tuesday at 10:00 PM EST.
  - Meeting URL: https://zoom-lfx.platform.linuxfoundation.org/meeting/91503008905?password=c2bcd124-eb36-4ff5-bf35-b001b1e9ad20
  - Next meeting August 27th.
- 1st deliverable: Model Openness Framework white paper updates:
  - A new section about MOT has been added to MOF whitepaper and under review.
  - MOF paper submitted for publish at ACM Journal on Responsible Computing.
  - Connection with Trustworthy and Ethical Assurance (TEA) Platform leads for potential opportunity for integration with MOF/MOT.
  - MOF Spec initial draft is ready and being reviewed by frameworks workstream members.
- 2nd deliverable: Enterprise Generative AI Reference Architecture updates:
  - Draft version is ready for framework workstream members to review

Models, Applications, and Data Update

Sachin Varghese, Raghavan Muthuregunathan, Nick Chase

(5 mins)

GENERATIVE AI COMMONS

LF AI & DATA

Model, Applications, and Data – Update

- We are working on a schedule of guest speakers - Your suggestions welcome!
- **R G Esteves** shared his work so far on the Gen AI applications Landscape in the meeting and follow up discussions
- Next meeting (August 29) we will have **Ke Ding** from OPEA team to discussion collaboration opportunities
- Discussion on next steps to improve workstream participation
- Action items:
  - Feedback on GenAI applications landscape document
  - Blog on building a **Code Search tool** using Open Source tools
- Next Meeting on August 29th

Education & Outreach Update

Ofer Hermoni & Raghavan Muthuregunathan

(5 mins)

Education & Outreach – Update

We are working on several objectives with target due dates:

- Planning our next quarterly webinar on "The role of Data in Gen AI" for September 12th
  - Host - Anni. Panelists: Lisa Cao, Datastrato, Denny Lee, Databricks, and Nick Schifano, Fastcatalog.ai
- Developers Community:
  - New blog post published: Developing GenAI Applications (RAG) Using Purely Open Source Technologies - by Raghavan Blog link
- General Public:
  - Outdoor workshop last week endorsed by LFAI ran by Jeremiah - linkedin post
  - Next Webinar - Create Your Own AI Image Model: A Hands-On Workshop for Artists and Educators - September 17th - Link
  - Next E&O Webinar - TBD
- Courses:
  - Building two new training courses on Gen AI and Risks in Conversational AI, targeted for August 2024 (on track)
- Next week we will have Nick Vidal to present the current status of the OSI definition
- Our workstream now has 58 team members contributing to the various initiatives!
- Our next meeting is scheduled for next Wednesday at 10am ET, 7 am pacific time

Responsible AI Update

Susan Malaika, Suparna Bhattacharya

(5 mins)

GENERATIVE AI COMMONS

LF AI & DATA

Responsible AI – Update for GAC on 2024-08-21

Meetings held every other Thursday after LF-AI TAC sessio:

- Leaders are: Suparna Bhattacharya, Andreas Fehlner, Susan Malaika Other Topics include : NIST AISIC Engagement |
- Working on Responsible (Generative) AI Framework Paper
  - https://docs.google.com/document/d/1CcKVLgtTpH2B7jH-Zq9GhE0GPchtlet9/edit?usp=sharing&ouid=111516505914 35635742&rtpof=true&sd=true
    - Co-authors may include Haluk Demirkan, Susan Malaika, suparna.bhattacharya@hpe.com; fehlner@arcor.de; Adel Zaalouk <azaalouk@redhat.com>; ronald.petty@rx-m.com; anni lai <annilai88@gmail.com>; 'Oita Coleman' <oita.coleman@openvoicenetwork.org>; Maureen McElaney <mmcelaney@us.ibm.com>; Karen Bennet <karen.bennet@gmail.com>
- Related Activities at: AIDev Hong Kong | Responsible AI Speaker List | Connections with ML Commons More here https://docs.google.com/document/d/1dR2zQEtZb5CiLPULpWS0UWH57x5GeiogAPK6D1x3Tg4/edit?usp=sharing
- Next Responsible AI meeting, on 2024-08-21 Thursday at 10am ET https://zoom-lfx.platform.linuxfoundation.org/meeting/91390993751?password=8a107f34-6986-409a-a101-b3ee093cb117

Responsible AI Materials https://drive.google.com/drive/folders/1GzUO_0jpM5Fsu2ft_6gu11a9ZarmD1Wg?usp=sharing

NIST Engagement

PI for LF-AI and Data engagement with AISIC @Matt White
PI for own company engagement with AISIC @Greg Lindahl

- wg #1 - Risk Management for Generative AI | @Susan Malaika |
- wg #2 - Synthetic Data - |@iot_mnovak|
- wg #3 - Capability Evaluations - | @Ali Hashmi |
- wg #4 - Red Teaming - | @Ofer Hermoni| @Vini Jaiswal |
- wg #5 - Safety & Security - | @Greg Lindahl (represents own company) | @Andreas Fehlner |

Outdated

What's Next?

Workstream Next Steps

Team meeting activities:

- Align your goals and activities with **Generative AI Commons Strategy** (approved on 3/6/24)
- Support **MOF GTM** (proposal approved on 3/20/2024)
- Publish schedule for your deliverables public release [here]
- (Make sure you have broken down objectives and deliverables into tasks [here])
- Make an impact!

Next Call

Wednesday September 11st, 2024 at 7:00 am PST

Tentative Agenda:

- Carry over items from meeting today
- Other topics or talks per members proposals (GAC-Chairs@lists.lfaidata.foundation)

Open Discussion

Past Important Slides for Record-keeping Purposes

Generative AI Commons Strategy Discussion

GENERATIVE AI COMMONS

LF AI & DATA

Near Term Goals & High Value Deliverables

by AI_Dev Paris (6/19-6/20)

Theme: Generative AI Commons is the champion for AI openness

1. Championing Openness in Generative AI
   a. Complete **Model Openness Framework (MOF)** proposal and peer reviews
   b. Build Model Evaluation tools and identify models to test
   c. Publish MOF and create a deck to promote MOF
2. Building Responsible AI
   a. Create and complete a **Responsible AI Framework (RAF)** proposal and peer
   b. Publish RAF and create a deck to promote RAF
3. Creating a Collaborative Platform to benefit Generative AI developers and ecosystem
   a. Publish the Generative AI Commons **Glossary**
   b. Create a Generative AI Commons Outreach deck (with value proposition)
   c. Create a Generative AI **Architecture** (draft)
   d. Create a Generative AI **Landscape** (draft)

Mid-Term Goals & High Value Deliverables

by AI_Dev HongKong (8/23)

Theme 1: Generative AI Commons is the champion for AI openness
Theme 2: The importance of community collaboration in building GenAI

