Building open-source AI

Article in Nature Computational Science · October 2023
DOI: 10.1038/s43588-023-00540-0

Yash Raj Shrestha
University of Lausanne

Building Open-Source AI

— Comment —
Preprint Manuscript

Yash Raj Shrestha[1], Georg von Krogh[2,3], Stefan Feuerriegel[4,5]

[1] Faculty of Business and Economics (HEC), University of Lausanne, Lausanne, Switzerland
[2] Department of Management, Technology, and Economics, ETH Zurich, Zurich, Switzerland
[3] ETH AI Center, ETH Zurich, Zurich, Switzerland
[4] Institute of AI in Management, LMU Munich
[5] Munich Center for Machine Learning, Munich Germany

Correspondence: Stefan Feuerriegel (feuerriegel@lmu.de)

Original link: https://rdcu.be/dpyod

Cite as: Shrestha, Y. R., von Krogh, G., Feuerriegel, S. (2023) Building Open-Source AI, Accepted for Publication at ***Nature Computational Science***

Standfirst

Artificial intelligence drives innovation across society, economy, and science. We argue for the importance of building artificial intelligence technology with open-source principles to foster accessibility, collaboration, responsibility, and interoperability.

Main

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 once founded to make scientific research openly available but has eventually restricted access to research findings. While such a strategy reflects a company's legitimate incentives to appropriate returns from research, restrictive protection increases the concentration of power over access to AI technology. 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.

Why exclusive proprietary AI technology is a problem

AI is a key technology that drives innovation across society, economy, and science. For example, large language models (LLMs) such as GPT-4 have become the backbone for text processing in many fields such as education, entertainment, media, and management. Thus, downstream innovations including novel business models and novel products and services may be at risk when widespread access to AI becomes restricted. Not surprisingly, the concentration of power over technology is known to hamper future innovation, fair competition, scientific progress, and thus welfare and human development at large[1].

Proprietary AI technology could also jeopardize inclusiveness and responsibility. When new AI technologies like LLMs are exclusively developed by a few companies, those may also arbitrarily decide which countries and languages to support in their systems and may thus exclude users, such as those from small markets (for instance, the Global South and rare languages). A certain level of openness of AI technology is further necessary for researchers to determine the safety, security, and fairness inherent in AI systems operations. Proprietary AI systems are difficult to appraise by the public, and thus to identify and fix errors.

The benefits of open-source principles in software development

The cardinal idea of open-source software (OSS) is that an organization not only relies on its internal knowledge sources and resources to innovate, but also draws on multiple external technical sources, such as software packages, bug reports, customer feedback, published patents, or communities[2]. Depending on the chosen license, OSS may not preclude commercialization: companies can combine OSS with additional products and services to generate revenue (for instance, RedHat offers Linux for free but charges for premium support, Amazon contributes to Apache but charges for hosting services in the cloud, and so forth). Nowadays, the open-source model safeguards effective and efficient software development. Yet, it took several decades for this model to mature and, for companies, to realize and utilize its full potential[3]. Also, an important lesson from OSS is that governments played an important role in boosting OSS adoption[4], suggesting that many ways to promote open-source AI will benefit from governments taking an active role.

OSS offers several benefits relative to proprietary software in terms of accessibility, collaboration, responsibility, and interoperability[5]. First, proprietary software is mostly available under licensing fees. In contrast, OSS is free and comes with no or only limited restrictions on use, inspection, and modification. Second, OSS tends to be developed and maintained by a community. Diversity in an OSS community stimulates software quality, faster innovation, and increased creativity of OSS development relative to proprietary alternatives[6]. Third, errors in OSS are detected and corrected with everyday use, and thus much faster than in closed-source software[7], thus making OSS applicable to even critical and highly reliable technical systems. Fourth, OSS typically relies upon open standards and modularity, which decouples dependencies among software components and leads to greater reusability and interoperability[6].

