Background Briefing on Open Source AI

On 10 November 2023, 17:30-18:30 CET, Women in AI Austria will host a panel discussion with Zuzanna Warsow (Open Future), Jeannette Gorzala (European AI Forum, AI Austria) and Alexander Baratsits (Creative Commons Austria, cba.media) to go deeper into the challenges and opportunities of open source AI.[1] The discussion will centre around the three themes of current regulatory developments around open source AI, the challenges around open washing, and finally, some best practices and positive outlooks for this developing field.

In this briefing, we will sketch some of the central themes of open source AI systems and attempt to roughly characterise a broad field marked by a variety of debates. Rather than define what open source AI is, we will try to show how open source AI relates to particular kinds of issues, concerns, movements and values. This briefing will be finalised and re-published after the event with the inputs provided in the course of through the discussion.

Open source movements

The first association made in connection with open source AI systems usually relates to open source movements, particularly around software. Initially, software was developed in an academic setting, bound to hardware but shared relatively freely within the community. Later, from the 1970s onwards, when computers became more widely available, software was developed into a proprietary product separate from the hardware; through these developments, software became subject of copyright laws, and thus entered the same legal domain as works of art. Pushback against this course of events led to the formation of, amongst others, the GNU project and the copyleft movement, which sought to liberate software development from the restraints imposed by copyright. The free software movement prioritised four fundamental freedoms for users: the freedom to run, study, modify, and distribute software. While the term 'free software' was initially used, the label 'open source' was popularized in the late 1990s, emphasising the practical benefits of collaboration and source code sharing. Over time, the movement expanded beyond software, influencing other domains such as hardware, data, and content, and has been integral to the development of countless technologies and platforms that underpin the modern digital age. The initial concerns of the advocates of free software are still relevant today, as the open source definition, while overlapping in many areas, did not always guarantee these initial freedoms, especially as industries found ways to monetize open source without granting users all these rights. On the other hand, ethical open software and AI developers explicitly reject the 'freedom' to run programs for biased or harmful purposes.

---

[1] https://www.linkedin.com/events/policychat-exploringopensourcea7115987928379535362/

With this brief characterisation, we can already begin to see some similarities but also differences between open source movements for software and for AI systems. Historically, the trajectory of enclosure - referring to previously public or shared resources - is at least similar, with much of AI research initially done in an academic setting and subsequently picked up and applied to generate profits. Yet the open source AI movement is complicated by the move of large tech companies investing in AI systems which are made easily accessible to others for their use and integration into applications. These AI systems are also called 'open' for the ease of access, although – and this distinction is important – the models released in such a way usually do not share the source code or allow reconfigurations by others.

Currently the Open Source Initiative is crafting a general definition of open source AI.[2] This definition has not yet become very concrete, yet it centres around licenses that make the study, use, modification and sharing of AI systems and their components by others possible. But the lack of a concrete definition currently has given rise to a variety of different conceptions. Sometimes, what is meant by open source AI includes both the algorithm and the training data; other times, the training data is not considered necessary for an AI system to be open source. Yet another approach is that the trained model is made available, without granting access to the algorithm and/or the data used to train it. What is most often seen as 'open' are AI models which can be used via API access, but these need not be open source in the more expanded view.

A common set of tools

Licensing has been and continues to be a preferred mechanism for ensuring free use of software, including AI systems. Several licenses are available (both restrictive and permissive), but permissive open source licenses are currently the most popular.[3] Open RAIL (Responsible AI Licenses) seek to integrate responsible practices into licensing and have become relatively widespread amongst the restrictive open source licenses.