1. Championing Openness in Generative AI
   a. Continue promoting **Model Openness Framework (MOF)** to increase adoption
   b. Build partnerships with Hugging Face, and/or other open source AI communities to increase MOF adoption
   c. Highlight these partnerships
2. Building Responsible AI
   a. Continue promoting **Responsible AI Framework (RAF)** to increase adoption
   b. Identify projects for RAF
   c. Highlight adoption stories
3. Creating a Collaborative Platform to benefit Generative AI developers and ecosystem
   a. Continue promoting Gen AI Commons and effectively articulate the value proposition to model producers, and users, AI developers and ecosystem
   b. Publish an initial draft of Generative AI **Architecture**
   c. Publish an initial draft of Generative AI **Landscape**
   d. Promote projects from all workstreams

Year-End Goals & High Value Deliverables

By AI_Dev Seattle (10/14-10/15), Tokyo (10/28-10/29)

Theme 1: Generative AI Commons is the champion for AI openness
Theme 2: The importance of community collaboration in building GenAI
Theme 3: Generative AI Commons is the community to expedite time to value of open source GenAI projects and ideas

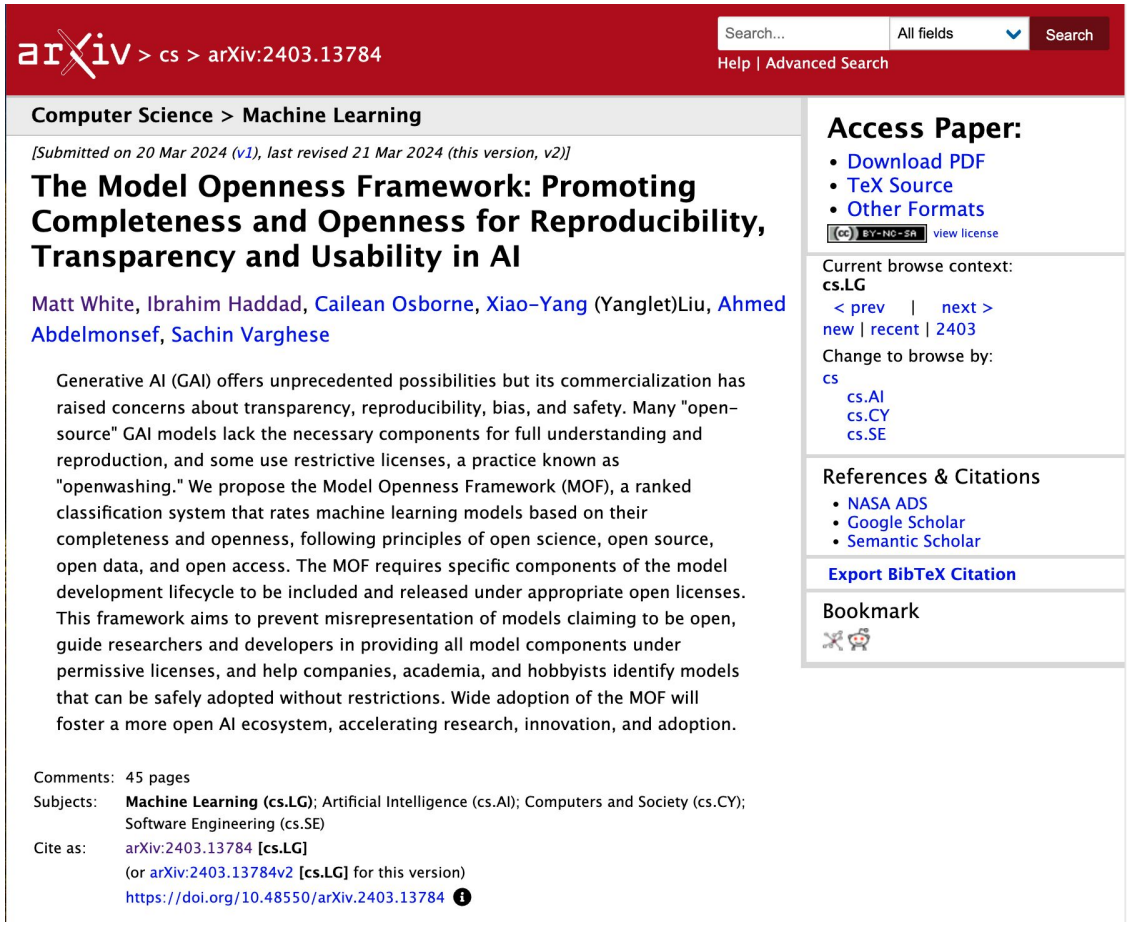
1. Championing Openness in Generative AI
   a. Continue promoting **Model Openness Framework (MOF)** to increase adoption
   b. Build/strengthen partnerships with Hugging Face, and/or other open source AI communities to increase MOF adoption
   c. Highlight more partnerships and adoption of MOF
2. Building Responsible AI
   a. Continue promoting **Responsible AI Framework (RAF)** to increase adoption
   b. Identify projects for RAF
   c. Highlight RAF project(s) and more adoption stories
3. Creating a Collaborative Platform to benefit Generative AI developers and ecosystem
   a. Continue promoting Gen AI Commons and effectively articulate the value proposition to model producers, and users, AI developers and ecosystem
   b. Publish and promote an approved Generative AI **Architecture**
   c. Publish and promote an approved Generative AI **Landscape**
   d. Promote projects from all workstreams

