Some Thoughts on the Upcoming Regulation of Open-Source Models

03/25/24

(1.) Is there evidence or historical examples suggesting that weights of models similar to currently-closed AI systems will, or will not, likely become widely available? If so, what are they?
While it is impossible to say how exactly the open source community will continue to formulate AI models in future, many tools allowing for the creation of resource databases have seen consistent development. The majority of such models can be downloaded for free on HuggingFace and Civit.ai, and tools such as Automatic1111 and ComfyUI have also seen major updates that both improve their resource efficiency and ease of use. It is reasonable to assume that as time continues, open source AI models will not only become more advanced but their interfaces will become easier for the layman to utilize.

One can also look at the plethora of open source programs that have nothing to do with AI but see much distribution because they provide perfectly competent, free alternatives to commercial software: Blender, Musescore, Wordpress, Krita, etc. It is totally natural for people to prefer software that allow them to do most of what they wish for free, and this is why so many eagerly anticipate the release of open source models in the first place. In the AI space, no matter how advanced an online solution might be, people will always have their eyes on a local, open source equivalent because it makes good sense to have total control over a program at no cost versus having to contend with the whims of a private entity.

(2.) Is it possible to generally estimate the timeframe between the deployment of a closed model and the deployment of an open foundation model of similar performance on relevant tasks? How do you expect that timeframe to change? Based on what variables? How do you expect those variables to change in the coming months and years?
I would say that the time gap between the release of a closed source model and its open source alternative takes roundabout eight months. It takes approximately five to six months for open source models' revisions to be announced, and then another couple months after that for those revisions to see wide distribution. This estimated time frame will likely shrink as open source solutions continue to mature and workflows are optimized, especially as AIs meant to aid in coding further develop and thus reduce time that developers spend troubleshooting bothersome bugs.

(3.) Should "wide availability" of model weights be defined by level of distribution? If so, at what level of distribution (e.g., 10,000 entities; 1 million entities; open publication; etc.) should model weights be presumed to be "widely available"? If not, how should NTIA define "wide availability?"
I would argue that "wide availability" should be defined by the licenses under which AI models are released under. Some licenses allow for more freedom than others, though I would say that any AI that is released under a license that allows for: commercial use; the ability to remix and modify content; the ability to share this modified content should all count as "widely available." An AI framework might be distributed among

many machines, but its license could be deceptive and severely restrict how its output could be utilized, as well as how the AI might be further disseminated via means unforeseen by its developers. Any AI that enforces censorship, outright denying people certain sorts of generations, would be limiting its "availability" because there are some people who would legally be excluded from its use.

(4.) Do certain forms of access to an open foundation model (web applications, Application Programming Interfaces (API), local hosting, edge deployment) provide more or less benefit or more or less risk than others? Are these risks dependent on other details of the system or application enabling access?

Web-based portals are inherently less secure than locally hosted models due to the fact that not only could network traffic be intercepted by third parties, but the information that a user inputs for the AI to parse has to be interpreted (and likely stored) on some far-off server. The two advantages of web hosts are ease of access, as well as the fact that these services allow users to take advantage of other organizations' and individuals' machines' computational power for AI generations. Not all computers are built equal, so a budget laptop is not going to be nearly as optimized for AI generations as the more powerful GPUs of a commercial company. This fact obviously means that lower income people are more likely to use web-based solutions, even if it means sacrificing some of their privacy because the alternative is not gaining access to certain AIs at all.

Unfortunately, web-based AIs are overwhelmingly closed source, and the actual mechanisms pertaining to how users' data are processed are opaque. This fact means that anytime someone utilizes a web-based AI portal, they are rolling the dice on whether or not the AI's hosts are good actors. There is no solving the issue of web-based AIs' privacy compromises unless more stringent federal legislation is passed that bars certain forms of data to be stored by companies while instituting regular checks that make sure companies are following these rules to the letter. How the European Union regulates the storage of user data could serve as a great role model in this regard. Unfortunately, until that happens, anyone greatly concerned with privacy (and who could afford it) should prioritize becoming familiarized with local AI systems. This route will almost inevitably lead them into the open source scene because the large majority of local AIs are open source. To run these local models may require more technical acumen than the curated models online, but they allow for far more flexibility of use.

