Before the
National Telecommunications and Information Administration
Washington, DC

In re

Dual Use Foundation Artificial Intelligence Models With Widely Available Model Weights

Docket No. 240216-0052

COMMENTS OF COMPUTER & COMMUNICATIONS INDUSTRY ASSOCIATION

The Computer & Communications Industry (CCIA)[1] submits the following comments in response to NTIA's February 26, 2024, Request for Comments.[2]

CCIA is an international, not-for-profit trade association representing a broad cross section of communications and technology firms. For more than fifty years, CCIA has promoted open markets, open systems, and open networks. CCIA members employ more than 1.6 million workers, invest more than $100 billion in research and development, and contribute trillions of dollars in productivity to the global economy.

CCIA members are at the forefront of research and development in technological fields such as artificial intelligence and machine learning. Many CCIA members also engage in open-source software development and have learned a number of lessons regarding the benefits and challenges of this approach. CCIA appreciates NTIA's consideration of our comments regarding how those lessons may be applied in the context of artificial intelligence.

I. Summary

Open-source software is widely regarded as a generally successful approach to software development. It now underlies much of our communications and the vast majority of Americans use systems based on open-source software (knowingly or unknowingly) on a daily basis. Now, AI systems are experiencing a similar moment. Open-source and partially public models are available, with new models and capabilities released regularly. Some of these models are formally open-sourced, with copyleft or MIT licenses, while others are open to a varying degree in terms of model weights, datasets, and code. This spectrum of approaches is not amenable to a simple "open/closed" binary, and NTIA should endeavor to analyze them on the basis of risk, rather than on what is and is not released and to whom.

This approach to AI development brings new challenges, but also significant benefits. Open models provide the potential for better security, less bias, and lower costs to AI developers and users alike. Open models also present advantages in AI governance, being easier to

[1] A list of CCIA members is available online at https://www.ccianet.org/about/members.
[2] 89 Fed. Reg. 14059 (Feb. 26, 2024) (hereinafter "Request").

understand and test. Open models also compete with closed models, resulting in enhanced innovation and benefiting consumers and AI deployers alike. While there are risks to open AI models, most of those risks are equally present in closed models and should not outweigh the significant benefits open development approaches can provide.

Finally, there are significant efforts on AI risk evaluation and governance taking place across industry, non-governmental international organizations, and government agencies. NTIA should work in partnership with these stakeholders to ensure that a common standard can be applied. NTIA should also consider how it could best support the integration of AI into society via mechanisms similar to NTIA's Internet For All Workforce Planning Guide.

II. Defining "Open" And "Widely Available"

Before discussing an appropriate definition of "open" or "widely available" models (referred to collectively as "public innovation" models hereafter), it is critical to emphasize that such a public innovation model is not necessarily a dual-use model, as that term is defined in the AI Executive Order. Maintaining a clear distinction between dual-use and non-dual use public innovation models in the ultimate rule is critical to ensuring that open-source development of AI models remains a viable approach. Absent a clear distinction, it is likely that these approaches will become effectively impossible in the United States for any but the best-funded AI developers. This will prevent researchers and scientists in smaller labs from taking advantage of AI's benefits. It is also likely to harm security research into AI as researchers will be limited to only those interactions that a closed model permits.

In order to minimize the possibility of such a conflation, it would be appropriate for any rules, regulations, benchmarks, recommendations, or other outputs from this process to employ an assessment of the dual-use nature of a model before any additional constraints based on its public innovation nature apply.

A. Defining "open"

As the Request acknowledges, there is no single category of public innovation model. Instead, there is a broad spectrum based on how much of the system is open and how that system is managed. A fully open-source model, where everything is public from the training data to the implementation code to the model weights, will have different concerns than an open model in which only weights and code are made publicly available and both will differ from a public innovation model in which implementation code and a sample model, but not a fully trained model, is publicly available. And beyond simple publication of aspects of a system, there are other traditional "openness" concerns observed in the context of open-source software to consider, such as contractual restrictions on modification or use of the released model. In each of these cases, the appropriate metrics and approaches will differ accordingly.

Instead of employing a binary open/not open categorization, NTIA should consider the full spectrum of possible public innovation approaches. NTIA should apply a purely risk-based approach uniformly across closed models and public innovation models alike, with the openness being just one factor in evaluating the risk of the system. Alternatively, NTIA could consider an approach in which certain aspects of any proposed rule are tied to specific characteristics of a model—for example, limiting some requirements or recommendations to only those models that publicly release weights.

B. Defining "widely available"

It is impossible to set a specific threshold for "wide availability" and NTIA should not attempt to do so. Instead, this categorization should be tied to risk assessment; "wide availability" of a conversational generative AI could require essentially completely public availability, while wide availability of a pesticide design AI might be triggered at a significantly lower number of distributions because of the foreseeable risk it might be used to create chemical weapons similar to a pesticide.

C. Thresholding dual-use foundation models

The current definition used for a dual-use foundation model is based on a particular amount of computation used in training the model. This is a poor proxy for the capabilities of a model. Instead, qualitative approaches in the vein of the AI Verify Toolkit should be used. This will allow a model to be assessed more directly for foundation capability, rather than relying on the imperfect proxy of compute resources.

III. Risks and Benefits of Public Innovation Models

Wide availability of a public innovation model provides both risks and benefits. While wide availability increases the possibility of misuse by making access easier to obtain, it also provides a number of benefits in terms of security, follow-on innovation, and reduced bias. NTIA should not put a thumb on the scale in favor of either public innovation or closed models, but rather assess based on the risk-benefit balance in any given case.