MOF Overview and Q&A

Model Openness Framework Go-to-Market Plan - Google Docs

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

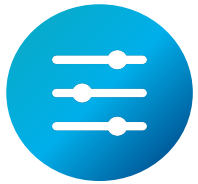
Bookmark

MOF Components


Datasets


Preprocessing Code


Model Architecture


Model Parameters


TVT Code


Inference Code


Evaluation Code


Evaluation Data


Evaluation Results


Technical Report


Model Metadata


Libraries and Tools


Model Card

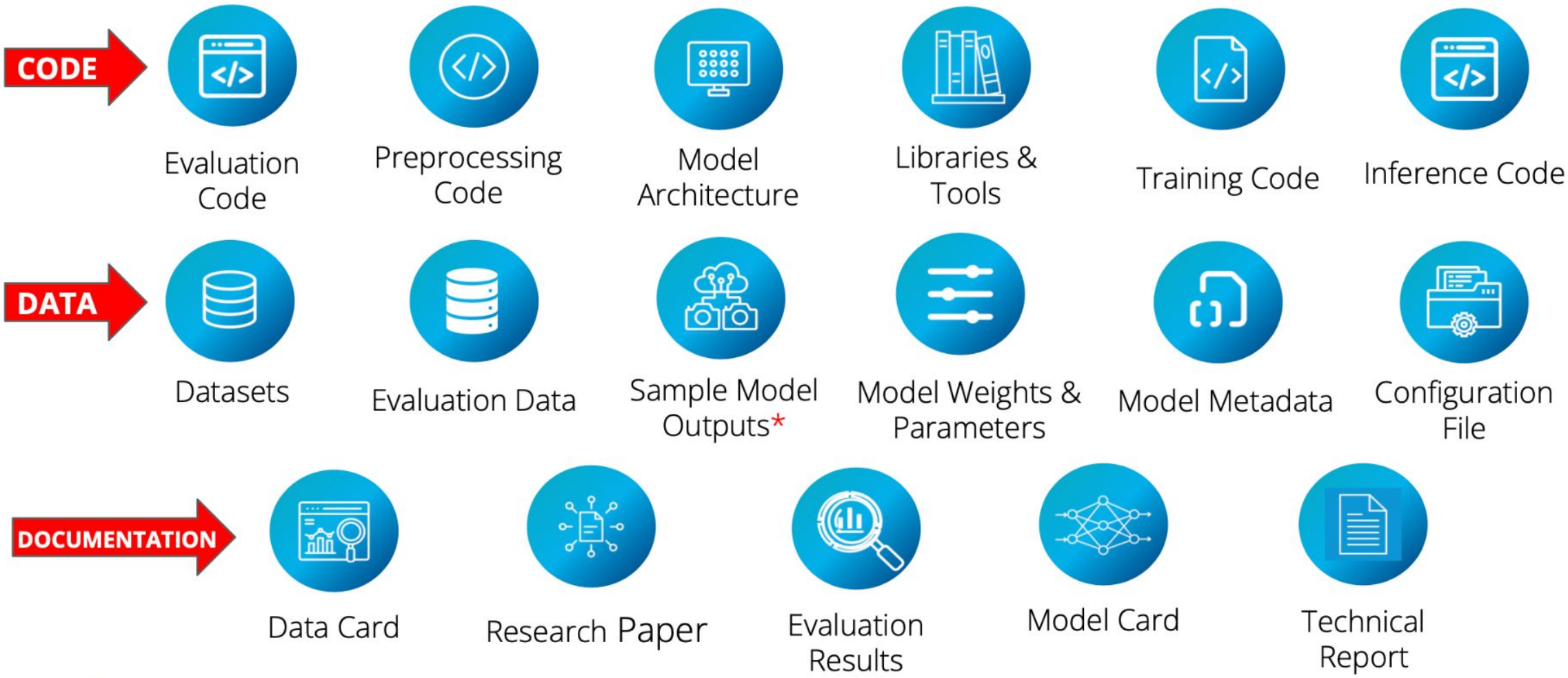

Data Card


Research


Sample Model Outputs


MOF File

MOF Components



GENERATIVE AI COMMONS

*Can be code or data

LF AI & DATA

MOF Acceptable Licenses

COMPONENT	DOMAIN	CONTENT TYPE	ACCEPTED OPEN LICENSE
Datasets	Data	Data	CDLA-Permissive-2.0, CC-BY-4.0 or equivalent permissive license (current version allows for unlicensed as well)
Data Preprocessing Code	Data	Code	OSI-approved
Model Architecture	Model	Code	OSI-approved
Model Parameters	Model	Data	CDLA-Permissive-2.0, CC-BY-4.0 or equivalent permissive license
Model Metadata	Model	Data	CDLA-Permissive-2.0, CC-BY-4.0 or equivalent permissive license
Training Code	Model	Code	OSI-approved
Inference Code	Model	Code	OSI-approved
Evaluation Code	Model	Code	OSI-approved
Evaluation Data	Model	Data	CDLA-Permissive-2.0,CC-BY-4.0 or equivalent permissive license
Evaluation Results	Model	Documentation	CC-BY-4.0, CC0 or equivalent permissive license
Supporting libraries and Tools	Model	Code	OSI-approved
Model Card	Model	Documentation	CC-BY-4.0, CC0 or equivalent permissive license
Data Card	Data	Documentation	CC-BY-4.0, CC0 or equivalent permissive license
Technical Report	Model & Data	Documentation	CC-BY-4.0, CC0 or equivalent permissive license
Research Paper	Model & Data	Documentation	CC-BY-4.0, CC0 or equivalent permissive license
Sample Model Outputs	Model	Data or Code	OSI-Approved for Code, CDLA-Permissive-2.0 for Data, CC-BY-4.0, CC0 or equivalent permissive license for content

MOF Classes

MOF Class	Components Included
Class III – Open Model	• Model Architecture • Model Parameters and Metadata (Final Checkpoints and Optimizer State) • Technical Report • Evaluation Results • Model Card • Data Card

