



# Regulating Machine Learning Open-Source Software

A PRIMER FOR POLICYMAKERS

**MAY 2024**

→ James Ostrowski<sup>1</sup>

## Introduction

In February 2023 Meta open sourced the code for LLaMA, its large language model (LLM). A week later, the model weights were leaked, which allowed open-source developers to replicate and build novel

---

<sup>1</sup> James Ostrowski is a JD candidate at the University of Washington School of Law.

applications from it.<sup>2</sup> In June, Senators Josh Hawley and Richard Blumenthal sent Mark Zuckerberg a bipartisan letter excoriating Meta for releasing LLaMA to the public: “Meta effectively appears to have put a powerful tool in the hands of bad actors.”<sup>3</sup> Jan Leike, the alignment lead at OpenAI, has written, “An important test for humanity will be whether we can collectively decide not to open source LLMs that can reliably survive and spread on their own. Once spreading, LLMs will get up to all kinds of crime.”<sup>4</sup> Yet machine learning open-source software (MLOSS) promotes innovation, decentralizes and commoditizes the power of AI, and improves security. The global value created by MLOSS was \$30 billion in 2022 alone.<sup>5</sup>

Still, the potential for misuse stoked a movement to regulate artificial intelligence (AI) and MLOSS. Policymakers, advocacy organizations, and businesses fear the externalities that come with a lack of government control over a powerful technology. As a result, they seek to control who creates AI systems and how they are created.

AI regulatory proposals fit into one of two categories: (1) regulation of AI’s development and (2) regulation of its deployment. Regulation aimed at development holds liable anyone developing or considering developing AI systems for downstream uses of these models. Regulation of deployment, on the other hand, involves controlling risky and harmful applications by

---

2 Ian Brown, “Expert Explainer: Allocating Accountability in AI Supply Chains,” Ada Lovelace Institute, June 29, 2023, <https://www.adalovelaceinstitute.org/resource/ai-supply-chains/>.

3 Richard Blumenthal and Josh Hawley to Mark Zuckerberg, June 6, 2023, <https://www.hawley.senate.gov/sites/default/files/2023-06/Hawley-Meta-LLAMA-Letter.pdf>.

4 Jan Leike, post on X, August 9, 2023, 11:45 a.m., <https://twitter.com/janleike/status/1689301898367787008>.

5 Max Langenkamp and Daniel Yue, “How Open Source Machine Learning Software Shapes AI,” AIES ’22, July 2022, <https://dl.acm.org/doi/10.1145/3514094.3534167>.

users. Regulating deployment shifts the liability burden from the developer to the user.

Under a development-focused regulatory regime, a nonprofit that built an image generation model would be required to undergo licensing, permitting, premarket approvals, red-teaming, or algorithmic audits before open sourcing its model. Under a deployment-focused regulatory regime, the person using the open-source model would be held liable for illegal applications, such as the nonconsensual creation or dissemination of sexual imagery.<sup>6</sup>

This paper outlines the state of MLOSS, catalogs its potential harms, and analyzes competing regulatory proposals. I argue that the regulation of AI development will stifle the machine learning open-source community by concentrating power in the hands of a few large companies. Policymakers should seek to regulate AI at the deployment level. This approach will allow MLOSS to thrive while adequately addressing concrete harms and holding accountable the actors most directly responsible for them.

## Defining Machine Learning Open-Source Software

There is no consensus definition for *machine learning* or *open source*.<sup>7</sup> For this paper, I define *machine learning* as a system that uses machines to improve its performance in approximating

---

6 S. 1042A, 2023–2024 Leg., Reg. Sess. (N.Y. 2023).

7 Ryan Calo, “Artificial Intelligence Policy: A Primer and Roadmap,” *University of California, Davis Law Review* 51, (2017): 399, [https://lawreview.law.ucdavis.edu/issues/51/2/symposium/51-2\\_Calo.pdf](https://lawreview.law.ucdavis.edu/issues/51/2/symposium/51-2_Calo.pdf); David Gray Widder, Meredith Whittaker, and Sarah Myers West, “Open (for Business): Big Tech, Concentrated Power, and the Political Economy of Open AI” (working paper, August 16, 2023), [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4543807](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4543807).

some aspect of human cognition.<sup>8</sup> Five components are essential to machine learning systems: development frameworks, computational power, data, labor, and models.<sup>9</sup> Practitioners rely on open-source *development frameworks*, such as PyTorch and TensorFlow, to construct and train machine learning models. Training models on large data sets requires *computational power*. (Technologists estimate that training an LLM such as GPT-3 could cost over \$4 million.)<sup>10</sup> High-quality, accurately labeled *data* are essential to performance. *Labor* includes data labeling, reinforcement learning with human feedback, engineering, and product development. The *model* is the algorithmic system that produces the final output.<sup>11</sup>

Machine learning software is considered open source when it (1) provides access to source code, documentation, and data; (2) is licensed to allow third parties to use the software; and (3) is situated to allow third parties to build off the software.<sup>12</sup> Openness exists on a spectrum.<sup>13</sup> Maximally open machine learning software will check all three boxes. For example, nonprofit EleutherAI's GPTNeo provides source code, training data, full documentation, and a broad usage license. Other developers may offer only API access, which allows authorized users to receive specific information about a model, rendering their "open source" claims little more than fanciful marketing. Figure 1 shows how openness varies among AI projects.

---

8 Calo, "Artificial Intelligence Policy."

9 Widder, Whittaker, and West, "Open (for Business)."

10 Jonathan Vanian and Kif Leswing, "ChatGPT and Generative AI Are Booming, but the Costs Can Be Extraordinary," *CNBC*, March 13, 2023, <https://www.cnbc.com/2023/03/13/chatgpt-and-generative-ai-are-booming-but-at-a-very-expensive-price.html>.

11 Widder, Whittaker, and West, "Open (for Business)."

12 Widder, Whittaker, and West, "Open (for Business)."

13 Irene Solaiman, "The Gradient of Generative AI Release: Methods and Considerations," preprint, February 5, 2023, <https://arxiv.org/pdf/2302.04844.pdf>.

