Openness and AI Open Source - Open Weights Open Data - Open Research



An Al Generated Overview of Al and Open Source - - - (WARNING - Hallucinations are possible)

Microsoft Copilot (ChatGPT 4) - 11/11/2024 ... Al and Open Source

(Prompt - Please provide an overview of openness and AI – in particular AI and open source, AI and open weights, AI and open data, and AI and open research. Please make your response roughly 800-1000 words in length and appropriate for a college audience.)

Openness and AI: An Overview

The concept of openness in AI encompasses various dimensions, including open source, open weights, open data, and open research. Each of these aspects plays a crucial role in fostering innovation, collaboration, and transparency in the development and deployment of AI technologies. This overview explores these dimensions and their implications for the AI landscape.

Al and Open Source

Open source AI refers to the practice of making AI software and tools freely available for anyone to use, modify, and distribute. This approach has several benefits:

- 1. Collaboration and Innovation: Open source AI fosters a collaborative environment where developers from around the world can contribute to and improve upon existing AI technologies. This collective effort accelerates innovation and leads to the development of more robust and versatile AI systems.
- 2. Accessibility: By making AI tools and frameworks open source, developers lower the barriers to entry for individuals and organizations that may not have the resources to develop AI technologies from scratch. This democratization of AI enables a broader range of participants to engage with and benefit from AI advancements.
- 3. Transparency and Trust: Open source AI promotes transparency, as the underlying code is available for scrutiny. This openness helps build trust in AI systems, as users can verify the functionality and fairness of the algorithms.

However, open source AI also presents challenges, such as the potential for misuse by malicious actors and the need for sustainable funding models to support ongoing development and maintenance.

AI and Open Weights

Open weights refer to the practice of sharing the trained parameters (weights) of AI models. This practice has significant implications for the AI community:

- 1. Reproducibility: Sharing model weights enhances the reproducibility of AI research. Researchers can validate and build upon existing models, ensuring that findings are robust and reliable.
- 2. Efficiency: By providing access to pre-trained models, open weights save researchers and developers the time and computational resources required to train models from scratch. This efficiency accelerates the development and deployment of AI applications.
- 3. Collaboration: Open weights facilitate collaboration across institutions and industries. Researchers can leverage pre-trained models to explore new applications and improve existing ones, fostering a more interconnected and innovative AI ecosystem.

Despite these benefits, open weights also raise concerns about intellectual property and the potential for misuse in applications that may harm society.

Al and Open Data

Open data involves making datasets freely available for use in AI research and development. The availability of open data is crucial for training and evaluating AI models:

- 1. Diversity and Inclusivity: Open data initiatives can help address biases in AI models by providing diverse and representative datasets. This inclusivity ensures that AI systems are fairer and more equitable.
- 2. Innovation: Access to high-quality open data enables researchers to develop and test new AI algorithms and applications. This access drives innovation and leads to the creation of more effective and efficient AI solutions.
- 3. Transparency and Accountability: Open data promotes transparency in AI research, as datasets used to train models are available for examination. This openness allows for greater accountability and helps identify and mitigate potential biases and errors in AI systems.

However, open data also poses challenges, such as ensuring data privacy and security, and addressing the ethical implications of using certain types of data.

Al and Open Research

Open research in AI refers to the practice of sharing research findings, methodologies, and tools with the broader community. This openness is essential for advancing the field of AI:

- 1. Knowledge Sharing: Open research facilitates the dissemination of knowledge and best practices across the AI community. Researchers can learn from each other's work, avoid duplicating efforts, and build on existing findings to push the boundaries of AI technology.
- 2. Collaboration: By making research outputs openly available, researchers can collaborate more effectively across institutions and disciplines. This collaboration leads to interdisciplinary innovations and a more holistic understanding of AI.
- 3. Transparency and Trust: Open research promotes transparency in the scientific process, as methodologies and results are available for peer review and replication. This transparency builds trust in AI research and ensures that findings are credible and reliable.

Despite its benefits, open research also faces challenges, such as the need for sustainable funding models and the potential for intellectual property disputes.

