

AI: A way to freely share technology and stop misuse already exists

February 12 2024, by Joseph Lindley and Jesse Josua Benjamin



Credit: AI-generated image

There are lots of proposed ways to try to place limits on artificial intelligence (AI), because of its potential to cause harm in society, as well as its benefits.

For example, the EU's AI Act places greater restrictions on systems

based on whether they fall into the category of general purpose and generative AI or are considered to pose limited risk, high risk or an unacceptable risk.

This is a novel and bold approach to mitigating any ill effects. But what if we could adapt some tools that already exist? Software licensing is one well-known model that could be tailored so that they could meet the challenges posed by advanced AI systems.

[Open responsible AI licenses \(OpenRails\)](#) might be part of this answer. AI that is licensed with OpenRail is similar to [open-source software](#). A developer may release their system publicly under the license. This means that anyone is free to use, adapt and re-share what was originally licensed.

The difference with OpenRail is the addition of conditions on using the AI responsibly. These include not breaking the law, impersonating people without consent or discriminating against people.

Alongside the mandatory conditions, OpenRails can be adapted to include other conditions that are directly relevant to the specific technology. For example, if an AI was created to categorize apples, the developer may specify it should never be used to categorize oranges, as doing so would be irresponsible.

The reason this model can be helpful is that many AI technologies are so general, they could be used for many things. It's really hard to predict the nefarious ways they might be exploited.

So this model allows developers to help push forward open innovation while reducing the risk that their ideas might be used in irresponsible ways.

Open but responsible

In contrast, [proprietary licenses](#) are more restrictive on how software can be used and adapted. They are designed to protect the interests of the creators and investors and have helped tech giants like Microsoft to [build vast empires](#) by charging for access to their systems.

Due to its broad reach, AI arguably demands a different, more nuanced approach that could promote the openness that drives progress. Currently many big firms are operating proprietary—closed—AI systems. But this could change, as there are several examples of companies using an open-source approach.

Meta's generative [AI system Llama-v2](#) and the image generator [Stable Diffusion](#) are open source. French AI startup Mistral, established in 2023 and now valued at US\$2 billion (£1.6 billion), is set to soon openly release its latest model, which is rumored to have [performance comparable to GPT-4](#) (the model behind Chat GPT).

However, openness needs to be tempered with a sense of responsibility to society, because of the potential risks associated with AI. These include the potential for algorithms to [discriminate against people](#), [replace jobs](#) and even pose [existential threats to humanity](#).

We should also consider the more humdrum and everyday uses of AI. The technology will increasingly become part of our societal infrastructure, a central part of how we access information, construct opinions, and express ourselves culturally.

Such a universally important technology brings its own kind of risk, distinct from the robot apocalypse, but still very worthy of consideration.

One way to do this is contrast what AI may do in the future, to what [free](#)

[speech](#) does now. The free sharing of ideas is not only crucial to upholding [democratic values](#) but it's also the engine of culture. It facilitates innovation, encourages diversity and allows us to discern truth from falsehood.

The AI models being developed today will likely become a primary means of accessing information. They will shape what we say, what we see, what we hear and, by extension, how we think.

In other words, they will shape our culture in much the same way that free speech has. For this reason, there is a good argument that the fruits of AI innovation should be free, shared and open. And, as it happens, most of it already is.

Limits are needed

On the [HuggingFace platform](#), the world's largest AI developer hub, there are currently over 81,000 models that are published using "permissive open-source" licenses. Just as the right to speak freely overwhelmingly benefits society, this open sharing of AI is an engine for progress.

However, free speech has necessary ethical and legal limits. Making false claims that are harmful to others or expressions of hatred based on ethnicity, religion, or disability are both widely accepted limitations. Providing innovators with a means to find this balance in the realm of AI innovation is what OpenRails do.

For example, [deep-learning technology](#) is applied in many worthy domains, but also underpins [deepfake videos](#). The developers probably did not want their work to be used to spread misinformation or create non-consensual pornography.

An OpenRail would have provided them with the ability to share their work with restrictions that would forbid, for example, anything that would violate the law, cause harm, or result in discrimination.

Legally enforceable

Can OpenRAIL licenses help us avoid the inevitable ethical dilemmas that AI will pose? Licensing can only go so far, with one limitation being that licenses are only as good as the ability to enforce them.

Currently, enforcement would probably be similar to enforcement for music copying and software piracy and would involve the sending of cease and desist letters with the prospect of potential court action. While such measures do not stop piracy, they do discourage it.

Despite limitations there are many practical benefits, licenses are well understood by the tech community, are easily scalable, and can be adopted with little effort. This has been recognized by developers, and to date, more than 35,000 models hosted on HuggingFace have adopted OpenRails.

Ironically given the company name, OpenAI—the company behind ChatGPT—does not license its most powerful AI models openly. Instead, with its flagship language models the company operates a closed approach that provides access to the AI to anyone willing to pay, while preventing others from building on, or adapting, the underlying technology.

As with the free speech analogy, the freedom to share AI openly is a right we should hold dearly, but perhaps not absolutely. While not a cure-all, licensing-based approaches such as OpenRail look like a promising piece of the puzzle.

This article is republished from [The Conversation](#) under a Creative Commons license. Read the [original article](#).

Provided by The Conversation

Citation: AI: A way to freely share technology and stop misuse already exists (2024, February 12) retrieved 29 April 2024 from <https://techxplore.com/news/2024-02-ai-freely-technology-misuse.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.