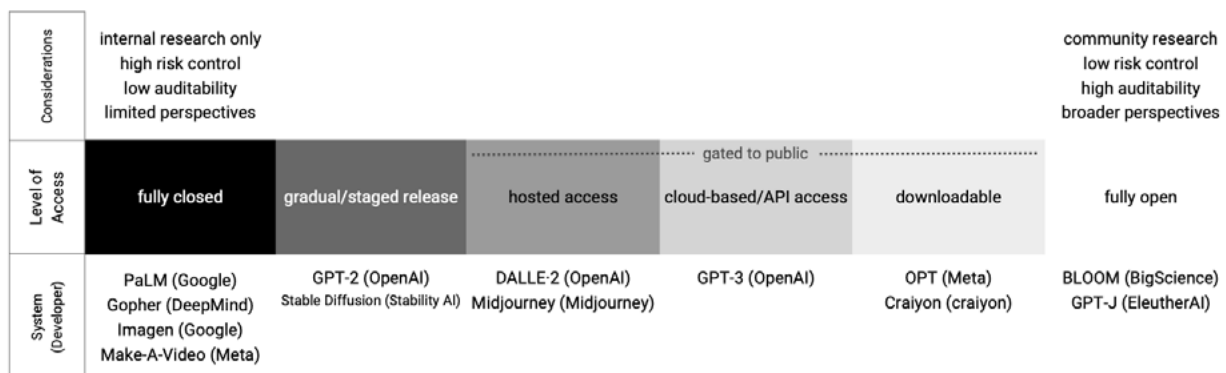
FIGURE 1 | How Open and Transparent Are 2023 Instruction-Following Text Generators?

Project <small>(maker, bases, URL)</small>	Availability					Documentation					Access methods		
	Open code	LLM data	LLM weights	RLHF data	RLHF weights	License	Code	Architecture	Preprint	Paper	Data sheet	Package	API
chatGPT	✓	✓	✓	✓	✓								
OpenAI						LLM base: GPT3.5, GPT4							<a href="https://chat.openai.com">https://chat.openai.com</a>
StableVicuna-13B	✓	✓	✓	✓	✓								
CarperAI						LLM base: LLaMA	RLHF base: oasst1, anthropic						<a href="https://huggingface.co/CarperAI/stable-vicuna-13b-delta">https://huggingface.co/CarperAI/stable-vicuna-13b-delta</a>
text-generation-webui	✓	✓	✓	✓	✓								
oobabooga						LLM base: various	RLHF base: various						<a href="https://github.com/Akegarasu/ChatGLM-webui">https://github.com/Akegarasu/ChatGLM-webui</a>
MPT-7B-Instruct	✓	✓	✓	✓	✓								
MosaicML						LLM base: MosaicML	RLHF base: dolly, anthropic						<a href="https://github.com/mosaicml/llm-foundry#mpt">https://github.com/mosaicml/llm-foundry#mpt</a>
Falcon-40B-Instruct	✓	✓	✓	✓	✓								
TII						LLM base: Falcon 40B	RLHF base: Balze (synthetic)						<a href="https://huggingface.co/tiiuae/falcon-40b-instruct">https://huggingface.co/tiiuae/falcon-40b-instruct</a>
minChatGPT	✓	✓	✓	✓	✓								
ethanyanjali						LLM base: GPT2	RLHF base: anthropic						<a href="https://github.com/ethanyanjali/minChatGPT">https://github.com/ethanyanjali/minChatGPT</a>
trlx	✓	✓	✓	✓	✓								
carperai						LLM base: various (pythia, flan, GPT)	RLHF base: various						<a href="https://github.com/carperai/trlx">https://github.com/carperai/trlx</a>
stanford_alpaca	✓	✓	✓	✓	✓								
Tatsu labs						LLM base: LLaMA	RLHF base: Self-Instruct (synthetic)						<a href="https://github.com/tatsu-lab/stanford_alpaca">https://github.com/tatsu-lab/stanford_alpaca</a>
Cerebras-GPT-111M	✓	✓	✓	✓	✓								
Cerebras, Schramm						LLM base: not open	RLHF base: alpaca (synthetic)						<a href="https://huggingface.co/SebastianSchramm/Cerebras-GPT-111M-instruction">https://huggingface.co/SebastianSchramm/Cerebras-GPT-111M-instruction</a>
OpenChatKit	✓	✓	✓	✓	✓								
togethercomputer						LLM base: EleutherAI pythia	RLHF base: OIG						<a href="https://github.com/togethercomputer/OpenChatKit">https://github.com/togethercomputer/OpenChatKit</a>
dolly	✓	✓	✓	✓	✓								
databrickslabs						LLM base: EleutherAI pythia	RLHF base: databricks-dolly-15k						<a href="https://github.com/databricks/dolly">https://github.com/databricks/dolly</a>
CharRWKV	✓	✓	✓	✓	✓								
BlinkDL						LLM base: RWKV-LM (own)	RLHF base: alpaca, shareGPT (synthetic)						<a href="https://github.com/BlinkDL/ChatRWKV">https://github.com/BlinkDL/ChatRWKV</a>
BELLE	✓	✓	✓	✓	✓								
LianjiaTech						LLM base: LLaMA, BLOOMZ	RLHF base: alpaca, shareGPT (synthetic)						<a href="https://github.com/LianjiaTech/BELLE">https://github.com/LianjiaTech/BELLE</a>
Open-Assistant	✓	✓	✓	✓	✓								
LAION-AI						LLM base: oasst1 (own)	RLHF base: OIG						<a href="https://github.com/LAION-AI/Open-Assistant">https://github.com/LAION-AI/Open-Assistant</a>
xmif	✓	✓	✓	✓	✓								
bigscience-workshop						LLM base: BLOOMZ, mT0	RLHF base: xP3						<a href="https://github.com/bigscience-workshop/xmif">https://github.com/bigscience-workshop/xmif</a>

Source: Andreas Liesenfeld, Alianda Lopez, and Mark Dingemans, “Opening Up ChatGPT: Tracking Openness, Transparency, and Accountability in Instruction-Tuned Text Generators,” in Proceedings of the 5th International Conference on Conversational User Interfaces (CUI’23), Eindhoven, Netherlands, July 19–21, 2023 (New York: Association for Computing Machinery, 2023), <https://doi.org/10.1145/3571884.3604316>.

Openness comes with important trade-offs, and these drive developers’ decisions about access. Figure 2 shows that as systems become more open they provide more utility to the community of practitioners, but developers lose control of their systems’ use.

FIGURE 2 | Considerations for Different Kinds of Artificial Intelligence System Access



Source: Irene Solaiman, “The Gradient of Generative AI Release: Methods and Considerations,” preprint, February 5, 2023, <https://arxiv.org/pdf/2302.04844.pdf>.

OpenAI cited concerns about control when it moved from an open system to a closed one. At its inception in 2015, OpenAI was a nonprofit dedicated to “openness,” as evidenced by its name. By 2019, it had become a capped-profit company with a \$1 billion investment from Microsoft, to whom it would exclusively license its GPT-3 model.<sup>14</sup> OpenAI cofounder Ilya Sutskever said, “At some point it will be quite easy, if one wanted, to cause a great deal of harm with those models. And as the capabilities get higher it makes sense that you don’t want to disclose them.”<sup>15</sup> The organization’s most recent GPT-4 report contained no details about the model’s architecture, hardware, computational power, data set, or training method.<sup>16</sup> Sutskever got his closure.