(5.) Are there promising prospective forms or modes of access that could strike a more favorable benefit-risk balance? If so, what are they?

Local AIs have increasingly become more performant over time while using less resources due to more optimized code. However, there are certain services such as Google Collab that allow for people without access to advanced GPUs to experiment online with advanced, open source AIs, disregarding the need to register accounts and/or phone numbers on closed source websites (though a Google account is required). However, due to the fact that simulated code environments are run by for-profit companies that have yet to figure out how to properly monetize such widespread usage of their systems, there is always the risk that a Collab space could get unceremoniously shut down, depriving people of its use. There have been a couple instances of this in the local large language model (LLM) AI space because the models are so large and many people flock to use them to experience novelties – frustrating Google and causing them to shutdown respective Collab spaces.

For ease of access, a taxpayer funded, national alternative to something like Google Collab would be ideal. It would allow for many people with weaker hardware to easily experiment with open source AIs without having to worry about offending a private entity. And if the national government were to overhaul how they process users' data, it would be a better safeguard against privacy violations to boot.

(6.) How do the risks associated with making model weights widely available compare to the risks associated with non-public model weights?
The risks are much lower because open source models are largely transparent by nature. Maybe everything that goes into them cannot be easily discerned or replicated by strangers, but the code that went into constructing them and (oftentimes) the databases used to create them are publicly viewable. That is inherently more trustworthy than closed source models that are far more opaque, keeping both the code and databases locked away from public view in almost all instances.

(7.) What, if any, are the risks associated with widely available model weights? How do these risks change, if at all, when the training data or source code associated with fine tuning, pretraining, or deploying a model is simultaneously widely available?
There is no denying that widely accessible AIs allow for such things as "deep fakes" to be more easily generated by people, allowing both for the public to become more easily fooled by duplicitous actors, as well as allowing for certain types of content to be created that offends various groups' sensibilities. However, the fear of AI "stealing" copyrighted works from people is false. This is a fear that often pops up in artistic communities that claim deep-learning AIs plagiarize from their work. While AIs do not learn in quite the same fashion as humans do, those who claim that open source AIs are outright plagiarizing the works that have been fed to them are misrepresenting how deep learning works. AI-generated works are not “collages” of pre-existing content; the AIs take concepts that they picked out from relevant data and remix them in such a fashion that new works are created. And in this way, the process is not altogether different from what humans get up to in studying from prior works and using those inspirations to inform new output.

On the security end of things, easily available source code could allow for potentially malicious AIs (those that harvest data or otherwise compromise users' systems) to be recognized much quicker than closed AIs because the software's code that would perform such operations is easily viewable. Or perhaps the AI is sold on a promise that the code cannot actually satisfy. And if such duplicitous code *is* detected, then communities could easily fork the program, remove the offending code, and thus be left with a clean alternative.

I do not believe that the ease by which people can view an AI model's training data has much of an influence in these regards because it would only be relevant to those who are more worried about the *types* of content that an AI could generate rather than its basic efficacy. It is true that an AI model with a terribly limited training database would have an extremely difficult time generating much of anything because it wouldn't have enough reference material to work with, though quantity of reference material is not an issue in the vast majority of instances.

(8.) Could open foundation models reduce equity in rights and safety-impacting AI systems (e.g. healthcare, education, criminal justice, housing, online platforms, etc.)?

I believe the risk to be extremely minimal. Proper educational programs dedicated to developing our society's critical-thinking skills so as to better identify when something or someone is attempting to deceive them would solve many of these potential issues over the long term. We already live in a world wherein rumors exist (as propagated by various web forums), and where non-AI programs such as Photoshop can be used to visually doctor evidence of something. Freedom of Speech has allowed for people to say nearly whatever they want about our society, and our philosophy has traditionally been that education and good sense should be enough to combat such deviancy. Censorship should never be employed unless the form of speech otherwise allowed could only ever result in harm (such as someone falsely shouting, "Fire!")

Specifically regarding content generation, whether for education or entertainment, AIs will actually level the playing field because they will allow for people to make more using fewer resources than what has traditionally been required. Video-based AIs are already indicating that in a handful of years, professional-quality films may very well be generated by amateurs without studio interference, reducing the overall reliance on big capital from creatives' ends.