Openness in AI, encompassing open source, open weights, open data, and open research, plays a vital role in fostering innovation, collaboration, and transparency in the field. While each dimension of openness offers significant benefits, it also presents challenges that need to be addressed to ensure the responsible and ethical development of AI technologies. By embracing openness, the AI community can create a more inclusive, equitable, and trustworthy AI ecosystem that benefits society as a whole.

Human Views on Al and Open Source - - - (WARNING - Humans are fallible)

Shorter Podcasts and Videos

Can Meta Build Open Source AGI?

Al for Humans - Gavin Purcell and Kevin Pereira (0.59) (January 29, 2024) [Video] https://www.youtube.com/watch?v=y4WJI6vRh5o

Has Open Source Inadvertently Screwed a Whole Generation of AI Startups?

Al Daily Brief - Nathaniel Whittemore (11:16)

(August 2, 2023) [Podcast/Video]

https://www.youtube.com/watch?v=ypBg0Cov8kc

Inside the AI Divide: Open Source vs. Closed Source Debates

A Beginner's Guide to AI (15:45)

(May 2, 2024) [Podcast]

https://www.listennotes.com/podcasts/a-beginners-guide/inside-the-ai-divide-open-NaKGn4boryZ/

Is Open Source AI Dangerous?

Al Daily Brief - Nathaniel Whittemore (12:20)

(July 23, 2023) [Podcast/Video]

https://www.youtube.com/watch?v= QHAvYamTJI

Moshi - An open source version of GPT-4o?

Last Week in AI - Episode 173 - Andrey Kurenkov and Jeremie Harris (1:00:45 - 1:04:07)

(July 7, 2024)[Video]

https://www.youtube.com/watch?v=gZGrA52A834

Open Source AI Part 1: Why is it important? |

Which is more dangerous for humanity: Open Source or Closed Source AI?

The Generative AI Podcast - Tony Wan (16:29)

(June 3, 2024) [Podcast]

https://podcasts.apple.com/ie/podcast/open-source-ai-part-1-why-is-it-important-which-is/id166080132 0?i=1000657654371

Open Source Al Part 2 | Meta Llama | Mistral | How open is open source?

The Generative AI Podcast - Tony Wan (13:28)

(June 10, 2024) [Podcast]

https://podcasts.apple.com/ie/podcast/open-source-ai-part-2-meta-llama-mistral-how-open-is/id16608 01320?i=1000658439176

Open Sourcing Al

Al for Humans - Gavin Purcell and Kevin Pereira (0.59)

(March 14, 2024) [Video]

https://www.youtube.com/watch?v=F0 ECx4L3Lw&t=854s

Open Sourcing AI: Accelerating Progress or Opening Pandora's Box?

A Beginner's Guide to AI (17:00) (October 8, 2023) [Podcast]

https://podcasts.apple.com/us/podcast/open-sourcing-ai-accelerating-progress-or-opening-pandoras/id 1701165010?i=1000630614562

Who will make AlphaFold3 open source? Scientists race to crack AI model

Last Week in AI - Episode 169 - Andrey Kurenkov and Jeremie Harris (1:00:45 - 1:04:07) (June 2, 2024)[Video]

https://www.youtube.com/watch?v=arqK GAvLp0&t=3645s

Why we need open-source Al

TEDAI - Heather Meeker (8:43) (October, 2023) [Video]

https://www.ted.com/talks/heather meeker why we need open source ai

Yann Lecun: Meta AI, Open Source, Limits of LLMs, AGI & the Future of AI

Lex Fridman Podcast - Episode #416 - Lex Fridman and Yann Lecun (1:43:48 - 1:47:48) (March 7, 2024) [Video]

https://www.youtube.com/watch?v=5t1vTLU7s40&t=6180s

Zuckerberg Wants to Build Open Source AGI

Al Daily Brief - Nathaniel Whittemore (12:28) (January 19, 2024) [Podcast/Video]

https://www.youtube.com/watch?v=xuNY0OYyIn0

Longer Podcasts and Videos

DisTrO and the Quest for Community-Trained AI Models

ai + a16z - Anjey Midha, Bowen Peng, and Jeffrey Quesnelle, Jeffrey (39:00) (September 27, 2024) [Podcast]

https://podcasts.apple.com/us/podcast/distro-and-the-quest-for-community-trained-ai-models/id1740178076?i=1000670993997