## History and Current Trends

The histories of open source and machine learning only recently converged. The open-source movement began in 1985 with the Free Software Foundation.<sup>17</sup> This organization provided access to programs at a time when code bases were primarily developed by closed corporations. Today, open-source software forms the bedrock of the US economy. One study that reviewed more than 1,700 code bases across 17 industries found that 96 percent of code bases contained open-source code.<sup>18</sup> And 100 percent of software from the aerospace, aviation, automotive, transportation, and

---

14 Widder, Whittaker, and West, “Open (for Business).”

15 Quoted in Widder, Whittaker, and West, 17.

16 OpenAI, “GPT-4 Technical Report,” March 27, 2023, <https://cdn.openai.com/papers/gpt-4.pdf>.

17 Free Software Foundation history page, accessed December 14, 2023, <https://www.fsf.org/history/>.

18 Synopsys, “2023 Open Source Security and Risk Analysis Report,” 2023, <https://www.synopsys.com/software-integrity/resources/analyst-reports/open-source-security-risk-analysis.html>.

logistics sectors contained open-source code.<sup>19</sup> AI entrepreneur Jeremy Howard explains,

*Today, nearly every website you use is running an open source web server (such as Apache), which in turn is installed on an open source operating system (generally Linux). Most programs are compiled with open source compilers, and written with open source editors. . . . [M]uch of the world of computers and the internet that you use today would not exist without open source.*<sup>20</sup>

Though MLOSS is born of this tradition, it differs from traditional open-source software in important ways. For traditional software, access to the source code and documentation can be sufficient to democratize access. Even the most open machine learning software, however, requires resources that cannot be open sourced. Building a model from scratch requires expensive computing power that is inaccessible to most. AI researchers David Widder, Meredith Whittaker, and Sarah West write, “The resources needed to build AI from scratch, and to deploy large AI systems at scale, remain ‘closed’—available only to those with significant (almost always corporate) resources.”<sup>21</sup> Even maximally open MLOSS relies on closed and expensive services from large tech companies.<sup>22</sup>

Despite these constraints, MLOSS continues to grow. Before 2012, there were few high-quality MLOSS projects.<sup>23</sup> Academia produced several open-source development frameworks after a deep

---

19 Synopsys, “2023 Open Source Security and Risk Analysis Report.”

20 Jeremy Howard, “AI Safety and the Age of Dislightenment: Model Licensing & Surveillance Will Likely Be Counterproductive by Concentrating Power in Unsustainable Ways,” fast.ai, July 10, 2023, <https://www.fast.ai/posts/2023-11-07-dislightenment.html>.

21 Widder, Whittaker, and West, “Open (for Business).”

22 AI Now Institute, “Democratize AI? How the Proposed National AI Research Resource Falls Short,” October 5, 2021, <https://ainowinstitute.org/publication/democratize-ai-how-the-proposed-national-ai-research-resource-falls-short>.

23 Langenkamp and Yue, “How Open Source Machine Learning Software Shapes AI.”

convolutional neural network won the 2012 ImageNet competition, which challenged algorithms accurately identify objects in a dataset.<sup>24</sup> Corporations responded by releasing their own open-source frameworks, such as Google's TensorFlow and Facebook's PyTorch.<sup>25</sup> Businesses have consistently benefited from providing their frameworks to the world for free. Meta CEO Mark Zuckerberg has noted that open sourcing the PyTorch framework made it easier to capitalize on new ideas developed externally.<sup>26</sup> Other businesses, such as Amazon and NVIDIA, may support open-source projects because these businesses have a general stake in increasing the adoption of AI.

In addition to the MLOSS projects driven by academia and commercial investment, there are several high-quality projects maintained by nonprofits. Nonprofit EleutherAI develops open-source AI models, providing maximally open documentation and licensing. EleutherAI operates by means of donations from CoreWeave, Hugging Face, Stability AI, the Mozilla Foundation, Google, former GitHub CEO Nat Friedman, and Lambda Labs.<sup>27</sup> This ecosystem of academic researchers, civil society, big corporations, and small companies churns out countless LLMs, with new ones emerging seemingly every week. Today, the vast majority of publicly available research code is written using open-source frameworks.<sup>28</sup>

MLOSS is as powerful as heavily funded, closed machine learning systems. A leaked Google memo states, "Open-source models are faster, more customizable, more private, and pound-for-pound more

---

24 Langenkamp and Yue, "How Open Source Machine Learning Software Shapes AI."

25 Langenkamp and Yue, "How Open Source Machine Learning Software Shapes AI."

26 Widder, Whittaker, and West, "Open (for Business)."

27 EleutherAI, about page, accessed December 15, 2023, <https://www.eleuther.ai/about>.

28 Sung Kim, "List of Open Sourced Fine-Tuned Large Language Models (LLM)," Medium, March 30, 2023, <https://sungkim11.medium.com/list-of-open-sourced-fine-tuned-large-language-models-llm-8d95a2e0dc76>.



capable. They are doing things with \$100 and 13B params that we struggle with at \$10M and 540B. And they are doing so in weeks, not months.”<sup>29</sup> But, though big tech companies readily open source their models today (e.g., Google’s BERT and OpenAI’s GPT-2), some researchers predict that the “incentives to release these models will diminish over time as they become more commercialized.”<sup>30</sup> There are already relatively few open-source models from nonprofit initiatives, leaving the field dependent on large technology companies. This could leave MLOSS a generation behind closed-source models.<sup>31</sup>

## Impact

AI improves access to and quality of health care, financial services, transportation, retail, agriculture, entertainment, energy, and aviation.<sup>32</sup> Open sourcing the technology behind these advancements accelerates innovation, decentralizes and commoditizes the power of AI, and improves security. As I noted earlier, the global value created by MLOSS was \$30 billion in 2022.<sup>33</sup> For every dollar invested in MLOSS tools, at least \$100 is created within the AI ecosystem.<sup>34</sup>

---

29 Google, “We Have No Moat, and Neither Does OpenAI,” published by Dylan Patel and Afzal Ahmad, *SemiAnalysis*, May 4, 2023, <https://www.semianalysis.com/p/google-we-have-no-moat-and-neither>.

30 Alex Engler, “The EU’s Attempt to Regulate Open-Source AI Is Counterproductive,” Brookings, August 24, 2022, <https://www.brookings.edu/articles/the-eus-attempt-to-regulate-open-source-ai-is-counterproductive/>.