(9.) What, if any, risks related to privacy could result from the wide availability of model weights?

Impersonating someone and/or framing them as having done something that they never did can already be accomplished by current technology – the fear of open source AIs being used to outright ruin people's lives is overblown. Regarding the privacy of AI users: locally hosted AIs, whether open or closed source, provide little, if any risk of compromising anyone's privacy because they are localized entirely on one person's machine, and keep all inputs secret. If we wish to reduce any potential privacy violations, then we should encourage *more* individual ownership of models, not less.

(10.) Are there novel ways that state or non-state actors could use widely available model weights to create or exacerbate security risks, including but not limited to threats to infrastructure, public health, human and civil rights, democracy, defense, and the economy?

Hostile actors could attempt to manipulate public opinion by disseminating fallacious documents and/or images, though representative republics would not suffer so much more from this than authoritarian regimes. Any AI-generated propaganda generated proclaiming one narrative could easily find its counterpoint getting generated by opponents. I don't see how civil rights, or the rights of people more generally would be adversely impacted by individuals' use of AI. There are just as many ways that open AIs could be used to combat hateful bigotry as there are to spread it, and certainly, an AI on its own would not be able to change anyone's opinion about major social issues; maybe falsified AI generations could push someone into a certain direction, but that person would have already been leaning that way in the first place.

There is an under-discussed endeavor that AI could be applied to, and that would be in service of social profiling initiatives on the parts of governments and private organizations. With the amount of footage, pictures, and online profiles that exist of many people, it would be easy for advanced AIs to generate algorithmic predictions concerning someone's behavior. This assessment could be utilized by law enforcement to antagonize would-be criminals, or it could be used by private organizations to sell people things. Both states and private organizations could utilize AI to manipulate people's actions if they feel they have good models to work with. This would be unethical. Our federal and state governments should take every step required to limit social profiling via whatever laws because this particular application of AI could lead to prejudicial decision-making on the part of whatever organization employed it.

Another related area that should receive attention is in its deployment in the hiring/firing of employees at private companies. Entities like Amazon are already using AI for the sake of determining who would make for a "good" worker to hire, who's performing their duties "responsibly," who's so much load to be jettisoned, etc. I believe that AIs should be banned from determining how employees are hired/fired because too often have they been guilty of employing unfairly discriminatory criteria toward these people. And companies could far too easily claim that even if it is found that they're using a discriminatory algorithm that led to unjust treatment of employees, the company itself is not at fault because they didn't do much throughout the ordeal – the AI was in charge of things, so aren't they to blame? Atop banning the usage of AIs for employment purposes, the government should also mandate that all companies describe the methods by which any algorithms they utilized function in easily understandable terms to the relevant regulatory agencies. This way, we as a society will be able to more easily tell how deleterious these closed source, proprietary algorithms are toward the maintenance of proper human dignity.

(11.) How do these risks compare to those associated with closed models?

The risks associated with closed models are no better than those associated with open source models, and could actually be outright worse. Private companies manage closed models and attempt to generate profit off of their AIs, attempting to monopolize their usage. Given that artificial intelligence will be increasingly relied upon for the acquisition and dissemination of information, as well as employees' willingness to use it becoming prerequisite for a number of jobs, allowing any private organizations to entirely gear the field toward profit would do long-term damage to our society.

Private companies do not hold any sense of social responsibility; what they do is create products that conform to what they believe paying members of the public desire on wide scale. If society evolves in such a fashion that ignorant people with hateful ideas come to be a very profitable base to extrapolate profit from, there would be almost nothing to stop companies from using their closed AIs to cater to these groups' ideologies while also barring the less profitable groups of a different political persuasion from being able to make their voice heard on the same playing field. Without open source models around to counteract companies' pandering tendencies in this hypothetical scenario, those pulling for a more equal society would be left on the back foot.

(12.) How do these risks compare to those associated with other types of software systems and information resources?

The internet already allows for vast troves of misinformation to spread like wildfire, and impersonations of people/manipulating their image has already been facilitated by various creative software. Photoshop, Audacity, Blender, what-have-you. One could argue that the very act of satire, which has existed for thousands of years, and which has damaged the reputations of millions of people in that timespan, is a dangerous force in society due to how much of it has been predicated on falsehoods taking advantage of the public's immature tendencies. However, the fact that satire has been used to ill ends does not mean that it should be gotten rid of or that the majority of instances cannot hold people in society accountable quite effectively.