The Future of Open Source AI Development (Hugging Face)

Eye on AI - Craig Smith and Jeff Boudier, J. (50:00)

(November 13, 2024) [Podcast]

https://podcasts.apple.com/us/podcast/218-jeff-boudier-the-future-of-open-source/id1438378439?i=10 00676817026

How Meta wants to shape our digital future with open source AI

The TED AI Show - Bilawal Sidhu and Ragavan Srinivasan (50:00)

(October 15, 2024) [Podcast]

https://podcasts.apple.com/us/podcast/the-ted-ai-show/id1741574582

How Mozilla Is Shaping the Future of Open-Source AI

Eye on AI - Craig Smith and Mark Surman (47:00)

(October 12, 2024) [Podcast]

https://podcasts.apple.com/us/podcast/213-mark-surman-how-mozilla-is-shaping-the-future/id1438378 439?i=1000672917434

The Promise and Peril of Open Source AI with Elizabeth Seger and Jeffrey Ladish

Your Undivided Attention - The Center for Humane Technology (38:44)

(November 21, 2023) [Podcast]

https://www.humanetech.com/podcast/the-promise-and-peril-of-open-source-ai-with-elizabeth-seger-and-jeffrey-ladish

Scholarly Journal, Preprint and Conference Articles

Eiras, F., Petrov, A., Vidgen, B., Schroeder, C., Pizzati, F., Elkins, K., ... & Foerster, J. (2024). **Risks and opportunities of open-source generative AI**. *arXiv* preprint arXiv:2405.08597. https://arxiv.org/abs/2405.08597

Abstract: Applications of Generative AI (Gen AI) are expected to revolutionize a number of different areas, ranging from science & medicine to education. The potential for these seismic changes has triggered a lively debate about the potential risks of the technology, and resulted in calls for tighter regulation, in particular from some of the major tech companies who are leading in AI development. This regulation is likely to put at risk the budding field of open-source generative AI. Using a three-stage framework for Gen AI development (near, mid and long-term), we analyze the risks and opportunities of open-source generative AI models with similar capabilities to the ones currently available (near to mid-term) and with greater capabilities (long-term). We argue that, overall, the benefits of open-source Gen AI outweigh its risks. As such, we encourage the open sourcing of models, training and evaluation data, and provide a set of recommendations and best practices for managing risks associated with open-source generative AI.

Gao, H., Zahedi, M., Treude, C., Rosenstock, S., & Cheong, M. (2024, October). **Documenting ethical** considerations in open source ai models. In *Proceedings of the 18th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement* (pp. 177-188). https://dl.acm.org/doi/abs/10.1145/3674805.3686679

Abstract: The development of Al-enabled software heavily depends on Al model documentation, such as model cards, due to different domain expertise between software engineers and model developers. From an ethical standpoint, AI model documentation conveys critical information on ethical considerations along with mitigation strategies for downstream developers to ensure the delivery of ethically compliant software. However, knowledge on such documentation practice remains scarce. The objective of our study is to investigate how developers document ethical aspects of open source AI models in practice, aiming at providing recommendations for future documentation endeavours. We selected three sources of documentation on GitHub and Hugging Face, and developed a keyword set to identify ethics-related documents systematically. After filtering an initial set of 2,347 documents, we identified 265 relevant ones and performed thematic analysis to derive the themes of ethical considerations. Six themes emerge, with the three largest ones being model behavioural risks, model use cases, and model risk mitigation. Conclusions: Our findings reveal that open source AI model documentation focuses on articulating ethical problem statements and use case restrictions. We further provide suggestions to various stakeholders for improving documentation practice regarding ethical considerations.