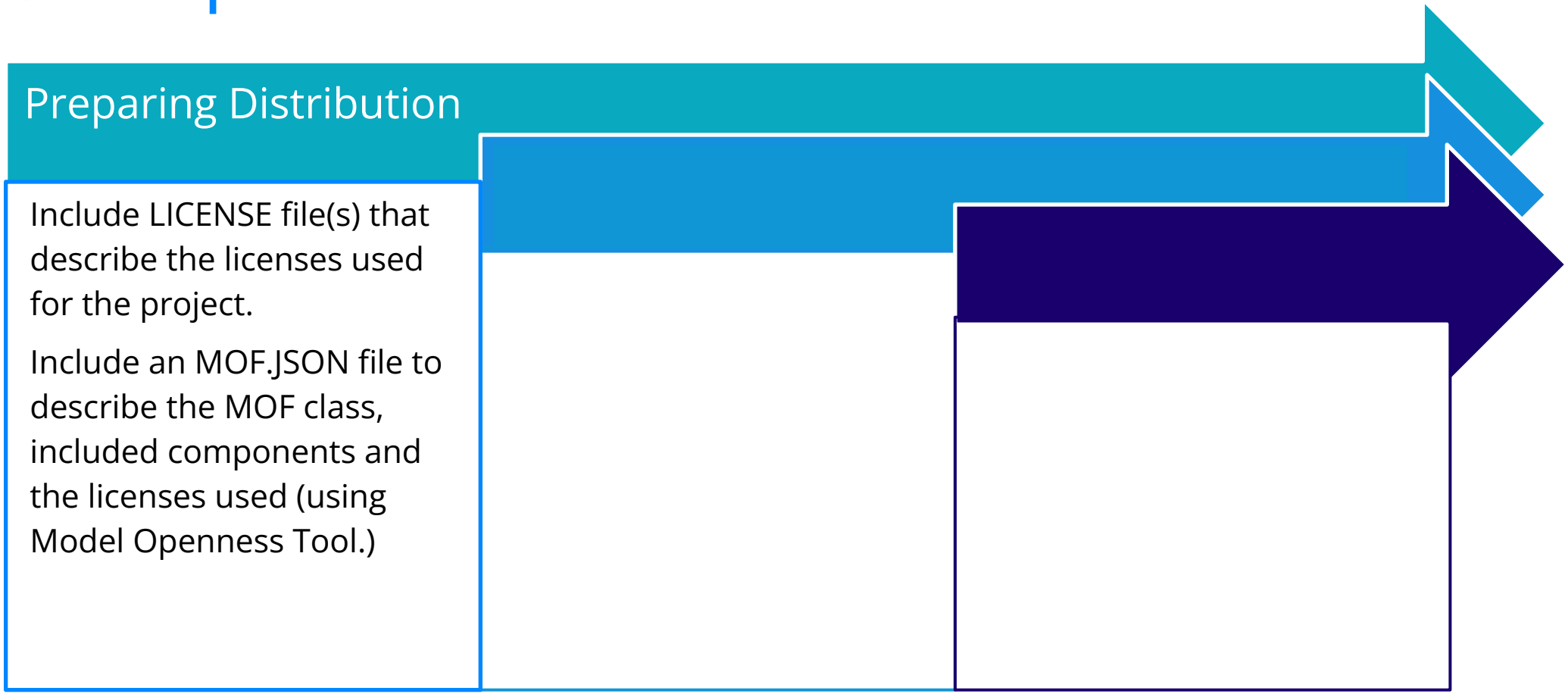
MOF Classes

MOF Class	Components Included
Class II – Open Tooling	• Training Code • Inference Code • Evaluation Code • Evaluation Data • Supporting Libraries & Tools • + All Class III Components
Class III – Open Model	• Model Architecture • Model Parameters and Metadata (Final Checkpoints and Optimizer State) • Technical Report • Evaluation Results • Model Card • Data Card

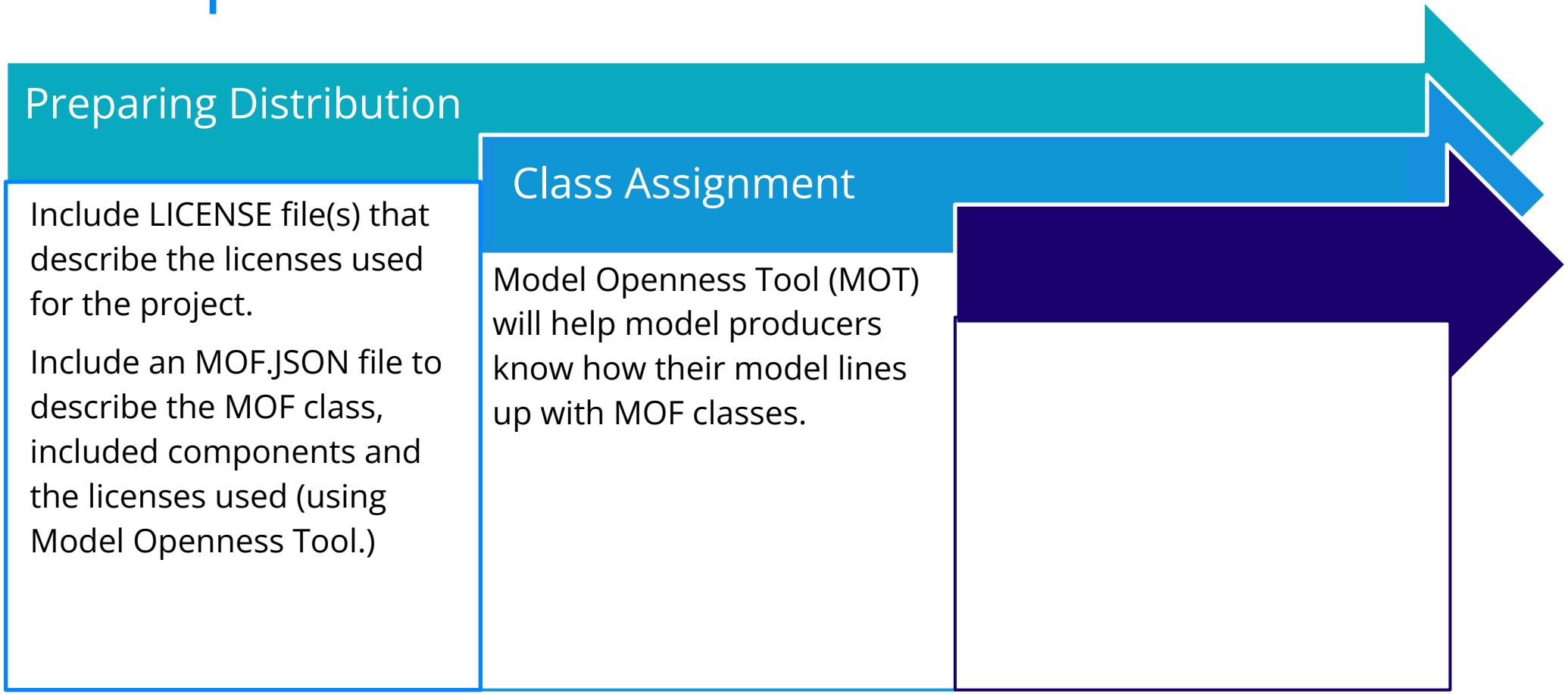
MOF Classes

MOF Class	Components Included
Class I – Open Science	• Research Paper • Datasets (can be unlicensed) • Data Preprocessing Code • Model Parameters and Metadata (Intermediate Checkpoints and Optimizer States) • + All Class II Components
Class II – Open Tooling	• Training Code • Inference Code • Evaluation Code • Evaluation Data • Supporting Libraries & Tools • + All Class III Components
Class III – Open Model	• Model Architecture • Model Parameters and Metadata (Final Checkpoints and Optimizer State) • Technical Report • Evaluation Results • Model Card • Data Card