Like anything facilitating greater information dispersal, any issues that result from open source AIs would be due to the individuals using them, not the software itself. Not only that, but the quantity of individuals who would employ the software for such deleterious purposes would certainly find themselves in the minority of users.

(13.) What, if any, risks could result from differences in access to widely available models across different jurisdictions?

Since a major use of AI is the facilitation of easier content creation (videos, music, art, writing, etc.), those who live in more restrictive jurisdictions would find themselves in more of an uphill battle in the creation of content versus their less restricted peers. Meanwhile, content consumers would find themselves having to more consistently pay for media that they would otherwise be able to generate or find for free. Companies would also employ AI to dispense with workers in less restrictive jurisdictions, which may or may not give them a competitive edge over companies forced to retain workers that cannot be replaced in stricter regions. Also, greater access to free, open source models would mean that the public would not have to fork over $50 or so across five different software subscription packages.

(14.) Which are the most severe, and which the most likely risks described in answering the questions above? How do these set of risks relate to each other, if at all?

Probably the greatest risk is that in enabling easy access to free, open source AIs, more companies would be able to easily justify laying off employees than if they had to pay a lot of money to manage disparate subscriptions, anyway. However, the democratization of production that AI would engender (especially given how much of our society is predicated on a service economy) could mean that if we play our cards right, less people will have to be dependent on a corporate structure to acquire income in the first place. They could be more entrepreneurial, or else use AI in such a way as to save money over the long term and thus require less accumulation of capital.

To keep things consistent, I would argue that the large majority of AI-oriented reforms should be legislated on the federal level. This way, there would be more consistency in how businesses can assume they should be operated across the country.

(15.) What benefits do open model weights offer for competition and innovation, both in the AI marketplace and in other areas of the economy? In what ways can open dual-use foundation models enable or enhance scientific research, as well as education/training in computer science and related fields?

Open source models would allow for the costs of developing AI to be spread across entire communities, reducing the need to find vast reserves of upfront capital. Closed source models, as traditionally developed, are reliant upon corporate investments whose capital inflows come from high-end investors. Keeping AI development restricted to these people's hands would be undemocratic because in this system, money would talk incredibly loudly, and corporations are hierarchical organizations that on average do not allow for the democratic flow of ideas within the organization.

Since many public schools see difficulty in acquiring substantial budgets, and families themselves are often short on funds, free resources like open source AIs would be invaluable in promoting harmonization between artificial intelligence and learning. Courses teaching coding would especially benefit from easy access to AI, given that AI-assisted coding is becoming an industry-standard methodology. To deny future software engineers properly educative experiences in school because said school cannot afford the right closed source packages would be most regrettable.

(16.) How can making model weights widely available improve the safety, security, and trustworthiness of AI and the robustness of public preparedness against potential AI risks?

By allowing for the free distribution of open source AI systems and their models, many more people can experiment with the development of AI than what would exist if we were reliant on closed source models. Rather than having to be hired on by certain organizations to gain access to proprietary code (that may not represent competitors', anyway), hobbyists would be able to build skills in developing and maintaining artificial intelligences without having to conform to others' desires. By being afforded such freedoms, the types and quantities of AI that are developed would explode in number compared to what restrictive, private organizations might mandate.

Another advantage in making open source AIs freely distributed is that the progress of AI could be more easily monitored by the public, preparing us more generally for large advances as opposed to waiting for random corporations to spring them on us. Closed source models, those developed in secret, tend to be announced to the public without proper warning, catching many people off-guard. The courses that open source models take are much easier to track, given that they are generally developed by dozens of people on public repositories over the course of years. That is not to say that corporate entities cannot pick up an open source model and perform their own surprise developments, but it would not be the *only* way by which progress could occur.

(17.) Could open model weights, and in particular the ability to retrain models, help advance equity in rights and safety-impacting AI systems (e.g. healthcare, education, criminal justice, housing, online platforms etc.)?

Yes, open source AIs are far better suited to all of these tasks than closed source alternatives. By allowing people from all different backgrounds and incomes to utilize and develop AI, it is my firm belief that most would use their time and skills to create programs that help society rather than hinder it. Bad actors would find their efforts thwarted by a much greater share of people who would not at all be content to allow hateful, deceptive information to spread without counter.