31 Elad Gil, “AI Platforms, Markets, & Open Source,” *Elad Blog*, February 15, 2023, <https://blog.eladgil.com/p/ai-platforms-markets-and-open-source>.

32 Adam Thierer, “Will AI Policy Become [sic] a War on Open Source Following Meta’s Launch of LLaMA 2?,” Medium, July 18, 2023, <https://medium.com/@AdamThierer/will-ai-policy-became-a-war-on-open-source-following-metas-launch-of-llama-2-b713a3dc360d>.

33 Langenkamp and Yue, “How Open Source Machine Learning Software Shapes AI.”

34 Langenkamp and Yue, “How Open Source Machine Learning Software Shapes AI.”

Rather than spending millions of dollars to train a new model or paying large firms for restricted access, anyone with a laptop can use MLOSS to customize an open-source model in a few hours.<sup>35</sup> This speeds up the pace of innovation by enabling everyone to iterate on existing models. Many clever new applications are created by innovators who would not have access to machine learning software in a closed-source environment. For example, a small document-processing company might use open-source machine learning by using an open source object detection model like Dectron2, programmed in an open source framework like Python, and trained on an open source dataset like COCO.<sup>36</sup>

MLOSS speeds up individual tasks for researchers. PyTorch Lightning, an MLOSS tool, can save a researcher hours of debugging.<sup>37</sup> MLOSS also accelerates innovation by developing standards and building communities, which improve productivity and interoperability.

The Google memo I quoted earlier highlights the democratic impact of open source: “Who would pay for a Google product with usage restrictions if there is a free, high quality alternative without them? . . . The modern internet runs on open source for a reason. Open source has some significant advantages that we cannot replicate.” The memo continues, “The innovations that powered open source’s recent successes directly solve problems we’re still struggling with. Paying more attention to their work could help us to avoid reinventing the wheel.”<sup>38</sup> If open source can solve even Google’s

---

35 Bruce Schneier and Jim Waldo, “Big Tech Isn’t Prepared for A.I.’s Next Chapter,” *Slate*, May 30, 2023, <https://slate.com/technology/2023/05/ai-regulation-open-source-meta.html>.

36 Langenkamp and Yue, “How Open Source Machine Learning Software Shapes AI.”

37 Langenkamp and Yue, “How Open Source Machine Learning Software Shapes AI.”

38 Google, “We Have No Moat.”

technical problems, it is highly likely to provide utility for smaller actors with fewer resources.

Some companies accelerate decentralization by promoting open source to weaken and fragment their competition. For instance, in the 1990s IBM invested \$1 billion in the open-source operating system Linux.<sup>39</sup> This was not an altruistic donation in service of decentralization, though increased decentralization was an externality. Rather, it was a strategic move to counterbalance Microsoft's dominance in server-side software—a plan to weaken IBM's competitor. Former Microsoft CEO Steve Ballmer tried painting open source as evil because it threatened Windows.<sup>40</sup> Steven Weber, a professor at University of California, Berkeley, believes that OpenAI is running scared from open source “in the same way that Microsoft ran scared of Linux in the late 1990s.”<sup>41</sup>

MLOSS contributes to responsible development by increasing security and reducing bias. For instance, researchers have compared MLOSS models' reliability in generating code and measuring environmental impact. Other researchers have used open-source LLMs to show how models display bias and how they may be manipulated to produce otherwise censored results.<sup>42</sup> This practice of community vetting mirrors bug bounty programs that incentivize individuals, often security researchers, to identify and responsibly disclose software vulnerabilities. These programs improve the security and resilience of an organization's digital assets by leveraging the collective expertise of a diverse community of researchers.

---

39 Gil, “AI Platforms, Markets, & Open Source.”

40 Sharon Goldman, “Senate Letter to Meta on LLaMA Leak Is a Threat to Open-Source AI, Say Experts,” *VentureBeat*, June 8, 2023, <https://venturebeat.com/ai/senate-letter-to-meta-on-llama-leak-is-a-threat-to-open-source-ai-at-a-key-moment-say-experts/>.

41 Quoted in Goldman, “Senate Letter to Meta.”

42 Engler, “EU's Attempt to Regulate Open-Source AI.”

Today they are a crucial component of cybersecurity.<sup>43</sup> In fact, OpenAI has launched its own bug bounty program, albeit to improve security rather than to root out bias.<sup>44</sup>

## Calls for Regulation

The rapid development of machine learning, particularly LLMs, is prompting increasingly loud calls from government, civil society, and industry for regulation. Policymakers and civil society organizations see MLOSS as too powerful and too difficult to control. Incumbent companies may see it as a threat. The types of regulation proposed vary significantly, from an outright ban on machine learning to heavy top-down regulation to various state-law, common-law and soft-law approaches. All proposed AI regulations would affect MLOSS, though some seek to regulate open source directly while others focus on AI broadly, sweeping up MLOSS in the process.

As I noted earlier, the most serious of these calls focus on either (1) regulating development or (2) regulating deployment. To understand which proposal is more effective, we must start by examining the threats that these regulations seek to address.

## Perceived Harms

Proponents of tighter controls on open source argue that AI is too dangerous to be in the hands of the public. In their letter to Meta, Senators Blumenthal and Hawley criticize the

---

43 Microsoft, "Microsoft Bug Bounty Program," accessed December 15, 2023, <https://www.microsoft.com/en-us/msrc/bounty>.

44 OpenAI, "Announcing OpenAI's Bug Bounty Program," April 11, 2023, <https://openai.com/blog/bug-bounty-program>.

company for releasing its LLaMA model in such an “unrestrained and permissive manner.”<sup>45</sup> Security researchers Bruce Schneier and Jim Waldo explain, “Having the technology open-sourced means that those who wish to use it for unintended, illegal, or nefarious purposes have the same access to the technology as anyone else.”<sup>46</sup> OpenAI cofounder Ilya Sutskever has expressed his fear: “These models are very potent and they’re becoming more and more potent. At some point it will be quite easy, if one wanted, to cause a great deal of harm with those models. And as the capabilities get higher it makes sense that you don’t want to disclose them.”<sup>47</sup> An influential white paper by leading researchers says, “It may be prudent to avoid potentially dangerous capabilities of frontier AI models being open-sourced until safe deployment is demonstrably feasible.”<sup>48</sup> All these watchdogs assume that companies will be more responsible than individual actors; however, it is unclear why using the powerful models obtained through a Microsoft cloud contract poses less danger than reusing an open-source model.<sup>49</sup>

Generative models can be used to commit financial fraud, spread election disinformation, create nonconsensual pornography, or share hazardous advice. In some cases, they already are being used for these purposes. Researchers recently manipulated a drug-developing AI to invent 40,000 potentially lethal molecules.<sup>50</sup> Another paper referred to the “proliferation” of AI capabilities,

---

45 Blumenthal and Hawley to Zuckerberg, June 6, 2023.

46 Bruce Schneier and Jim Waldo, “Big Tech Isn’t Prepared for A.I.’s Next Chapter,” *Slate*, May 30, 2023, <https://slate.com/technology/2023/05/ai-regulation-open-source-meta.html>.