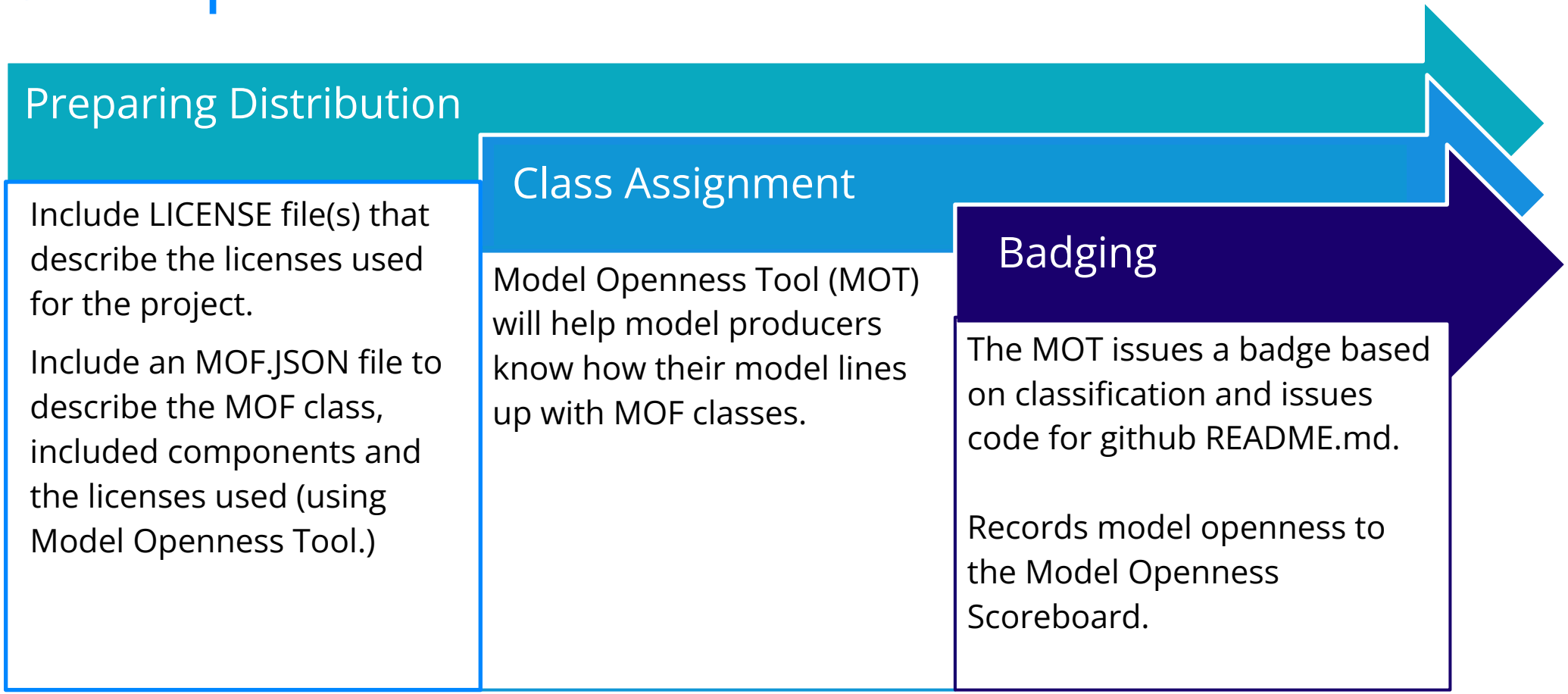
MOF Implementation



MOF Implementation



MOF Implementation



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.

Next Steps for MOF & MOT

(Free Discussion - 15 minutes)

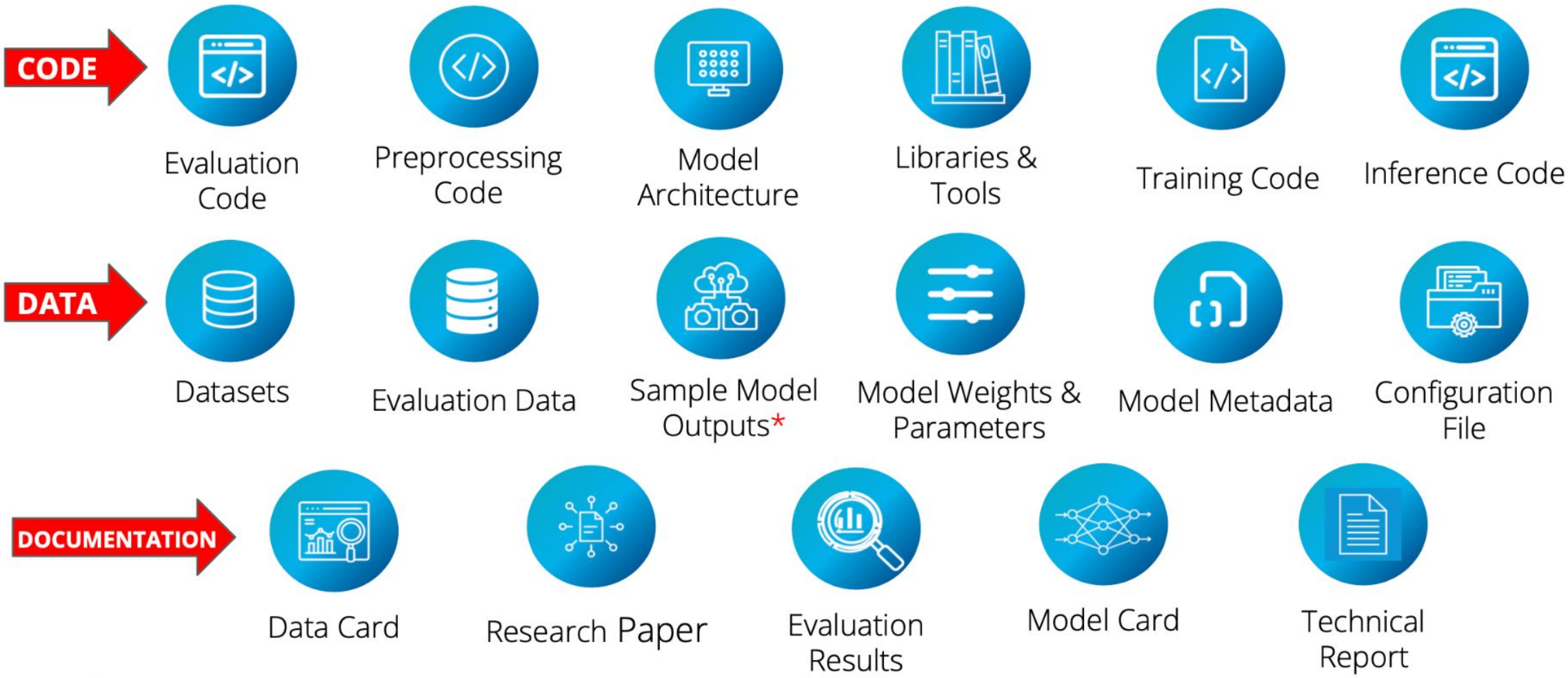
Proposal: Next Step for MOF & MOT

1. MOF 1.0 for GA (Framework workstream leads) – due 10/1
   a. Feedback collection and revision of the MOF
   b. Produce an official "LF AI & Data MOF 1.0" paper and presentation for GA
2. MOT 1.0 for GA (Framework workstream leads) – due 10/1
   a. Feedback collection and revision on the MOT
   b. Produce an online README/Instruction manual for MOT 1.0
3. Promote MOF 1.0 and MOT 1.0 in the model producer community, so the model producers will self-disclose their licenses on MOT/isitopen.ai (Model & Data workstream leads) - Ongoing
4. Identify 5+ AI influencers to support MOF 1.0 & MOT 1.0 (Outreach workstream leads) – due 10/1
5. MOF 1.0 & MOT 1.0 official launch/GA release at AI_Dev Tokyo 10/28-10/29 (Marketing leads) – due 10/1

MOT Call to Action - We need your help!