In addition to cost savings, companies also benefit from OSS in multiple ways. Companies can build business models around complementary goods and services while adopting as backbone OSS software at a low cost[8]. In addition, companies can gain trust and reputation by contributing to OSS, which fosters recruiting top talents and the dissemination of their own technologies and products. By participating in OSS activities, companies can also steer the direction of innovation and control the further development of the technology on which they depend (for instance, by introducing a standard that helps the company compete more effectively). By contributing to OSS, companies also gain valuable feedback on their technologies and are able to identify potential issues with the products early on. Finally, the presence of OSS drives innovation and competitiveness of commercial solutions.

Promoting open-source AI technology

The development of open-source AI has several similarities with open-source software. Yet there are also some important differences that require a tailored approach to building open-source AI. While conventional software is programmed with explicit rules to perform a task, AI is programmed to learn to perform a task. As a result, AI technology has three essential components: datasets for training, source codes for formalizing the training task, and models that eventually store the trained weights. Training AI models requires substantial hardware resources and comes with high operating costs. Furthermore, the use of AI may expose society to large risks, for example, malicious use of AI to create misinformation, thus mandating a responsible approach to open-source AI technology. In the following, we discuss a tailored approach to open-source AI complementary to proprietary AI by fostering (1) accessibility, (2) collaboration, (3) responsibility, and (4) interoperability (see **Figure 1**).

Improving accessibility

To foster accessibility, policy-makers should proactively encourage the development and adoption of open-source AI. Since AI innovation is considerably more capital-intensive than regular software development given the need for data and infrastructures for building contemporary AI models, additional resources (such as funding and access to large-scale infrastructure and data) are needed to kickstart and scale open-source AI technology. Importantly, existing computational resources are often not of sufficient magnitude to build state-of-the-art AI technology comparable to that of for-profit companies. For example, the development of LLMs is estimated to cost between EUR 300 and 400 million. Another limiting factor is that, even if resources are made available, they are bound to academia and are thus inaccessible to other stakeholders such as non-profit organizations that seek opportunities where AI could be leveraged for social benefits. A promising counterexample is the U.S. roadmap to offer broader access to computational resources, including public-private partnerships. Since scientists are currently often unable to replicate AI technology from companies due to a lack of resources, such roadmaps can help facilitate reproducibility (for instance, via the ML Reproducibility Challenge).

To broaden access to data and models, policy-makers could support the development of open repositories for hosting both under a trustworthy and responsible governance model. Importantly, open data from public institutions are often large and originate from diverse sources, which is beneficial in practice. Furthermore, public institutions can actively incentivize data-sharing partnerships, which, in combination with federated learning, may promote AI across institutional boundaries while ensuring data privacy. For example, the German government recently launched a consortium called Mobility Data Space where different stakeholders in the mobility sector (such as public transport companies, private car-sharing providers, and car manufacturers) are able to access shared data, even those of competitors.

However, data sharing comes with challenges. First, opening up datasets increases the likelihood of privacy breaches and raises ethical issues around confidentiality, data misrepresentation, and informed consent. Second, organizing open data and maintaining fairness in terms of distribution rights and acknowledgments for its contributors is challenging. Fortunately, there has been recent progress with respect to the development of governance frameworks to tackle these challenges, such as the FOT-Net Data Sharing Framework designed for connected automated driving under the General Data Protection Regulation in the European Union. Such frameworks could be useful starting points in improving accessibility while tackling ethical, legal, and organizational challenges.

Finally, much educational material on state-of-the-art AI is managed by for-profit companies (such as Coursera and Udemy) and is often hidden behind paywalls. Hence, to promote the adoption of open-source AI, more effort is needed to improve access to high-quality educational materials. As a result of the above, the barrier of entry for contribution and access will considerably drop.

Improving collaboration

AI technologies may be jointly developed and maintained by diverse and inclusive communities of developers, users, and stakeholders. This collaborative approach may significantly reduce the cost of development and contribute to solving scaling problems. This will result in broad participation by stakeholders who can make the future of AI more inclusive and fairer.