Osborne, C., Ding, J., & Kirk, H. R. (2024). The AI community building the future? A quantitative analysis of development activity on Hugging Face Hub. *Journal of Computational Social Science*, 1-39. https://link.springer.com/article/10.1007/s42001-024-00300-8

Abstract: Open model developers have emerged as key actors in the political economy of artificial intelligence (AI), but we still have a limited understanding of collaborative practices in the open AI ecosystem. This paper responds to this gap with a three-part quantitative analysis of development activity on the Hugging Face (HF) Hub, a popular platform for building, sharing, and demonstrating models. First, various types of activity across 348,181 model, 65,761 dataset, and 156,642 space repositories exhibit right-skewed distributions. Activity is extremely imbalanced between repositories; for example, over 70% of models have 0 downloads, while 1% account for 99% of downloads. Furthermore, licenses matter: there are statistically significant differences in collaboration patterns in model repositories with permissive, restrictive, and no licenses. Second, we analyse a snapshot of the social network structure of collaboration in model repositories, finding that the community has a core-periphery structure, with a core of prolific developers and a majority of isolate developers (89%). Upon removing these isolates from the network, collaboration is characterised by high reciprocity regardless of developers' network positions. Third, we examine model adoption through the lens of model usage in spaces, finding that a minority of models, developed by a handful of companies, are widely used on the HF Hub. Overall, the findings show that various types of activity across the HF Hub are characterised by Pareto distributions, congruent with open source software development patterns on platforms like GitHub. We conclude with recommendations for researchers, and practitioners to advance our understanding of open AI development.

Quillivic, R., & Mesmoudi, S. (2024). **Is open source software culture enough to make AI a common?** *arXiv* preprint arXiv:2403.12774. https://arxiv.org/abs/2403.12774

Abstract: Language models (LM or LLM) are increasingly deployed in the field of artificial intelligence (AI) and its applications, but the question arises as to whether they can be a common resource managed and maintained by a community of users. Indeed, the dominance of private companies with exclusive access to massive data and language processing resources can create inequalities and biases in LM, as well as obstacles to innovation for those who do not have the same resources necessary for their implementation. In this contribution, we examine the concept of the commons and its relevance for thinking about LM. We highlight the potential benefits of treating the data and resources needed to create LMs as commons, including increased accessibility, equity, and transparency in the development and use of AI technologies. Finally, we present a case study centered on the Hugging Face platform, an open-source platform for deep learning designed to encourage collaboration and sharing among AI designers.

Shrestha, Y. R., von Krogh, G., & Feuerriegel, S. (2023). **Building open-source AI.** *Nature Computational Science*, *3*(11), 908-911. https://papers.ssrn.com/sol3/Delivery.cfm?abstractid=4614280

Abstract: The computer science community has a long tradition of embracing open-source principles. However, companies increasingly restrict access to AI innovations. An example is OpenAI, which was founded to make scientific research openly available but which eventually restricted access to research findings. Although such a strategy reflects a company's legitimate incentive to obtain financial returns, such protection increases concentration of power, restricting access to AI technology. Further down the road, concentrated power could lead to growing inequality in AI research, education and public use. Here we discuss why proprietary AI technology should be complemented by open-source AI across the essential components for building AI technology: datasets, source codes and models.

Widder, D. G., Whittaker, M., & West, S. M. (2024). Why 'open' Al systems are actually closed, and why this matters. *Nature*, *635*(8040), 827-833.

https://www.nature.com/articles/s41586-024-08141-1.pdf

Abstract: This paper examines 'open' artificial intelligence (AI). Claims about 'open' AI often lack precision, frequently eliding scrutiny of substantial industry concentration in large-scale AI development and deployment, and often incorrectly applying understandings of 'open' imported from free and open-source software to AI systems. At present, powerful actors are seeking to shape policy using claims that 'open' AI is either beneficial to innovation and democracy, on the one hand, or detrimental to safety, on the other. When policy is being shaped, definitions matter. To add clarity to this debate, we examine the basis for claims of openness in AI, and offer a material analysis of what AI is and what 'openness' in AI can and cannot provide: examining models, data, labour, frameworks, and computational power. We highlight three main affordances of 'open' AI, namely transparency, reusability, and extensibility, and we observe that maximally 'open' AI allows some forms of oversight and experimentation on top of existing models. However, we find that openness alone does not perturb the concentration of power in AI. Just as many traditional open-source software projects were co-opted in various ways by large technology companies, we show how rhetoric around 'open' AI is frequently wielded in ways that exacerbate rather than reduce concentration of power in the AI sector.