In general, licenses try to bring the logic used for open source software and code into the framework of AI systems; that is, they offer control over their use through contractual means, effectively securing a legal right to prevent usage under specific conditions. How effective the legal right to prevent usage actually is will be put to the test by the rise of AI systems trained on copyrighted data, which we will discuss next. Another approach for more transparency considered by Wikimedia is governance is governance of AI systems through model cards, allowing the public to engage with how the model is made and suggest edits in the same collaborative vein as with Wikipedia entries.[4] However, a different set of tools also needs to be taken into account, namely the platforms used to share open source AI systems. The rise of HuggingFace as a well-known, central distribution point has made it vastly easier to share AI systems and to draw on the work of others for evaluating and developing applications. It also facilitates research about the kind of AI systems developed in different fields and allows for others to contribute to development. Yet its funding sources indicate that the openness of HuggingFace, as a platform, could be less than

---

[2] See https://opensource.org/deepdive/drafts/
[3] See https://openfuture.pubpub.org/pub/growth-of-responsible-ai-licensing/release/2
[4] https://arxiv.org/abs/1810.03993

permanent and potentially a strategy of enclosure akin to similar marketplace platforms.

What has been sidestepped so far are the infrastructures required for developing open source AI systems, especially those with high computational demands. While it should not be taken for granted that all AI systems require large, indeed ever-increasing, amounts of computing power and data to provide good outputs, the currently popular methods for AI models do rely on computational infrastructure that is in most cases unavailable to research outside of industry.

The question of data

As with all Machine Learning-based AI systems, open source AI systems are heavily reliant on the data used to train them (and all AI systems rely on data for testing). The difficulties start with obtaining datasets which can be used for AI development. Smaller initiatives often have difficulties obtaining sufficiently large datasets for training and testing their systems - this can be due to API restrictions (for example on social media data), to copyright or GDPR considerations, or to difficulties in labelling the data. In tandem with the increasing pushback of copyright holders in connection with generative AI, data available for scraping is quickly dwindling. At the same time, AI-generated content is increasingly spawning all over the Internet, with severe consequences for the usability of online content for the development of generative AI systems: as reported by Carl Franzen,[5] there is evidence that the output of generative AI systems deteriorates if trained on data that consists of too much AI-generated content.[6] Existing copyleft initiatives, for instance around the Creative Commons licenses, could collaborate in the development of open source AI systems – that is, if the computational infrastructure were also available to them or their users.

But open source AI and data intersect in another matter, namely in the issue of licensing. As software, copyright applies to AI systems, hence the focus on licensing that allow open usage (whether restricted or unrestricted). We effectively have analogies of legal frameworks applied to content, which was subsequently extended to software and now becomes applicable to AI systems. The implications of this still need to be explored – for instance, does that mean that open source AI systems should necessarily have open datasets? To which point do open source AI systems need to be open in order to allow for the pursuit of community values?

Points of power

Open source AI systems do exist, but mostly have restrictive licenses, e.g. Apache 2.0. Generally, they are open to varying degrees: most make the model available, without opening up the algorithm or the training data. A famous example is LLaMA, an LLM developed by Meta which was leaked and immediately taken up for experimentation by developers and researchers; yet the non-disclosure of training data (amongst other things) puts the actual openness of the model in questionable. OpenAI's Whisper, a speech-to-text model, and Jukebox, a song generation model, are two of the few models actually released with open licenses from this house. Mistral.ai is one of

---

[5] https://venturebeat.com/ai/the-ai-feedback-loop-researchers-warn-of-model-collapse-as-ai-trains-on-ai-generated-content/
[6] https://arxiv.org/pdf/2305.17493v2.pdf

the European open source developers focusing on LLMs, while EleutherAI's research and releases of open source datasets and LLMs has shifted to increasing interpretability and what they term 'alignment' of AI systems. BLOOM, an open source LLM, was developed by AI researchers and a consortium of companies to allow wider access to LLM capabilities. For many open source developers, affording compute infrastructure and access to data is the biggest challenge. Therefore, funding can be a challenge and is frequently obtained from cloud providers or other big tech industry. In addition to pre-trained models, there are also a large number of open source libraries and other components publicly shared, to make the development of AI systems possible, which are owned by for-profit companies or foundations alike.