To promote collaboration in open-source AI technology, clear steps should be taken for community building across academia, non-profit organizations, companies, and public institutions. Given that the development of AI models is less decomposable and that task division is more difficult as compared to standard software development, further effort is needed to develop suitable collaboration practices that allow for more iterative and parallel development processes. Here, the lessons learned from the Project BigScience[9] developing an LLM called BLOOM[10] are an important first step. Furthermore, policymakers should fund large-scale initiatives to produce open-source as complements to proprietary LLMs.

Creating synergies and networks between universities, research centers, government, and industry may lead to new ecosystems around open-source AI and become a driver for future innovation. Building such ecosystems is especially relevant for start-up firms, and small- and medium-sized enterprises[11] as they often lack dedicated infrastructure and capacities to boost AI technology.

Improving responsibility

It is important to establish clear barriers against the misuse of AI technology. To this end, access control, similar to existing norms for open data, is needed to enforce a responsible use of open-source AI in practice. Consider, for example, MIMIC-III, a large, freely available health-related dataset. Given the sensitive nature of medical data, MIMIC-III is open to researchers only after they undergo compulsory ethics training. Notwithstanding, the access control for open-source AI should follow a layered approach that varies across datasets, source codes, and models to ensure responsible use, including safety, security, and privacy.

In addition, novel licenses are required – inspired by OSS but carefully tailored to open-source AI[12]. Such licenses must ensure broad user access while enforcing guidelines that prohibit malicious practices (such as abusing LLMs for automatically generating propaganda campaigns) under legally enforceable premises. Furthermore, such licenses for open-source AI should include sub-clauses that define permissive and restrictive use and also how technology can or cannot be repurposed. Prominent examples are the RAIL licenses, which prevent irresponsible and harmful applications of AI technologies by granting permission only for certain use cases. Over time, customized variants of licenses for open-source AI could be developed, so that high-risk applications of AI technology become subject to a more restrictive use.

Similar to OSS, AI technology under open-source principles will be especially effective in addressing bias in AI systems and steering innovations in a fair, ethical, and trustworthy direction. First, due to the diversity of inputs from stakeholders from around the world, there will be a greater emphasis on removing bias. Addressing bias will be especially important both when curating datasets as well as when training models. Second, a common concern is that open-source AI may not have the same level of quality control and testing as proprietary solutions, leading to potential bugs accidentally introduced by its developers. To this end, collaboration is important as it naturally leads to extensive testing.

Further, the development of AI in open communities may introduce decentralized organizations (that is, with no authority hierarchies based on employment contracts). Many open communities have developed effective organizational structures based on meritocracy, effort, and expertise that are effective at resolving both coordination and cooperation issues including how to manage conflicts. For instance, the Debian community developed a constitution that determines the decision-making rights of contributors and a set of rules the community can resort to in case of conflicts or accountability issues. Lessons from communities such as Debian could be incorporated into establishing a functional organizational structure and effective governance for open-source AI communities. Likewise, as designated bodies for maintaining adherence to legal frameworks are typically missing and as questions around accountability are often unclear, there can be legal challenges that originate from regulatory compliance. Nevertheless, open-source AI technology brings important principles to the table that go beyond existing regulatory frameworks for a responsible and trustworthy use of AI.

It is also worth noting that there are privacy and security threats associated with the use of open-source AI. For example, malicious actors could perform backdoor attacks in which they manipulate a small portion of the training data to make an AI model learn additional, hidden functionalities[13]. In general, vulnerabilities in open-source AI are often exposed publicly, which can make attacks but also their identification easier. Furthermore, there are also risks for society since open-source AI can be used for nefarious purposes. Examples are the use of open-source AI technologies for the development of weapons and AI-generated propaganda campaigns[14]. Still, the benefits are likely to outweigh the downsides of open-source AI, especially if a responsible open-source approach with clear barriers against misuse is pursued, as we laid out above.

Improving interoperability

Over time, AI technology will need to build upon more standardized and modular building blocks within software libraries (such as prompt templates and standardized prompt optimizers in the case of LLMs) that allow for easier adoption and customization in downstream applications. Interoperability of pre-trained models across platforms should also drastically reduce the need to retrain large models. The result is a greater reusability of AI technologies, which can reduce the need to "reinvent the wheel" and thus promote faster iterations during development. Reusability is not only important for rapidly building AI applications but also for reusing high-quality source codes and models designed in a responsible and robust manner.