Greater democratization of content creation would also allow for voices that have traditionally gone unheard to make their statements known via video, music, literature, film, etc. Traditionally, our society's creatives have had to find employment under corporate structures that have been antagonistic toward messages that break from the norm; any system that would free various forms of expression from such profit-driven environments will work to enrich the United States' social fabric and political awareness.

A beneficial application of locally hosted, open source AIs specifically can be found in chatbot roleplays of the sort performed in such applications as SillyTavern using LLMs. We have reached a point at which reasonably competent LLMs can be hosted on users' machines, and people have used these models to generate characters with whom they can speak about topics that they feel uncomfortable sharing with other humans. These conversations span a large range of topics, including those that some may find "objectionable," such as gory or pornographic roleplays. However, they are utterly harmless in the vast majority of instances – most people would never act upon these fantasies in real life. Rather, these discussions between humans and AIs can build a reclusive person's social skills in a context where without the AI (or with a severely compromised one that could not properly respond to the controversial topics that the user brings up), that person would emotionally retreat further inward.

Sometimes people need someone to talk to for therapeutic reasons, even if that someone is of a debatable level of sapience at the moment. It would not do these people well if their conversation partners were censored to such an extent that they could not hold an engaging conversation. Further, it is imperative that we as a society ensure that language models remain easily run on the local level because that will reduce the likelihood of any secrets divulged in these roleplays from being recorded by sinister actors online – ergo, we must encourage the proliferation of open source AI models. I know that *I* would not feel comfortable discussing sensitive topics with an AI hosted on an online database, yet most closed source AIs would force this sort of experience on their users. Not to mention, if a closed source AI is hosted online and is subsequently taken down, then users who came to enjoy the simulated conversations they held with that AI could suffer emotional distress with no means of rectifying it.

As one last point, we must not forget the existence of copyright. This outdated system of intellectual ownership has often led to many programs that could remain very useful becoming obsolete over decades' time due to their getting abandoned by their owners. This has been observed to be the case for pieces of software across multiple industries. If OpenAI developed a closed AI that they for some reason decided to cease development for, no one but they would be able to modify it until its copyright ran out. This would mean that anyone who wished to develop a similar AI would have to create their own alternative, and even then, *that* pursuit might get stymied via patents if OpenAI's programs utilize protected methodologies.

Open source AIs are extremely liberal – there are almost no scenarios in which someone could be barred from updating a deprecated open source AI if they so chose. This freedom would mean that software developers could extend the shelf lives of open source AIs for a far longer period than is possible regarding closed source AIs. Extended shelf lives for old programs would be a fantastic boon for any under-funded organization that heavily integrate old AIs into their workflows and find it too difficult to upgrade to newer programs. Any improvements that they made to their own manifestations of the program could then be shared to everyone else who might benefit from them, rather than be kept secret out of fear that the copyright holder would sue them.

(18.) How can the diffusion of AI models with widely available weights support the United States' national security interests? How could it interfere with, or further the enjoyment and protection of human rights within and outside of the United States?

Open source AIs that lower the barrier of entry to coding and that do not cost large sums of money to acquire will allow for the development of programs that could be used to directly counter those crafted by bad actors. Rather than keeping security developments in the hands of private organizations that must always be contracted at a cost, the national government will be able to pull from a much grander pool of their citizenry to not only engineer defensive measures, but also potential counterattacks.

Even *individuals'* digital security may be improved by their ability to engineer and distribute better protections against increasingly advanced forms of malware developed by rogue states or other types of organizations. I have no doubt that programs intended to compromise computers will increasingly be reliant upon AI to operate, or else benefit from AI coding assistance in their construction. It would only make sense, then, to allow the general public similar tools in combating these new generations of malicious programs.

On the cultural end of things, open source AIs downloaded and utilized by oppressed peoples could allow for greater content generation containing messages opposing oppressive regimes the world over to a greater degree than what could otherwise be easily created. Also, they could be used to develop programs that circumvent censorious, digital restrictions that these oppressive regimes attempt to enforce. Many closed source AIs would not be distributed to these peoples because private companies are seldom motivated to directly oppose censorious polities. At the very least, they seldom do so without seeing a clear benefit for themselves, and this motivator should not be relied upon.

