Carnegie
Mellon
University
Libraries

Open Forum for Artificial Intelligence (OFAI)

December 10, 2024

CNI Fall 2024 Membership Meeting

Sayeed Choudhury – @sayeedc.bsky.social
Associate Dean for Digital Infrastructure
Director of Open Source Programs Office (OSPO)
Executive Director of Open Forum for AI (OFAI)

Open Forum for AI (OFAI)

Carnegie Mellon University (CMU) is launching the Open Forum for AI (OFAI) to provide holistic, objective expert advice regarding AI, particularly as it relates to research, technical prototypes, policy recommendations, community engagement, and inclusion. The overarching vision is that AI should augment human capability rather than replace it; that AI development needs to be human centered. Leveraging CMU's AI expertise and capacity, OFAI will include partners that offer a range of complementary expertise necessary for collective action for AI strategy and associated policy development.

Contact

Sayeed Choudhury
sayeedc@cmu.edu
Associate Dean for Digital Infrastructure, University Libraries
Director, Open Source Programs Office

Founding Reason

OFAI's launch signals a powerful call for deeper technical understanding and more rigorous scrutiny in AI, including but also extending beyond open source software. OFAI will advocate in an objective manner for wider involvement in the AI ecosystem to shape policy-making and address issues such as AI safety, privacy, and equity, as underscored in memoranda like the White House AI Executive Order. There is growing consensus that an open AI ecosystem will mitigate risks while fostering innovation and inclusion, but little clarity about the characteristics of such an ecosystem. OFAI aims to address this gap in a comprehensive and inclusive manner.

ENGIN in the News

- **CMU Launches New Initiative for Human-Centered AI** [library.cmu.edu]

- **Can Artificial Intelligence be Open Sourced?** [cacm.acm.org]

- **CMU OSAID Statement**

Carnegie Mellon University

OFAI (Current) Partners

Georgia Tech Open Source Program Office

GW Open Source Program Office

The University of Texas at Austin — Open Source Program Office

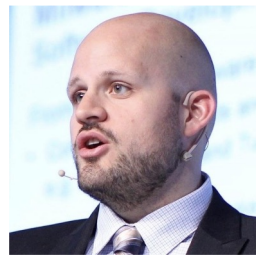open source initiative®

Atlantic Council

creative commons

SCIENCE CONS

Carnegie Mellon University

OFAI Fellows

Open Forum for AI (OFAI) Leadership

Sayeed Choudhury	Hoda Heidari	Vince Sha	Deb Bryant	Trey Herr	Rachel Dzombak
Director	Research	Technical Prototype	Community Engagement	Policy	Talent for Service
Library	Block Center	Dietrich	OSI	Atlantic Council	Heinz

The "Open" in OFAI

- Value of openness as a design principle – fostering innovation, participation, etc. for internet and web (e.g., TCP/IP, HTTP)

- Benefits of open source software – $8.8 trillion in value, 3.5x reduction in production costs

- Questioning and testing assumptions – focus on evidence and research

- Produce and release outputs in more rapid, open manner (e.g., preprints, policy briefs)

Openness in AI Framework

- Technical framework, developed with community engagement, that rigorously defines openness and the ensuing or corresponding value

- Move from open source to open governance (e.g., transparency, reproducibility, accessibility)

- Need to consider specific context and use case with emphasis on the Ws of Why, What, Who, Whom, When, and Where

What's Open Source AI?

Following the same idea behind Open Source Software,
an Open Source AI is a system made available under terms that grant users the freedoms to:

Use   Study   Modify   Share

Use the system for any purpose and without having to ask for permission.

Precondition to exercise these freedoms is to have access to
the preferred form to make modifications to the system, and to the means to use it.

Carnegie Mellon University

Preferred form to make modifications to machine-learning systems

The preferred form of making modifications to a machine-learning system must include all the elements below:

- **Data Information:** Sufficiently detailed information about the data used to train the system so that a skilled person can build a substantially equivalent system. Data Information shall be made available under OSI-approved terms.

  - In particular, this must include: (1) the complete description of all data used for training, including (if used) of unshareable data, disclosing the provenance of the data, its scope and characteristics, how the data was obtained and selected, the labeling procedures, and data processing and filtering methodologies; (2) a listing of all publicly available training data and where to obtain it; and (3) a listing of all training data obtainable from third parties and where to obtain it, including for fee.

- **Code:** The complete source code used to train and run the system. The Code shall represent the full specification of how the data was processed and filtered, and how the training was done. Code shall be made available

The growing focus on open source AI makes it all the more important that we establish a shared understanding of what open source AI is. A definition should outline what must be shared and under what terms or conditions. Without this clarity, we risk a fragmented approach, where companies label their products as "open source" even when they aren't, where civil society doesn't have access to the AI components they need for testing and accountability, and policymakers create regulations that fail to address the complexities of the issue.

The Open Source Initiative (OSI) has recently released a new draft definition of open source AI, marking a critical juncture in the evolution of the internet. This moment comes after two years of conversations, debates, engagements, and late-night conversations across the technical and open source communities. It is critical not just for redefining what "open source" means in the context of AI; it's about shaping the future of the technology and its impact on society.

OFAI Landscape Map

	Research	Technical Prototypes	Community Engagement	Policy Recommendations	Talent for Service
Data					
Design & Architecture					
Code					
Compute					
Model Weights					
Human Feedback					
Intended vs Unintended Usage					
Expertise, Education & Training					
Oversight					

Carnegie Mellon University

Technical Prototype Demo

Resources

OFAI website
Can Artificial Intelligence be open sourced?
Open Source AI Definition (OSAID)
CMU Faculty statement of support for OSAID
George Washington OSPO endorsement of OSAID
Georgia Tech endorsement of OSAID
Mozilla statement of support for OSAID
OFAI landscape map

Acknowledgements and Q&A

Carnegie Mellon University Libraries

Omidyar Network™

NobleReach™ Foundation