In terms of standardization, various regulatory bodies such as the International Organization for Standardization have several standards under draft that aim at the harmonization of AI technology. The current initiatives cover various aspects including life cycle management, data quality, risk management, and auditing. Such standardization roadmaps are helpful for developing trustworthy AI systems in high-risk applications (for example, through standardized conformity checks). Crucially, standardization must not only be on 'pen-and-paper' but eventually filled with life through software libraries for developing AI technology. In this regard, public funding to support the development of open-source libraries could be necessary, as well as corresponding educational resources and long-term maintenance.

As a result of a growing harmonization, the dependence on a specific AI technology will diminish, so that end-users avoid "lock-in" effects and benefit from reduced switching costs (for instance, when changing from the LLM of company A to that of company B). For developers, interoperability can eventually help counteract the growing inequality in the development, access, and use of AI technology, while also promoting effective competition. In this regard, a concern from a corporate perspective may be that, if AI research is forced to be open, then companies may not see value in investing as much in research and development as they would otherwise do. For example, the motivation of companies to develop new AI technologies may be reduced in the presence of open-source alternatives, which may hamper innovation more broadly and could eventually also lead to gatekeeping behavior of established companies. However, as we argue, the presence of open-source AI complementary to proprietary alternatives may also increase healthy competition, which can also make commercial products better.

Author contributions

All authors wrote, edited, and approved the manuscript.

Ethical declarations

Competing interests

The authors declare no competing interests.

References

1. Aghion, P., Harris, C., Howitt, P., & Vickers, J. Competition, imitation and growth with step-by-step innovation. *The Review of Economic Studies* **68**, 467–492 (2001).
2. Chesbrough, H. W. Open innovation: The new imperative for creating and profiting from technology. *Harvard Business Press* (2003).
3. Fitzgerald, B. The transformation of open-source software. *MIS Quarterly* **30**, 587–598 (2006).
4. Lerner, J., & Schankerman, M. The comingled code: Open source and economic development. MIT Press, Boston, MA (2013).
5. Haefliger, S., Von Krogh, G., & Spaeth, S. Code reuse in open-source software. *Management Science* **54**, 180–193 (2008).
6. Von Krogh, G., & Von Hippel, E. The promise of research on open-source software. *Management Science* **52**, 975–983 (2006).
7. Paulson, J. W., Succi, G., & Eberlein, A. An empirical study of open-source and closed-source software products. *IEEE Transactions on Software Engineering* **30**, 246–256 (2004).
8. Golden, B. Succeeding with open-source. Addison-Wesley Professional, Boston, MA (2005).
9. Akiki, C., Pistilli, G., Mieskes, M., Gallé, M., Wolf, T., Ilić, S., & Jernite, Y. BigScience: A case study in the social construction of a multilingual large language model. *NeurIPS Workshop on Broadening Research Collaborations in ML* (2022).
10. Scao, T. L., et al. BLOOM: A 176B-parameter open-access multilingual language model. *arXiv:2211.05100* (2022).
11. Jacobides, M. G., Brusoni, S., & Candelon, F. The evolutionary dynamics of the artificial intelligence ecosystem. *Strategy Science*. **6**, 412–435 (2021).
12. Contractor, D., et al. Behavioral use licensing for responsible AI. In *ACM Conference on Fairness, Accountability, and Transparency (FAccT)* (2022).
13. Saha, A., Subramanya, A., & Pirsiavash, H. Hidden trigger backdoor attacks. In *AAAI Conference on Artificial Intelligence* (2020).
14. Feuerriegel, S, DiResta, R., Goldstein, J. A., Kumar, S., Lorenz-Spreen, P., Tomz, M., & Pröllochs, N. Research can help to tackle AI-generated disinformation. *Nature Human Behaviour.* In Press (2023).

Display items

Figure 1. Key approaches to promote open-source AI technology. The suggested actions should foster accessibility, collaboration, responsibility, and interoperability.