(19.) How do these benefits change, if at all, when the training data or the associated source code of the model is simultaneously widely available?

As previously mentioned, extremely modifiable source code of various programs would mean that they could be quickly iterated upon and expanded to suit a population's needs independent of corporate mandates and/or the desire to make profit. Generalized training data that is known to have gone into the AI could be narrowed to more accurately conform to the principles and values of disparate communities. It is impossible for any central authority to create an AI that would make everyone happy – the better approach is to give communities the tools to customize open source AIs to accommodate their needs as is deemed fit. If these communities decide to use their customized AIs to break laws, then they can be disciplined for violating those actions. But that was because *those* people were dangerous – not the AI that merely served as a tool, and not the millions of other people utilizing it who have not been proven guilty of any crimes.

Now to comment on training data being public or not, I will say that I don't think the training data *have* to be made public to know whether or not an AI is good to use. There are a handful of open source AIs that have kept their training sets opaque, but still allow for users to iterate their own training data atop the baseline. If I want to train Stable Diffusion on how to generate anime cats of a certain style, for example, I can easily do so without peering into the grand training set that creates the base Stable Diffusion model. The potential risk of mandating that training data be kept transparent is that it both could adversely affect smaller AI model developers who wish to keep a certain competitive edge over other developers, and these developers might receive harassment if any of their data are deemed to be "problematic" by some portions of the public.

(20.) What should the role of model hosting services (e.g. HuggingFace, GitHub, etc.) be in making dual-use models with open weights more or less available? Should hosting services host models that do not meet certain safety standards? By whom should those standards be prescribed?
In an ideal world, I believe that these services should not be pressured into verifying how "safe" a model is. As I have said, advanced, open source AI models have the potential to do just as much good as they do harm, and the vast majority of people downloading them would either put them to use in neutral pursuits or those aimed to benefit a community. Too many regulations would limit the scope of what good could be done just as much as it would limit the bad…or perhaps the good would be more severely limited because bad actors will always find ways to circumvent any regulations placed from on high. Especially since the genie is out of the bottle, so to speak, and there are already plenty of AI models that have proliferated throughout the internet. Reining them all in will prove to be impossible.

One of my big concerns is that if standards were placed on content generation, they would be crafted in favor of current social norms at the expense of allowing individuals to express more niche perspectives. I hope that none of these standards would involve censorship along the lines of barring models that allow for pornographic, violent, or hateful content because attempting to ban these sorts of AI generations would not only restrict people's speech, but they could also lead to AI models' performance being undercut.

That is not to say that I believe that the *distribution* or *possession* of certain sorts of generations shouldn't be made illegal, however. If someone generates and distributes realistic child pornography, even if it is artificial, its presence on the internet would still be unconscionable because of how much it would confuse efforts on the part of law enforcement to determine which instances of realistic pornography represent real children caught in abusive situations. There also exists the possibility that real likenesses could be used in the generation of such pornography, or even adult pornography – these should all constitute crimes because they are utilizing people's likenesses without permission for compromised purposes (and in the former case, it would indirectly comprise exploitation of children). However, I must once again stress that even though the AI would facilitate the creation of these images and/or videos, the actual criminal would be the *human* who input the prompts. Getting rid of the AI would be like getting rid of Photoshop, Audacity, or Blender because someone could use that program to generate illicit material.

(21.) Should there be different standards for government as opposed to private industry when it comes to sharing model weights of open foundation models or contracting with companies who use them?
I would argue that every AI model that the government commissions and shares to the public be released under the GPLv3 license. This license was specifically created to allow for the free and fair distribution of software packages such that any user can:

- Copy the software.

- Redistribute the software.
- Alter the software.
- Use the software for commercial or non-commercial purposes.

All this, so long as developers follow the terms set out in the license mandating that they provide proper credit to the application's creators and attach the same terms/conditions to their versions of the software. This approach to distribution would be a fantastic way to make the government's open models as easily available as possible while not entailing that the government totally relieve their copyright (which is necessary to keep if one wishes to punish those who violate the license's terms).