1. Documentation: We need help for this. Signup to help write a paragraph or a few sentences in this page. Instructions are at the top. Thanks! https://docs.google.com/document/d/12xmyw-Fh9x5rkBUGSF_I2QzbQ5PVd25QFNVqSGjBChI/edit
2. Models: Signup to do the research for models.
   a. Instructions are here https://docs.google.com/document/d/1qHtaweze1bvZnnVin_xcXKr-XMghCwwK_kz2zzL8u8M/edit?usp=drive_link
   b. list of models to validate are here: https://docs.google.com/spreadsheets/d/1yUpxTszIjdPscfaIKMqXKN2P9-cOpDS4IfV6aTk38WA/edit?usp=drive_link

MOF Components



GENERATIVE AI COMMONS

*Can be code or data

LF AI & DATA

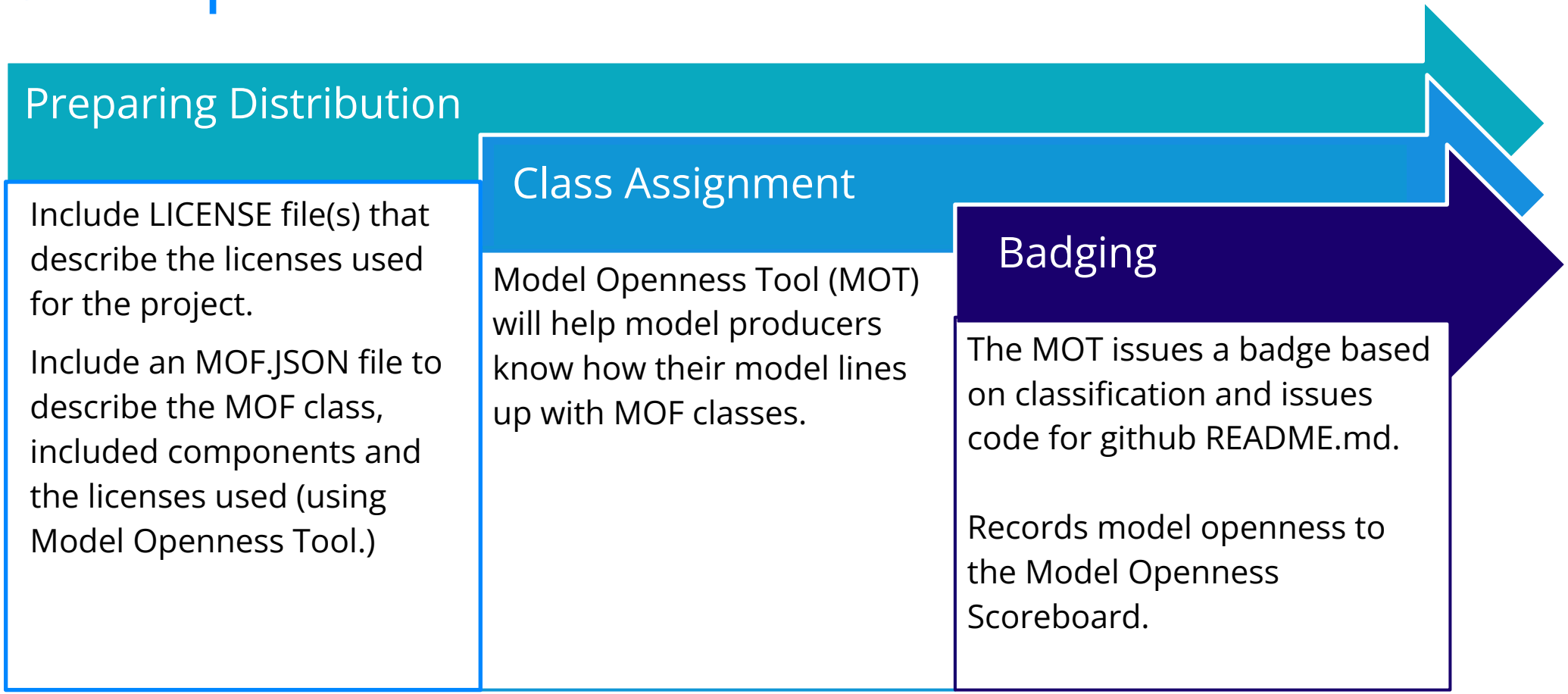
MOF Acceptable Licenses

COMPONENT	DOMAIN	CONTENT TYPE	ACCEPTED OPEN LICENSE
Datasets	Data	Data	CDLA-Permissive-2.0, CC-BY-4.0 or equivalent permissive license (current version allows for unlicensed as well)
Data Preprocessing Code	Data	Code	OSI-approved
Model Architecture	Model	Code	OSI-approved
Model Parameters	Model	Data	CDLA-Permissive-2.0, CC-BY-4.0 or equivalent permissive license
Model Metadata	Model	Data	CDLA-Permissive-2.0, CC-BY-4.0 or equivalent permissive license
Training Code	Model	Code	OSI-approved
Inference Code	Model	Code	OSI-approved
Evaluation Code	Model	Code	OSI-approved
Evaluation Data	Model	Data	CDLA-Permissive-2.0,CC-BY-4.0 or equivalent permissive license
Evaluation Results	Model	Documentation	CC-BY-4.0, CC0 or equivalent permissive license
Supporting libraries and Tools	Model	Code	OSI-approved
Model Card	Model	Documentation	CC-BY-4.0, CC0 or equivalent permissive license
Data Card	Data	Documentation	CC-BY-4.0, CC0 or equivalent permissive license
Technical Report	Model & Data	Documentation	CC-BY-4.0, CC0 or equivalent permissive license
Research Paper	Model & Data	Documentation	CC-BY-4.0, CC0 or equivalent permissive license
Sample Model Outputs	Model	Data or Code	OSI-Approved for Code, CDLA-Permissive-2.0 for Data, CC-BY-4.0, CC0 or equivalent permissive license for content

MOF Classes

MOF Class	Components Included
Class I – Open Science	• Research Paper • Datasets (can be unlicensed) • Data Preprocessing Code • Model Parameters and Metadata (Intermediate Checkpoints and Optimizer States) • + All Class II Components
Class II – Open Tooling	• Training Code • Inference Code • Evaluation Code • Evaluation Data • Supporting Libraries & Tools • + All Class III Components
Class III – Open Model	• Model Architecture • Model Parameters and Metadata (Final Checkpoints and Optimizer State) • Technical Report • Evaluation Results • Model Card • Data Card

MOF Implementation



Next Steps for MOF & MOT

(Open Discussion - 15 minutes)

MOF & MOT Update

1. MOT spreadsheet (300 model spreadsheet) – Nick
2. MOT source code – Arnaud
3. MOF spec process discussion – Arnaud
4. Open items:
    a. MOF/MOT Documentation – need volunteers – Framework workstream chairs: Ahmed and Alexy
    b. Invite Model producers to the Model and Data workstream meetings and be MOF/MOT influencers. – Model workstream chairs – Nick and Sri
    c. Is AI_Dev Tokyo, 10/28-29 still a reasonable timeframe for MOF & MOT GA?