47 Quoted in Widder, Whittaker, and West, “Open (for Business).”

48 Markus Anderljung et al., “Frontier AI Regulation: Managing Emerging Risks to Public Safety,” preprint, submitted July 6, 2023, <https://arxiv.org/abs/2307.03718>.

49 Widder, Whittaker, and West, “Open (for Business).”

50 Justine Calma, “AI Suggested 40,000 New Possible Chemical Weapons in Just Six Hours,” *The Verge*, March 17, 2023, <https://www.theverge.com/2022/3/17/22983197/ai-new-possible-chemical-weapons-generative-models-vx>.

adopting the loaded term that often appears in arguments for nuclear arms control.<sup>51</sup>

## Regulating Development

Policymakers who fear the decentralization of powerful MLOSS may seek to control who may create machine learning systems and how they may create them. These policymakers want to regulate the development of AI. Their policies may require licensing, permits, premarket approvals, audits, transparency labels, or agency oversight.<sup>52</sup> These proposals treat AI as inherently dangerous, and they require developers and potential developers to comply with a set of requirements before launching new applications.

Examples of development proposals include the Algorithmic Accountability Act, Senators Hawley and Blumenthal's Framework for U.S. AI Act, and the European Union's Artificial Intelligence Act.<sup>53</sup> The Algorithmic Accountability Act would require businesses to produce impact assessments and attempt to mitigate any "likely material negative impact" that an AI decision will have on a consumer's life. It would also establish a new Bureau of Technology within the Federal Trade Commission to vet the assessments.<sup>54</sup> Senators Hawley and Blumenthal's framework calls for a "licensing

---

51 Markus Anderljung et al., "Frontier AI Regulation."

52 Adam Thierer, "Flexible, Pro-innovation Governance Strategies for Artificial Intelligence," R Street Institute, April 20, 2023, <https://www.rstreet.org/research/flexible-pro-innovation-governance-strategies-for-artificial-intelligence/>.

53 Algorithmic Accountability Act of 2022, H.R. 6580, 117th Cong. (2022); Josh Hawley and Richard Blumenthal, "Bipartisan Framework for U.S. AI Act," accessed December 15, 2023, <https://www.blumenthal.senate.gov/imo/media/doc/09072023bipartisanaiframework.pdf>; *Proposal for a Regulation of the European Parliament and of the Council on Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts*, COM (2021) 206, September 5, 2023, [https://www.europarl.europa.eu/meetdocs/2014\\_2019/plmrep/COMMITTEES/CJ40/DV/2023/05-11/ConsolidatedCA\\_IMCOLIBE\\_AI\\_ACT\\_EN.pdf](https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/CJ40/DV/2023/05-11/ConsolidatedCA_IMCOLIBE_AI_ACT_EN.pdf).

54 Algorithmic Accountability Act of 2022, H.R. 6580.

regime” for companies developing sophisticated general-purpose AI models. It also calls for a new independent oversight body to ensure compliance.<sup>55</sup>

Those proposals are in line with OpenAI CEO Sam Altman’s recommendation that the government adopt licensing and registration requirements for large AI models.<sup>56</sup> Microsoft has contributed its own top-down vision, recommending a “multitiered licensing regime” to enforce advance notification of large training runs, comprehensive risk assessments, and extensive prerelease testing. Microsoft also recommends KY3C compliance, by which it means that developers must know the cloud on which their models are developed and deployed, must know their customers, and must provide labels or watermarks on AI-generated content.<sup>57</sup> Policy analyst Adam Thierer notes that industry-specific regulations are often triggered by firm size, but Microsoft’s and OpenAI’s proposals use “highly capable models” as the threshold for regulation. As a result, start-ups and open-source providers would be covered immediately.<sup>58</sup>

This approach comes with serious costs. The MLOSS community may crack under a top-down regulatory regime. Compliance requirements and the threat of liability for downstream uses could price out open-source projects that operate on already tight budgets.

---

55 Hawley and Blumenthal, “Bipartisan Framework for U.S. AI Act.”

56 Sam Altman, “Written Testimony of Sam Altman, Chief Executive Officer, OpenAI, before the U.S. Senate Committee on the Judiciary, Subcommittee on Privacy, Technology, & the Law,” accessed December 15, 2023, <https://www.judiciary.senate.gov/imo/media/doc/2023-05-16%20-%20Bio%20&%20Testimony%20-%20Altman.pdf>.

57 Microsoft, “Governing AI: A Blueprint for the Future,” May 25, 2023, <https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RW14Gtw>.

58 Adam Thierer, “Microsoft’s New AI Regulatory Framework & the Coming Battle over Computational Control,” Medium, May 29, 2023, <https://medium.com/@AdamThierer/microsofts-new-ai-regulatory-framework-the-coming-battle-over-computational-control-1bcc014272c0>.

Larger players, on the other hand, can more easily shoulder these costs. For example, Microsoft promised to defend its customers from copyright litigation over the use of its generative AI products.<sup>59</sup> It used the same strategy at the beginning of the 21st century to beat out open-source software makers such as Linux when it offered compensation to partners and customers using its software. In 2017, Microsoft again offered to protect customers of its cloud products.<sup>60</sup> Open-source systems, on the other hand, cannot possibly offer similar services because they do not know the users of their programs and cannot afford to cover the legal costs of all their users.

Proposals to regulate development attach liability to MLOSS, which would have a chilling effect on the creation of open-source models. The European Union's proposed Artificial Intelligence Act would impose significant compliance requirements on open-source developers of foundation models, including the obligation to achieve performance, predictability, interpretability, corrigibility, security, and cybersecurity throughout the models' life cycle.<sup>61</sup> Hugging Face CEO Clem Delangue posted,

*"Requiring a license to train models would be like requiring a license to write code. [In my opinion], it would further concentrate power in the hands of a few & drastically slow down progress, fairness & transparency."*<sup>62</sup>

---

59 Brad Smith and Hossein Nowbar, "Microsoft Announces New Copilot Copyright Commitment for Customers," Microsoft, September 7, 2023, <https://blogs.microsoft.com/on-the-issues/2023/09/07/copilot-copyright-commitment-ai-legal-concerns/>.