But not all open source development faces funding challenges; some companies engage in open source-ish development strategies. Usually, proprietary development ensures the economic viability of products by enclosing them and making use of the system conditional on some exchange of value, be it monetary, expressed in data, or in other forms. However, an equally attractive strategy might be the reconfiguration of others' systems in a way which creates dependence on the developer, thereby exerting an infrastructural pull that can be harnessed (for profit) at opportune moments in time. Effectively, open source with restricted licenses (and the platformisation this strategy effects) provides an opportunity for companies to create revenue streams from standardised services that have become ubiquitous. For instance, the base source code of one of Alphabet's most commonly used services, Android OS, is available under an open source license, but it is only made public once a new major version has been released. In addition, most Android distributions come with a significant amount of (Google) proprietary software, which (alongside the Android trademark) can only be used through individual licensing agreements. These mechanisms of platformisation point to different pathways towards enclosure for different kinds of market actors. How open source AI will fit into the existing patchwork of strategic enclosure will be interesting to observe.

Value creation

Free and open source movements have historically been closely associated with ideas of commons – that is, communities of practice collaborating to manage shared resources. In a context where AI systems are increasingly embedded into a variety of contexts and products, open source AI systems promise avenues of scrutiny by the public and ways for broader engagement with the development of AI systems, both at a technical and at a political level.

But there are significant value differences between open source movements, ranging from permissionless innovation to actively creating counterpoints to dependencies on large tech companies. In the meantime, the label of 'open' has gained a positive connotation, signifying transparency and inclusion; it falls into a similar line of signification as the efforts around 'democratising' AI. With both meanings, it is important to understand who mobilises them when – and there has been a decided shift by large companies towards communicating (marketing) in terms of community values (which may or may not be put into practice). In this context, it is also important that European political leaders have framed open source AI development as a pathway towards open strategic autonomy - or 'digital sovereignty' -, which (amongst other things) is a geopolitical

effort to secure independence from providers of strategically important technologies from certain regions. Also perhaps strategically important - but not at all 'open' in terms of known - is the environmental impact of open source AI systems: although in general the environmental impact of AI systems has not yet been fully established, and is significantly more difficult to establish for smaller companies,[7] open source AI systems promise a lower environmental footprint when considering their distribution as the environmental cost of training is distributed amongst a wider user base.

Additionally, the question of value accrual has not been fully settled. The high computational costs of developing AI systems and the challenges with data, as well as the ubiquitous question of remuneration in an open source/open data context, present challenges to the evolving field of open source AI. With proprietary AI systems, the first court cases are popping up to secure remuneration for copyright holders; it remains to be seen whether the stewardship of outputs, often performed by gig workers at low cost, will be equally counted as contributions to the functioning of these AI systems. What this might mean in a less-resourced open source context remains to be seen.

Recommended reading

For those interested in staying up-to-date on open source AI, Open(ish) Machine Learning News might be a valuable resource. For a very recent review of open source, Demos prepared a policy brief as input to the UK's AI Safety Summit. We also liked this Medium article if you want to go deeper into the challenges related to data colonialism in an open source AI context.

---

[7] https://blog.eleuther.ai/fmti-critique/

Profiles of panelists

Zuzanna Warso[8]

Zuzanna works at the intersection of law, ethics and technology, both within academia and through her policy and advocacy work. With a PhD in law and an M.A. in English Literature, she has received a number of highly celebrated scholarships and the Marshall Memorial Fellowship. Currently, she is an advisory board member of the Institute for the Ethics of AI at the Technical University Munich and Director of Research at Open Future.

Jeannette Gorzala

Jeannette is a celebrated member of the Austrian legal community focused on technology, Legal Community Lead and board member of AI Austria, and the vice-president of the European AI Forum. She acts as Associate Professor at the Woxsen University India and has published extensively on the legal implications of AI, blockchain and other emerging technologies.

Alexander Baratsits[9]

Alexander Baratsits has extensively worked as legal counsel, e.g. at the Institute of Science and Technology Austria, and has been a founding board member of Cultural Broadcasting Archive, an association for the promotion of digital communication (cba.media). In addition, he has been extensively involved in the Austrian Creative Commons community, where he has acted as Chapter Lead since 2020.

---

[8] https://openfuture.eu/author/zuzanna/
[9] http://www.baratsits.at/cv/