A. Benefits

The public innovation model in AI development benefits from many of the same advantages that open-source development has shown over the past decades. This includes enhanced security via external code review and the ability to patch vulnerabilities without being forced to rely on a single vendor, reducing barriers to entry for SMEs by providing access to basic tools they can apply in their particular problem domain, positive economic impacts through competition with closed models and reductions in cost of operations due to lower software costs, and follow-on innovation via code being reused and repurposed for new applications. And, while a minor benefit in some circumstances, the ability of a public innovation model to be run without a network connection may prove critical for AI intended for use in situations like disasters where telecommunications services cannot be guaranteed. Alongside this benefit, the fact that a public innovation model can typically be run locally reduces some privacy and cybersecurity concerns.

These benefits are not necessarily tied to any specific aspect of a model being open. Security review may rely more on code and less on weights, while the inverse may be true for the equity benefits of public innovation models.

In public innovation AI models, there is an additional benefit—equity can be improved. While a developer should consider equity as they create a model, it is not guaranteed that they will do so or that they will be successful. However, if the model is a public innovation model, others may be able to enhance the fairness and reduce biases in the model via their own independent contributions. Those contributions then benefit other users of that model as they are incorporated into the public innovation model project.

The traditional benefits are even more apparent in the AI sphere because of the extremely high cost of initial training of a model. This cost can easily run into millions or tens of millions of dollars, limiting access to AI models for those who cannot afford this entry cost. Public innovation AI models allow others to benefit from training, as has been observed with the plethora of models built off of Meta's LLaMA.[3]

The economic benefits are a major portion of the advantages of a public innovation model. Even with the limited availability of public innovation models to date, a huge amount of innovation has come out of reuse of those models. For example, the public release of Stable Diffusion's image generation model has allowed development of new tools for artists.[4] Recent releases of multi-modal open-source models such as LLaVa are likely to extend use cases into applications such as accessibility (automatic generation of alt-text for web images that lack it) and multimedia production such as Apple's MGIE[5] instruction-based image editing. As additional public innovation AI models with different and new capabilities continue to be released[6] the number and variety of applications will continue to increase alongside them.

B. Risks

While public innovation models create many benefits, there are some risks. However, those risks should not be overestimated and should be compared with the same risks from non-public innovation models. For example, while a malicious actor could employ a public innovation AI model, they could also employ prompt injection attacks to similarly employ a closed model.[7] Remote desktop tools provide important capabilities in computing, but can be abused by cybercriminals. Similarly, pen-testing tools provide critical capabilities to IT professionals, despite the possibility of their misuse by criminals. The single largest differential risk is the inability to "claw back" a public innovation model. Once publicly released, it will remain public. However, this risk presumes that misuse of the model is easier than misuse of a closed model, which likely will not be the case in many circumstances, and that the risks of release outweigh the risks, costs, and lack of competition in a closed-only AI ecosystem.

Further, there is a key difference between a public innovation model and a closed model in terms of risk surface. With a closed model, there is the additional filter of whatever 'glue' is employed between the user and the model itself. This is often where mitigations are placed, and because of this, bypass attacks are a frequent vector to try to misuse closed AI models. In contrast, in a public innovation AI model, any mitigations must fundamentally be part of the model. This may be more difficult to accomplish, but it also makes a public innovation far more resistant to attack after a mitigation is achieved.

Only in the rare case where a true differential risk exists should NTIA focus on the open/closed spectrum, and the focus must go both ways—addressing situations, such as equity, where a public innovation model is likely to be less risky than a closed model.

[3] https://www.technologyreview.com/2023/05/12/1072950/open-source-ai-google-openai-eleuther-meta/

[4] https://www.alpacaml.com/

[5] https://github.com/apple/ml-mgie

[6] *See, e.g.*, https://ai.google.dev/gemma/.

[7] https://simonwillison.net/2023/Nov/27/prompt-injection-explained/

IV. Mechanisms For Risk Management

Risk evaluation and benchmarking is a continually—and rapidly—evolving area of artificial intelligence technology. Government agencies such as NIST, international organizations like the G7, private entity forums like the AI Alliance or AI Verify, and traditional standards bodies like ISO all are developing metrics for quantifying and qualifying risks of AI systems.

Governance of AI systems is another area that is benefited by public innovation models. While it can be difficult to characterize a closed model due to limitations on access,[8] public innovation models do not suffer from this problem. As governance entities seek to create risk benchmarks and standards for AI models, this process will be made simpler and more effective because of the access to the model provided by public innovation approaches. For example, benchmarking equity concerns related to a model is far simpler when that model is publicly available, and ensuring that any fixes are truly part of the model rather than being bandages slapped on top of a model that is still flawed is much simpler in a public innovation environment.

CCIA urges NTIA to work with NIST on common standards for risk management, in line with the NIST AI Risk Management Framework, as well as other bodies engaged in the creation of tools and standards for AI risk management and evaluation. A common approach will help minimize compliance burdens on SMEs who wish to engage in AI development, which is likely to become even more common as public innovation models continue to be released.

V. Conclusion

CCIA appreciates NTIA's consideration of our comments and its attention to this important issue.

Respectfully submitted,

Joshua Landau
Senior Counsel, Innovation Policy
Computer & Communications Industry Association
25 Massachusetts Ave NW
Suite 300C
Washington, DC 20001
jlandau@ccianet.org

[8] *See, e.g.*, https://hackingsemantics.xyz/2023/closed-baselines/; https://simonwillison.net/2023/Jun/4/closed-model-training/.