As AI technology grows in both power and popularity, the question of regulation is a natural one. However, as an information technology, any attempt to restrict the production and distribution of AI models is ultimately an attempt to restrict and control the sharing of information. These models are only ones and zeros on a hard drive, and can be freely shared infinitely by anyone with access to a computer and the internet. Restricting these models based on any quantitative factor such as file size, training computation, or distribution scale, is like trying to cut the head from a hydra. The technology can easily be redesigned to morph around these limitations, maintaining the constant pace of evolution while still fitting within the letter of the law. Reactive adjustments to the law will always be a step behind the cutting edge of the machine learning community.  
  
I believe that such quantitative restrictions miss the mark, and will not truly provide the safety that regulation seeks to ensure. The raw size, power, and distribution of a model are not what make it inherently dangerous. The information that a model is trained on is the primary determining factor in how a model will behave. For example, a model that is trained on a dataset that includes instructions on the production of dangerous weapons, or other nefarious acts such as hacking and scamming, will be a dangerous model. This is true regardless of model size and intelligence; even a relatively small model that isn’t powerful or state-of-the-art can be dangerous when trained in such a way. Conversely, a very large and powerful model that has been trained on a well curated dataset, with careful attention paid to remove “dangerous” instructional information, will not be a dangerous model.  
  
As training datasets are often hundreds of billions of words long, any attempt to regulate the content of training datasets seems to be a fool’s errand. As technology progresses, the computational and skill requirements will continue to decrease, and the ability to train machine learning models will become easier to the point of triviality. Eventually anyone will be able to train their own models at home. We are already to the point that skilled programmers can do this right now. AI models cannot realistically be regulated by limiting their size, computational power, or distribution, and any attempt to do so will only retard the growth of the USA as this revolutionary technology changes the world around us.

Models can only be judged and regulated after the fact, based on their actions. If a user of a specific model has been shown to cause illegal harm with a model, then they can be punished or sanctioned within the framework of our existing law, based on the crimes that they have committed. Then once the model has been demonstrated to be dangerous, it can be regulated on a case by case basis. Restricting models before they have been produced, based on arbitrary quantitative limits to size, power, and distribution, will only restrict the technological development of the country, without providing any real additional safety.  
  
In conclusion, while the impulse to regulate AI models is understandable given their increasing power and potential for misuse, the proposed regulations based on quantitative factors such as size, computational power, and distribution are misguided. These restrictions would be ineffective in practice, easy to circumvent, and would only serve to hinder the United States' progress in this transformative field. Instead, we should focus on promoting responsible AI development practices, particularly in the curation of training datasets, and address harmful applications of AI models through our existing legal framework on a case-by-case basis. By taking a more targeted and adaptive approach to AI regulation, we can foster innovation while still protecting society from the potential risks of this technology. Preemptive quantitative restrictions, however well-intentioned, are not the answer.