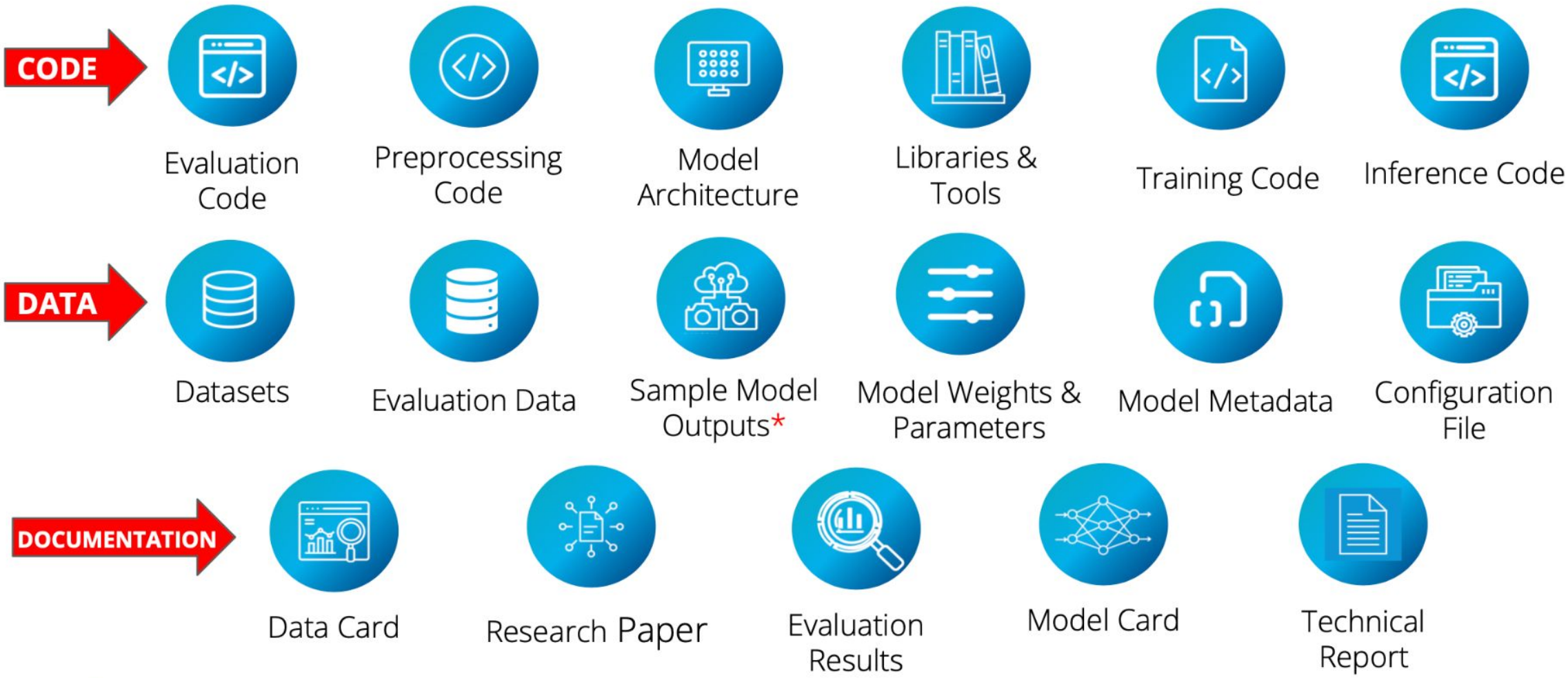
Proposal: Next Step for MOF & MOT (proposal)

1. MOF specification Working Draft
   a. Put a first draft together – focuses on the what rather than the why.
   b. Text still needs tightening but feedback welcome. Will need to decide whether this is something we want to pursue.
2. MOF 1.0 for GA (Framework workstream leads) – proposed due date 10/1
   a. Feedback collection and revision of the MOF
   b. Produce an official "LF AI & Data MOF 1.0" paper and presentation for GA
3. MOT spec 1.0 for GA (Framework workstream leads) – proposed due date 10/1
   a. Feedback collection and revision on the MOT
   b. Produce an online README/Instruction manual for MOT 1.0
4. Promote MOF 1.0 and MOT 1.0 in the model producer community, so the model producers will self-disclose their licenses on MOT/isitopen.ai (Model & Data workstream leads) - Ongoing
5. Identify 5+ AI influencers to support MOF 1.0 & MOT 1.0 (Outreach workstream leads) – due 10/1
6. MOF 1.0 & MOT 1.0 official launch/GA release at AI_Dev Tokyo 10/28-10/29 (Marketing leads) – due 10/1

MOT Call to Action - We need your help!

1. Documentation: We need help for this. Signup to help write a paragraph or a few sentences in this page. Instructions are at the top. Thanks! https://docs.google.com/document/d/12xmyw-Fh9x5rkBUGSF_I2QzbQ5PVd25QFNVqSGjBChI/edit
2. Models: Signup to do the research for models.
   a. Instructions are here https://docs.google.com/document/d/1qHtaweze1bvZnnVin_xcXKr-XMghCwwK_kz2zzL8u8M/edit?usp=drive_link
   b. list of models to validate are here: https://docs.google.com/spreadsheets/d/1yUpxTszIjdPscfaIKMqXKN2P9-cOpDS4IfV6aTk38WA/edit?usp=drive_link