60 Dina Bass, "Microsoft Says It Will Protect Customers from AI Copyright Suits," *Bloomberg News*, September 7, 2023, <https://news.bloomberglaw.com/artificial-intelligence/microsoft-says-it-will-protect-customers-from-ai-copyright-suits>.

61 European Parliament Briefing, EU Legislation in Progress, PE 698.792, "Artificial Intelligence Act" (June 2023), [https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS\\_BRI\(2021\)698792\\_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI(2021)698792_EN.pdf).

62 Clem Delangue, Twitter, May 17, 2023, 8:56 a.m., <https://twitter.com/ClementDelangue/status/1658818666203344897>.



KYC (know your customer) requirements would hit MLOSS particularly hard. Fortune writer Jeremy Kahn explains, “By their very nature, those offering open-source AI software are unlikely to be able to meet Microsoft’s KYC regime, because open-source models can be downloaded by anyone and used for almost any purpose.”<sup>63</sup> Only well-funded open-source projects could fulfill these obligations. As a result, there will be fewer open-source projects, which hurts innovation because smaller businesses rely on these free building blocks.

Companies, including Microsoft and OpenAI, support regulation that raises their rivals’ costs.<sup>64</sup> Senator Dick Durbin called OpenAI’s testimony before Congress “‘historic,’ because he could not recall having executives come before lawmakers and ‘plead’ with them to regulate their products.”<sup>65</sup> Regulation weighs heavier on smaller firms that lack established compliance teams and budgets. OpenAI embraces regulation now, after it has already built its LLM free from the regulatory oversight it now demands. Economist Lynne Kiesling writes, “Firms have incentives to argue for regulation in their industry, particularly if they have achieved substantial market share or are an early or first mover.”<sup>66</sup> The result of such regulation would be an AI ecosystem run by a small number

---

63 Jeremy Kahn, “Microsoft: Advanced A.I. Models Need Government Regulation, with Rules Similar to Anti-fraud and Terrorism Safeguards at Banks,” *Fortune*, May 25, 2023, <https://fortune.com/2023/05/25/microsoft-president-says-the-u-s-must-create-an-a-i-regulatory-agency-with-rules-for-companies-using-advanced-a-i-models-similar-to-anti-fraud-safeguards-at-banks/>.

64 Naoufel Mzoughi and Gilles Grolleau, “Raising Rivals’ Costs,” in *Encyclopedia of Law and Economics*, ed. Alain Marciano and Giovanni Battista Ramello (New York: Springer, 2021), [https://doi.org/10.1007/978-1-4614-7883-6\\_403-2](https://doi.org/10.1007/978-1-4614-7883-6_403-2).

65 Sue Halpern, “Congress Really Wants to Regulate A.I., but No One Seems to Know How,” *New Yorker*, May 20, 2023, <https://www.newyorker.com/news/daily-comment/congress-really-wants-to-regulate-ai-but-no-one-seems-to-know-how>.

66 Lynne Kiesling, “ShackledAI,” *Knowledge Problem*, May 16, 2023, <https://knowledgeproblem.substack.com/p/shackledai>.

of companies with the massive resources to develop foundational models and comply with burdensome regulations.

If only a few companies generate progress in a field, not only will innovation slow because fewer groups are working on it, but big technology corporations may bury transformative products that would cannibalize their own offerings. Tim Wu refers to this as the Kronos effect.<sup>67</sup> For example, in the twentieth century RCA, the leading electronics firm, helped develop FM radio, but it sat on the technology for years despite innovative use cases. RCA feared that the technology would upend its dominance in AM radio. In the meantime, consumers lost out on high-quality broadcasts, more radio stations, and a greater diversity of programming.<sup>68</sup> Remember, Google had its own LLM that it did not release until after ChatGPT went public. Technology columnist Parmy Olson explains, “The innovator’s dilemma has forced Google to keep LaMDA hidden away, fearful it could cannibalize its own search results.”<sup>69</sup>

Top-down regulations will also stifle the security and safety research that comes with an open-source environment. Security researchers Sayash Kapoor and Arvind Narayanan highlight these risks:<sup>70</sup>

---

67 Tim Wu, *The Master Switch: The Rise and Fall of Information Empires* (New York: Alfred A. Knopf, 2010).

68 Wu, *Master Switch*.

69 Parmy Olson, “Google Will Join the AI Wars, Pitting LaMDA against ChatGPT,” *Washington Post*, February 5, 2023, [https://www.washingtonpost.com/business/google-will-join-the-ai-wars-pitting-lamda-against-chatgpt/2023/02/03/439c2716-a3cd-11ed-8b47-9863fda8e494\\_story.html](https://www.washingtonpost.com/business/google-will-join-the-ai-wars-pitting-lamda-against-chatgpt/2023/02/03/439c2716-a3cd-11ed-8b47-9863fda8e494_story.html).

70 Sayash Kapoor and Arvind Narayanan, “Licensing Is Neither Feasible nor Effective for Addressing AI Risks,” *AI Snake Oil*, June 10, 2023, <https://www.aisnakeoil.com/p/licensing-is-neither-feasible-nor>.

- Monoculture may multiply security risks. If one model powers all commercially available apps, they will be subject to the same vulnerabilities.
- Monoculture may lead to outcome homogenization. If one model powers all resume screening apps, then a candidate could be receive a blanket rejection across many companies.
- Monoculture may confine the boundaries of acceptable speech. If most people use generative models created by a handful of developers, these developers can define acceptable speech and influence people's views.

Regulation of development also suffers from definitional issues. There is no straightforward, consensus definition of artificial intelligence. The term frequently describes large, resource-intensive machine learning systems. But how large and how resource-intensive do these systems need to be? Current discourse and threat scenarios mostly concern image, text, and video generation models, and one gets the sense that these are what regulators are thinking of when they say "AI." Even so, it is unclear at what threshold generative AI becomes "sophisticated" or "highly capable" enough to fall under the proposed regulations.