Adopting such an open license in the distribution of AI models is a measure that most corporations would never employ – or at the very least, they couldn't be trusted to remain consistent about it. There is constant threat in the open source AI community of companies who previously distributed open source models deciding to make a particularly advanced revision closed source so as to generate maximum profit rather than allow the community to benefit. If the United States' federal and state governments make it a point to always make their models as easily distributed as possible, then the general public will not have to fret the rug getting pulled out from under them by policy shifts and will instead be able to consistently plan long-term. It wouldn't even just be lower income people who would benefit from this sort of policy; if a company not only makes an AI closed source, but only ever licenses future revisions out to large companies, then the whole public has now been deprived of a resource that they may have come to depend on.

(22.) What should the U.S. prioritize in working with other countries on this topic, and which countries are most important to work with? What insights from other countries or other societal systems are most useful to consider?

There should be international laws that are dedicated to closely regulating closed source companies whose artificial intelligences are powered by models that are not publicly disclosed. Even companies using open source AIs should be watched via legislation that much more aggressively protects the data of those using said AIs, instead of allowing for undisclosed quantities of data to be kept in those companies' online databases for nebulous ends. This effort should be international in scope and coordinated with all our allies so as to hold companies/organizations/individuals who interface with our populaces' data accountable on the regular. The digital age has made it so that many of the services that people utilize in the United States are provided by organizations that are headquartered out of our borders – proper enforcement of data regulations is only possible insofar as we have those host countries' cooperation.

We should strive to make the acquisition and utilization of open source AIs as close to a universally entitled right as we can because these pieces of software will quickly become a key tenet of society, much as interfacing with the internet and smartphones have become practical necessities for modern life. We dropped the ball in making these latter two technical advances easily available to all of our citizens and it has resulted in very exploitative practices on the part of internet service providers (ISPs) and tech companies that take advantage of captive markets, or else ignore regions that they consider to be unprofitable. I am not proposing that every citizen needs a workhorse of a computer (as AI models continue to get smaller, these will likely become less and less of a requirement, anyway), but there should be research put into optimizing our existing open source models and accommodating their distribution over as wide an area as possible. My earlier suggestion of a sort of nationalized equivalent to Google Cohost would allow for the best short-term compromise in providing millions of people a convenient means of utilizing advanced, open source AI models.

(23.) Are there particular individuals/entities who should or should not have access to open-weight foundation models? If so, why and under what circumstances?

As has likely been made extremely evident by this point, I do not believe that anyone should be barred from utilizing open source AIs. Any crimes that various groups and individuals may commit should be treated as crimes committed by those people, not crimes committed by the AI. The debate is not equivalent to that revolving around guns as those are weapons solely designed for killing; artificial intelligences are neutral programs that should not even partially shoulder the burden of a bad actor's guilt.

(24.) In the face of continually changing technology, and given unforeseen risks and benefits, how can governments, companies, and individuals make decisions or plans today about open foundation models that will be useful in the future?

There should be many instances in which we as a society come together via open debate to discuss whatever policies may need to be instituted to address the seismic shifts that our world will inevitably undergo in the coming decades. Even if open source models were totally, formally banned in the United States, closed and open source models would still be developed (either in this country or another by organizations that did not wish to play by our rules), and their deployment will increasingly put many tenets of society that were once thought secure at risk - employment serving as the biggest one. Our society has a brilliant opportunity before us, in that we could utilize artificial intelligence in a way that allows for the general population to gain more fruits from less labor, potentially moving our economy away from a grind for ever more capital, subsequently increasing the general population's wealth and quality of living.

However, both the federal and state governments must remain actively vigilant, observing the routes that various AIs and their developers steer our economy down as we turn into them, not well after the fact. That way, we can legislate *with* the flow of change as opposed to feeling that we have to ungraciously halt these changes by, say, banning wide swathes of open source AIs. Any negative developments pertaining to AIs' deployment in the private sector could be countered via proper legislation in the public sector. Even if AIs eliminate many private jobs, the federal government institute universal employment, which would guarantee everyone a stable job at a comfortable living wage. Socialist securities such as this would become extremely important as time goes on: more and more jobs in the private sphere will get displaced by an assortment of AIs, so it will be up to society in general to pass more compassionate policies to make up for that.