MOF Components



GENERATIVE AI COMMONS

*Can be code or data

LF AI & DATA

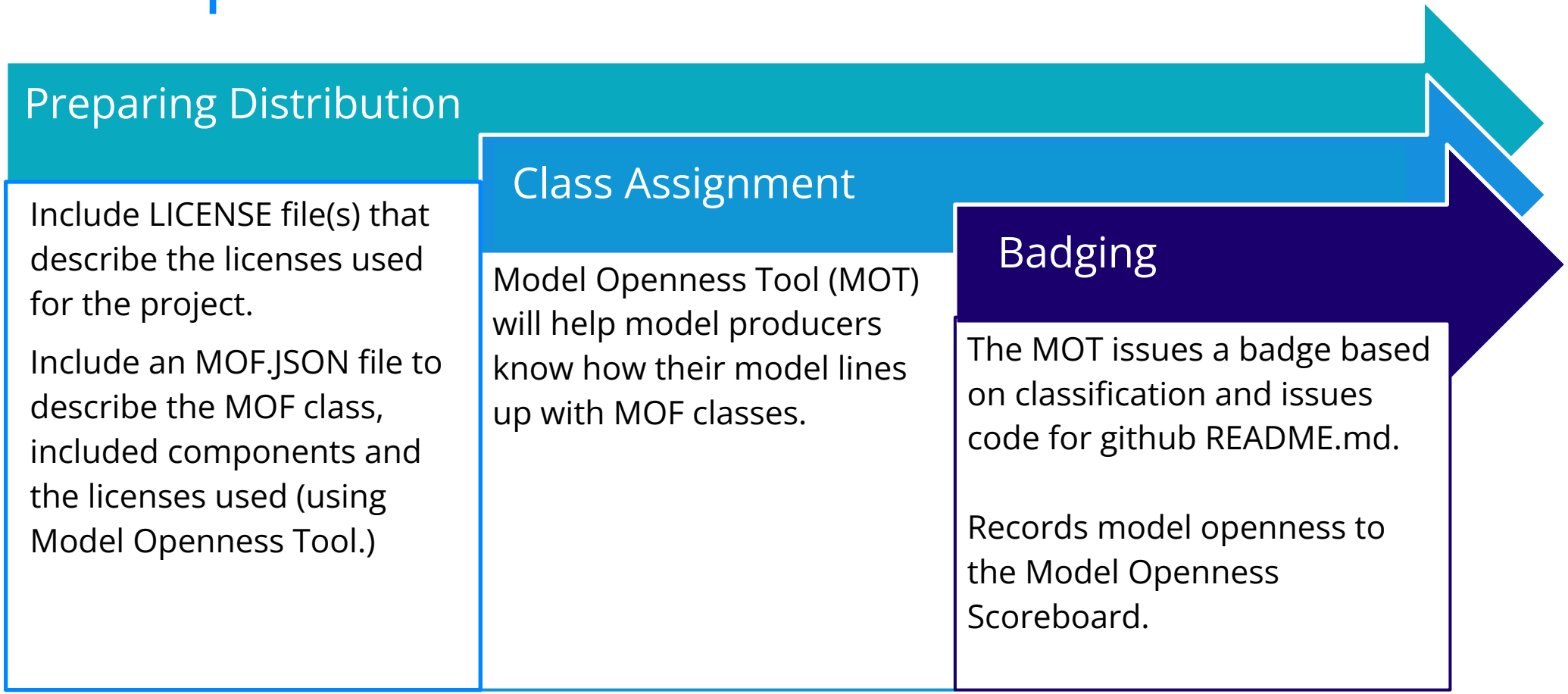
MOF Acceptable Licenses

COMPONENT	DOMAIN	CONTENT TYPE	ACCEPTED OPEN LICENSE
Datasets	Data	Data	CDLA-Permissive-2.0, CC-BY-4.0 or equivalent permissive license (current version allows for unlicensed as well)
Data Preprocessing Code	Data	Code	OSI-approved
Model Architecture	Model	Code	OSI-approved
Model Parameters	Model	Data	CDLA-Permissive-2.0, CC-BY-4.0 or equivalent permissive license
Model Metadata	Model	Data	CDLA-Permissive-2.0, CC-BY-4.0 or equivalent permissive license
Training Code	Model	Code	OSI-approved
Inference Code	Model	Code	OSI-approved
Evaluation Code	Model	Code	OSI-approved
Evaluation Data	Model	Data	CDLA-Permissive-2.0,CC-BY-4.0 or equivalent permissive license
Evaluation Results	Model	Documentation	CC-BY-4.0, CC0 or equivalent permissive license
Supporting libraries and Tools	Model	Code	OSI-approved
Model Card	Model	Documentation	CC-BY-4.0, CC0 or equivalent permissive license
Data Card	Data	Documentation	CC-BY-4.0, CC0 or equivalent permissive license
Technical Report	Model & Data	Documentation	CC-BY-4.0, CC0 or equivalent permissive license
Research Paper	Model & Data	Documentation	CC-BY-4.0, CC0 or equivalent permissive license
Sample Model Outputs	Model	Data or Code	OSI-Approved for Code, CDLA-Permissive-2.0 for Data, CC-BY-4.0, CC0 or equivalent permissive license for content

MOF Classes

MOF Class	Components Included
Class I – Open Science	• Research Paper • Datasets (can be unlicensed) • Data Preprocessing Code • Model Parameters and Metadata (Intermediate Checkpoints and Optimizer States) • + All Class II Components
Class II – Open Tooling	• Training Code • Inference Code • Evaluation Code • Evaluation Data • Supporting Libraries & Tools • + All Class III Components
Class III – Open Model	• Model Architecture • Model Parameters and Metadata (Final Checkpoints and Optimizer State) • Technical Report • Evaluation Results • Model Card • Data Card

MOF Implementation



MOT Notes

Models | Model Openness Tool (isitopen.ai)

Instructions: https://docs.google.com/document/d/1qHtaweze1bvZnnVin_xcXKr-XMghCwwK_kz2zzL8u8M/edit

Spreadsheet: https://docs.google.com/spreadsheets/d/1yUpxTszIjdPscfaIKMqXKN2P9-cOpDS4IfV6aTk38WA/edit#gid=1034334420 (edited)

MOF & MOT plan

Soft launch at AI_Dev Europe (June), official launch at OSS Europe (Sept)

Volunteers needed:

1. MOT testing – 2 people (Farshad, Alberto and Raghavan all have admin access and were asked to beat the tires)
2. MOT documentation – 2 people (Nick chase + ??). * this is critical
3. MOF & MOF website update – Marketing (Jen Shelby is lead)
4. Update the spreadsheet, the most critical piece! (2-4 people)

Note: The Framework workstream leads will lead the entire effort with Nancy's help.

Generative AI Commons Landscape Proposal

(Free Discussion - 1 minutes)

Generative AI Commons Landscape Proposal

A landscape produced by each of the following workstream:

1. Model & Data workstream
2. Application workstream
3. Responsible AI workstream

Note: Here are all the ones that this tool builds for lots of other Foundation:
https://github.com/cncf/landscape2-sites#landscape2-sites