Even if these regulations are passed, implementation will present challenges. Dozens of MLOSS projects have already shared their code bases and training methodologies. This information will continue to circulate. Meanwhile, China and other countries will not stop their massive investments in AI technology.<sup>71</sup> Absent an international agreement, technologies developed abroad will

---

71 Adam Thierer, "Existential Risks and Global Governance Issues around AI and Robotics" (Policy Study No. 291, R Street Institute, June 2023), [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4174399](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4174399).

inevitably spread to the US. In the 1990s the US tried to control the spread of encryption technology; it failed for these reasons.<sup>72</sup>

Proponents of developmental regulations highlight several benefits to this approach. First, they may argue that top-down legislation is consistent with the tort principle of the *cheapest cost avoider*: liability should be placed on the party who can fix the problem while incurring the least cost.<sup>73</sup> According to this principle,

*"The chosen loss bearer must have better knowledge of the risks involved and of ways of avoiding them than alternate bearers; he must be in a better position to use that knowledge efficiently to choose the cheaper alternative; and finally he must be better placed to induce modifications in the behavior of others where such modification is the cheapest way to reduce the sum of accident and safety costs."<sup>74</sup>*

The companies developing the models have the greatest knowledge of their systems, they can implement changes to their systems that result in better outcomes, and their conduct can have a great impact on the experience of downstream users. However, this doctrine involves difficult empirical questions and is complicated by the nature of AI development.

Take deepfakes for example. Deepfakes have been used by criminals pretending to be a person to trick that person's bank or family into sending money.<sup>75</sup> In the case of defrauded family members,

---

72 Brown, "Expert Explainer."

73 Paul Rosenzweig, "Cybersecurity and the Least Cost Avoider," Lawfare, November 5, 2013, <https://www.lawfaremedia.org/article/cybersecurity-and-least-cost-avoider>.

74 Guido Calabresi, "Concerning Cause and the Law of Torts: An Essay for Harry Kalven, Jr.," *University of Chicago Law Review* 43, no. 1 (1975): 69-108.

75 Heather Chen & Kathleen Magramo, "Finance worker pays out \$25 million after video call with deepfake 'chief financial officer,'" CNN, February 4, 2024, <https://www.cnn.com/2024/02/04/asia/deepfake-cfo-scam-hong-kong-intl-hnk/index.html>.

recovery is limited by the difficulty that victims will have finding the bad actors.<sup>76</sup> The developer of an enterprise deepfake generation platform is well positioned to implement KYC protocols and automatically review its generated deepfakes for potential scams. A deepfake enterprise platform has better knowledge of the methods for avoiding risk than the average person. The developer can implement the previously mentioned protocols at a lower cost than an individual. In the case of bank fraud, however, banks are more aware of the risk and methods for avoiding risk than the enterprise deepfake generation platforms. Fraud is endemic to banks. As a result, banks already deploy verification methods, such as voice authentication.<sup>77</sup> It is likely cheaper for banks to implement additional verification methods, including deepfake detection, at a lower cost than enterprise generation platforms. Banks only need to verify that a customer is who they say they are, whereas enterprise deepfake platforms must ensure that a user is not creating content that will scam an indeterminate party. The bank has decidedly narrower remit. These scenarios also ignore the fact that enterprise deepfake generation platforms are not always the developers of the models they deploy. The models may be open source. Open-source developers cannot easily implement these safeguards, as discussed above.

Proponents of developmental regulation also argue that transparency labels would help users deploy AI more effectively and safely—this disclosure of information is one of the appeals of MLOSS in the first place. Transparency labels explain how an AI model works. Jeremy Howard compares them to nutrition labels: “Whilst

---

76 Jack Langa, “Deepfakes, Real Consequences,” *Boston University Law Review*, 761 no. 101 (2021): 795.

77 Joseph Cox, How I Broke Into a Bank Account With an AI-Generated Voice, *VICE*, February 23, 2023, <https://www.vice.com/en/article/dy7axa/how-i-broke-into-a-bank-account-with-an-ai-generated-voice>.

we don't ban people from eating too much junk food, we endeavor to give them the information they need to make good choices."<sup>78</sup>

Regulating AI's development also solves a legibility issue for the government. It may be easier to identify and monitor resource-intensive developers via computing power than to monitor every downstream harm. Computing power is quantifiable and therefore easier to monitor. The US government already has experience implementing compute governance as a result of its export controls, so it has the capability to restrict usage at the compute level.<sup>79</sup>

## Regulating Deployment

Regulating AI's deployment involves controlling risky and harmful applications by users. This regulatory strategy shifts the liability burden from the developer to the user. Regulating deployment is less speculative than regulating development because it addresses harms that have actually occurred and holds accountable the actors most directly responsible for them. The barriers to entry in building MLOSS are high. Regulating deployment would allow open-source developers to continue operating without fear of being priced out of the industry by compliance requirements and downstream liability.

One benefit of this approach is that it does not require new laws to address many AI harms. The Food and Drug Administration reviews medical devices that involve machine learning before they hit the market, and the National Highway Traffic Safety Administration

---

<sup>78</sup> Howard, "AI Safety and the Age of Dislightenment."

<sup>79</sup> Lennart Heim, "Video and Transcript of Presentation on Introduction to Compute Governance," *Lennart Heim*, May 17, 2023, <https://blog.heim.xyz/presentation-on-introduction-to-compute-governance/>.

produces guidelines for driverless cars. Both agencies possess recall authority. The Federal Trade Commission monitors “unfair and deceptive practices,” including those wrought by AI.<sup>80</sup> The Equal Employment Opportunity Commission, the Federal Trade Commission, the Consumer Financial Protection Bureau, and the Department of Justice have pledged to address algorithmic discrimination.<sup>81</sup> The White House Office of Science and Technology Policy is organizing a National AI Initiative to oversee interagency coordination on algorithmic policy. State laws offer additional consumer protection from algorithmic systems,<sup>82</sup> and state common-law torts are flexible and have evolved to meet challenges posed by new technologies.<sup>83</sup>

New laws that have failed to account for existing laws’ coverage will result in a competing mess of regulations that unnecessarily hamper MLOSS. For harms not covered by existing regulation, lawmakers can pass targeted legislation. For example, New York passed a bill making it illegal to disseminate nonconsensual deepfake pornography.<sup>84</sup> The legislation addressed the direct harm without upending open-source developers.

Regulating AI deployment makes sense given that AI systems can be black boxes. In some cases, neither developers nor deployers know exactly what a model will do in any given scenario. Each prompt is,

---

80 Federal Trade Commission, “A Brief Overview of the Federal Trade Commission’s Investigative, Law Enforcement, and Rulemaking Authority,” revised May 2021, <https://www.ftc.gov/about-ftc/mission/enforcement-authority>.

81 Adam Thierer, “The Most Important Principle for AI Regulation,” R Street Institute, June 21, 2023, <https://www.rstreet.org/commentary/the-most-important-principle-for-ai-regulation/>.

82 Kyle Wiggers, “NYC’s Anti-bias Law for Hiring Algorithms Goes into Effect,” *TechCrunch*, July 5, 2023, <https://techcrunch.com/2023/07/05/nycs-anti-bias-law-for-hiring-algorithms-goes-into-effect/>.