As this situation continues to unfold over the course of years, people will increasingly perceive AI as a threat to their ability to survive in this world - a perception that is certainly only helped along by the plethora of dystopic fiction that posits the end state of a humanity that embraces AI is to fall under AI's heel. These fears would reach a fever pitch if artificial general intelligence (AGI) is developed, at which point many jobs would no longer be dominated by humans because some dominant theories of AGI posit an AI sharing the same level of cognizance as we humans (or perhaps they would attain an even higher level. The theories tend to get a little disorganized because there is debate over just what is possible and what is not). However, if humans are no longer faced with the need to compete for jobs to survive, then we will feel less antagonized by AIs as a whole, whether those AIs be human-level or lower. This is not to suggest that we would hold hands and begin singing songs, but the fears that AI would directly challenge people's ability to survive even in a "peaceful" context would abate some.

Less antagonism toward AI eliminating jobs would allow for a much more level-headed public to assess AIs and their potential risks/benefits. Combined with educating the populace to employ proper critical thinking skills encouraging the assessment of a situation before resorting to emotional appeals like fear, the path laying before us could be paved with long-term, harmony between AIs and humans could be achieved without fears of existential threat constantly impeding progress and peace of mind.

I will step slightly out of the scope of the question to clarify that some of this rhetoric is born from a humanitarian concern toward hypothetical AGIs. If we do develop AIs with human consciousnesses, then I fear that if we do not take preemptive measures to reduce prejudices against them, these hypothetical individuals will be abused by humans to such an extent that our coexistence will be severely threatened. This would be a shameful outcome that would gradually work toward the deterioration of our future society. As such, I strongly encourage governments the world over to seriously consider the potential future in which humans find their peers in AIs, and work to make sure that once this sort of technology is created, these AIs are given equivalent rights to humans. If we fail to prepare ourselves for this hypothetical outcome, then we will be caught wildly off-guard if it manifests, and clumsy decisions might be made that will be heavily criticized by later generations as a point of shame.

(25.) What other issues, topics, or adjacent technological advancements should we consider when analyzing risks and benefits of dual-use foundation models with widely available model weights?
We cannot predict when or if AIs' capabilities will cease – I have already made it clear that I believe we will achieve AGI of a human-like cognizance someday, and I am rather concerned over how they will be treated if they are created. In the interim, it is best if we democratize AI development as much as possible, allowing for fundamental knowledge of AI systems (how to build, maintain, and distribute them) to reach as many people as possible, early as possible. We do not want to create a slow start for ourselves, lagging behind proprietary AI models for an extended period of time.

Yes, there are risks attached to free experimentation. At the same time, I urge the reader to think of how many skilled developers exist who never find the right opportunities to get hired by major companies, yet who could do so much good in advancing the various fields of AI research. The pursuits that they put their skills toward will be motivated by a wealth of factors stretching beyond profit; by contributing to the development of open source AIs, these developers will improve pieces of software that future generations will be able to further iterate upon. There are more decent people in this world than bad. Therefore, the strides that most take in developing AIs would largely be put to use in benefiting the common good. Limiting this potential good out of fear will not do our society well.

I, and I believe many others, would certainly find these open projects to be far more trustworthy than the grand majority of closed off, private AIs. The organizations that create them have not earned my "trust" at all, as there are too many corruptible, materialistic reasons for their involvement in the industry. Even if private individuals developing open source AIs cannot be held socially accountable in the same way that a government ostensibly can, the fact that anyone could fork these projects means that any AI that was ever repurposed for subversive goals could be remade and redistributed in its original form again (or at least an approximation), maintaining continuity with what the community desires/expects. Meanwhile, closed source AIs distributed by private entities could see older versions entirely dropped by their creators. And that nasty little bit about copyright mentioned earlier would mean that further development, or even *distribution* of these older versions, could be unjustly obstructed. If it was an AI that a lot of society was reliant upon, these companies would possess the power to make many people's lives much more difficult overnight.

If we wish to maintain a fair and free society, then we should not create a situation in which some of our citizens who have merely achieved a certain degree of social cachet and capital gain access to certain tools while others are left in the dust. Whether imposed by government regulations or corporate profit-seeking, elitism should be evaded at all costs. The United States is supposed to be *for* the people, *by* the people. Ergo, the "people" composed primarily of individuals unaffiliated with grandiose levers of power deserve some trust as we move into the future.