Policy Papers

Azoulay, P., Krieger, J. L., & Nagaraj, A. (2024). **Old moats for new models: Openness, control, and competition in generative AI** (No. w32474). *National Bureau of Economic Research*. https://www.nber.org/papers/w32474

Abstract: Drawing insights from the field of innovation economics, we discuss the likely competitive environment shaping generative AI advances. Central to our analysis are the concepts of appropriability—whether firms in the industry are able to control the knowledge generated by their innovations—and complementary assets—whether effective entry requires access to specialized infrastructure and capabilities to which incumbent firms can ration access. While the rapid improvements in AI foundation models promise transformative impacts across broad sectors of the economy, we argue that tight control over complementary assets will likely result in a concentrated market structure, as in past episodes of technological upheaval. We suggest the likely paths through which incumbent firms may restrict entry, confining newcomers to subordinate roles and stifling broad sectoral innovation. We conclude with speculations regarding how this oligopolistic future might be averted. Policy interventions aimed at fractionalizing or facilitating shared access to complementary assets might help preserve competition and incentives for extending the generative AI frontier. Ironically, the best hopes for a vibrant open source AI ecosystem might rest on the presence of a "rogue" technology giant, who might choose openness and engagement with smaller firms as a strategic weapon wielded against other incumbents.

Luna, Angela. (2024, September 3). **Open-source AI: The debate that could redefine AI innovation**. *American Action Forum*.

https://www.americanactionforum.org/insight/open-source-ai-the-debate-that-could-redefine-ai-innovation/

Abstract: A major debate around artificial intelligence (AI) is whether to promote closed or open-source AI; proponents of open-source — AI systems with their components available for further study, use, modification, and sharing — argue they promote AI development, greater transparency, and reduce market concentration, while opponents argue they pose safety risks and put some AI companies at a competitive disadvantage.

Meanwhile, the California legislature recently passed a bill that, in an attempt to mitigate potential safety risks from advanced AI models, would require large AI models' developers to implement strict safety and security protocols and assume liability of open-source models and their derivatives – though such requirements might discourage AI developers from opening their models.

As Congress considers the debate between open or closed AI, lawmakers should be aware of the

potential pitfalls of premature regulations that could hinder the development of open-source AI models and AI development more generally.

Roy-Chowdhury, Rahul. (2023, December 22). **Why open-source is crucial for responsible AI development**. *World Economic Forum*.

https://www.weforum.org/stories/2023/12/ai-regulation-open-source/

Abstract: At the heart of regulating AI is the need for a unified approach to responsible AI development and deployment. A wide range of technologists must be involved in regulatory conversations to ensure that regulation is developed without bias. Without varied viewpoints, resulting regulations may overlook open-source models as a cornerstone of responsible AI.

Websites

Aschenbrenner, Leopold. (2024, June). Lock Down the Labs: Security for AGI. Situational Awareness: The Decade Ahead. https://situational-awareness.ai/lock-down-the-labs/

The nation's leading AI labs treat security as an afterthought. Currently, they're basically handing the key secrets for AGI to the CCP on a silver platter. Securing the AGI secrets and weights against the state-actor threat will be an immense effort, and we're not on track. ... The United States has an advantage in the AGI race. But we will give up this lead if we don't get serious about security very soon. Getting on this, now, is maybe even the single most important thing we need to do today to ensure AGI goes well.

Zuckerberg, Mark. (2024, July 23). **Open source AI is the path forward.** Meta. https://about.fb.com/news/2024/07/open-source-ai-is-the-path-forward/

"Today we're taking the next steps towards open source AI becoming the industry standard. We're releasing Llama 3.1 405B, the first frontier-level open source AI model, as well as new and improved Llama 3.1 70B and 8B models. In addition to having significantly better cost/performance relative to closed models, the fact that the 405B model is open will make it the best choice for fine-tuning and distilling smaller models."