83 Thierer, “Flexible, Pro-innovation Governance Strategies.”

84 S. 1042A, 2023–2024 Leg., Reg. Sess. (N.Y. 2023).

in effect, a roll of the dice. There is always some level of risk associated with bad, unpredictable outcomes. Under a deployment liability regime, the entity that implements a model into its product bears the risks. For example, a nonprofit would be free to build an open-source computer vision model that identifies moving objects. And a self-driving car company that chooses to implement that computer vision model into its cars would assume the risk (liability for a car accident) because it is responsible for the stakes (putting the model into a driverless car).

The US has achieved success by focusing on the downstream liability of the parties most responsible for violative behavior rather than on the tools they used. Section 230 of the Communications Decency Act says that “interactive computer services” like social media platforms are not legally responsible for content posted by “speakers” like social media users.<sup>85</sup> The underlying point in deployment regulation and section 230 cases is to assign liability to the user rather than to the intermediary tools. This is why section 230 was essential in enabling the rise of the modern internet.<sup>86</sup> In the same manner, AI regulation should aim to place liability on the deployers (like the “speakers” in 230) rather than on the model developers (the “interactive computer services” in 230).

The deployment approach is iterative. Harms are identified and mitigated on a case-by-case basis. This allows courts and policymakers to test what works on a smaller scale and refine their approach. Incrementalism, rather than sweeping legislation, leaves room for technology to evolve. Figure 3 shows the exponential

---

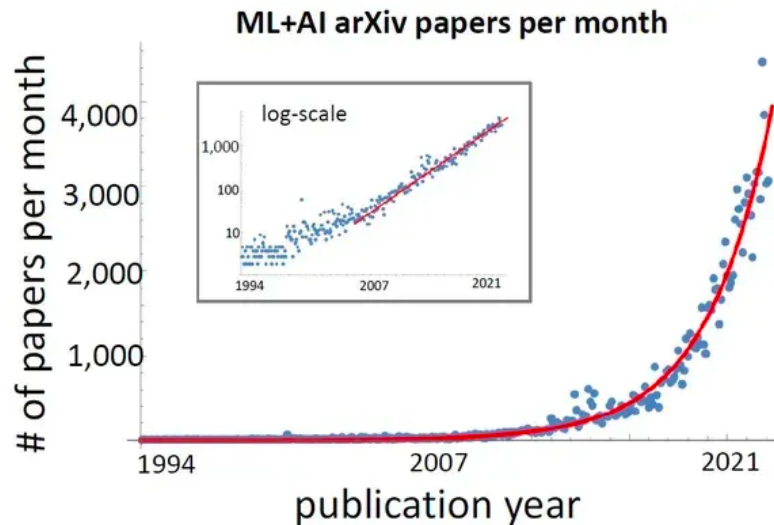
85 47 U.S.C. § 230.

86 Jeff Kosseff, *The Twenty-Six Words That Created the Internet* (Ithaca, NY: Cornell University Press, 2019).



growth of AI research. A breakthrough could upend rigid top-down policies, while incrementalism provides flexibility.

FIGURE 3 | Machine Learning and Artificial Intelligence Papers per Month



Source: Mario Krenn et al., “Predicting the Future of AI with AI: High-Quality Link Prediction in an Exponentially Growing Knowledge Network,” preprint, submitted September 23, 2022, <https://arxiv.org/abs/2210.00881>.

Some contend that developmental controls should be implemented *because* of the rapid pace of progress. These arguments are unsurprising: a period of panic often follows the introduction of new, transformative technology. AI, particularly generative AI, may be on a trajectory known as the “tech panic cycle.”<sup>87</sup> Fears increase rapidly, then they slowly decline as the public becomes familiar with the new technology. However, businesses and policymakers may use public disorientation to push through regulation that helps corporations or increases the power of regulators. Other technologies, such as the printing press, phonograph, and video cameras, followed this course:

87 Patrick Grady and Daniel Castro, “Tech Panics, Generative AI, and the Need for Regulatory Caution,” Center for Data Innovation, May 1, 2023, <https://datainnovation.org/2023/05/tech-panics-generative-ai-and-regulatory-caution/>.

*"In each panic, an innovation in the creative sector makes it much easier to produce new content. Some people, especially incumbents and elites, tend to fear the implications of this new content and concern reaches a boiling point as policymakers and alarmists work together to slow its progress, with news media unable to resist the drama. Eventually, however, the public embraces the tools and moves on."<sup>88</sup>*

As policy analysts Patrick Grady and Daniel Castro point out, this phenomenon is readily observable in the national media:

- *THE ATLANTIC*: "The College Essay Is Dead" (by Stephen Marche, December 6, 2022)
- *DAILY STAR*: "Attack of the Psycho Chatbot" (by Meg Jorsh, February 18, 2023)
- *THE NEW YORK TIMES*: "How ChatGPT Hijacks Democracy" (by Nathan E. Sanders and Bruce Schneier, January 15, 2023)
- *TIME*: "New AI-Powered Bing Is Threatening Users. And That's No Laughing Matter" (by Billy Perrigo, February 17, 2023)
- *NEW YORK POST*: "Rogue AI 'Could Kill Everyone,' Scientists Warn as ChatGPT Craze Runs Rampant" (by Ben Cost, January 26, 2023)

The companies behind these technologies can make regulation work for them. For example, Microsoft and OpenAI paint AI as too powerful to be open sourced to the public but not too capable to be sold as Microsoft cloud API contracts.<sup>89</sup> Regulating harmful applications, rather than regulating the process of building AI,

---

<sup>88</sup> Grady and Castro, "Tech Panics."

<sup>89</sup> Widder, Whittaker, and West, "Open (for Business)."

decreases our chances of falling for the tech panic trap. This, in turn, will protect MLOSS from being regulated out of existence.

## Conclusion

Focusing on regulating AI's deployment rather than its development is a more effective solution to managing AI's externalities with fewer downsides. The rapid advancement of machine learning and the rise of MLOSS has profound implications for the US. A broad commitment to open-source accelerates this innovation by democratizing machine learning and scaling its impact across the entire economy. Many fear the power of MLOSS and seek to limit its reach using regulation.

Regulatory proposals vary significantly, from proposals to regulate AI's development to proposals to regulate its applications. The heavy controls that come with the regulation of development will hurt MLOSS by adding regulatory and liability regimes that make it impossible for open-source systems to exist. This would have a deleterious effect on MLOSS and wind back at least \$30 billion in economic value.<sup>90</sup> Application-based regulations, on the other hand, would hold responsible those most directly linked to harm while allowing MLOSS to flourish and spread the benefits of AI across the economy.

---

90 Langenkamp and Yue, "How Open Source Machine Learning Software Shapes AI."