Regulating AI: Six Principles and Their Consequences

By Jason Furman[1]

Basic research and initial deployments of artificial intelligence have proceeded very rapidly. Workers are rapidly incorporating it into their work (Deming 2024). So far, however, productivity growth has not increased and estimates of its impact going forward range from 0.1 percentage point per year to as much as 1 percentage point per year—with the actual number also likely to depend on choices by policymakers (Acemoglu 2024; Filippucci et al. 2024). Applications also have enormous potential in education, health, climate change, and other areas (US Department of Education 2023; World Economic Forum 2024).

The United States is well ahead of Europe in AI development for many reasons including more flexible and deeper U.S. capital markets and the nature of immigration to the United States (Arnold et al. 2024). An additional factor is the European Union's more extensive AI rules have slowed progress there (Jin, Wagman, & Jia 2019), in contrast to the more voluntary and disclosure-based approach pursued to date in the United States (White House 2023).

The following regulatory principles would have the consequence of increasing the pace of AI development while also maintaining safety[2]:

1. Balance benefits and risks. The European approach to regulation has been based on the "precautionary principle" that supports delaying AI until it is proven absolutely safe (Guha et al. 2023). An alternative regulatory approach is cost-benefit analysis which would require regulators to think not only about the risks of AI but also the risks from slower AI development, such as more cancer deaths because of delayed drug discovery, worse educational outcomes because students lack personalized digital tutors, more car accidents because of delays in self-driving cars, and worsening climate change because of a slowdown in discovering better materials for grid-level battery storage. AI regulation would generate more net benefits, almost by definition, under a benefit-cost framework than the precautionary principle.

2. Compare AI with humans, not to the Almighty. Autonomous cars crash—but how do they compare with human drivers? AI may show biases, but how do these stack up against human prejudices (Kleinberg et al. 2017)? In fact, maximizing social welfare should not require AI to perform at parity with humans given the dynamic consideration that many AI technologies, like self driving cars, offers significant convenience and has greater potential for improvement over time than humans have. Recent research suggests that autonomous driving systems have a significantly lower crash rate than human drivers (Kusano et al. 2024). AI is learning much faster

---

[1] Aetna Professor of the Practice of Economic Policy jointly at the Harvard Kennedy School and the Economics Department at Harvard University. Also a Non-Resident Senior Fellow at the Peterson Institute for International Economics. An earlier version of this essay was published in the *Wall Street Journal*.

[2] See Coglianese (2023) for a broad overview of AI regulation with some overlap, but also differences, from the below.

than humans are and the future gains this learning will generate belong on the benefit side of the ledger.

3. Address how existing regulations are hindering progress. The most obvious are permitting and other obstacles to the expansion of data centers and the power sources they will need; these immediate hurdles delay critical expansions needed to support the growing computational demands of AI (McKinsey & Company 2024). Data centers, which are the backbone of AI infrastructure, face numerous regulatory challenges, often stemming from complex zoning laws, stringent environmental regulations, and local resistance to large-scale infrastructure projects. A potentially more consequential issue over time is the dozens of state laws regulating AI that have already been passed and the hundreds more that have been proposed, creating an unpredictable and fragmented regulatory environment for AI developers (National Conference of State Legislatures 2024). Federal pre-emption of state rules would have the consequence of ensuring the U.S. remains a digital single market—unlike the fractured EU.

4. Where new regulation is warranted, it would be overseen by existing domain-specific regulators rather than a new super-regulator. We don't have a super-regulator for all products that use linear regression or electricity; instead, regulators specialize in areas where these are used, such as auto safety, stock trading, and medical devices. Existing regulators, such as the Food and Drug Administration (FDA) and the Federal Aviation Administration (FAA), can adapt their processes to accommodate AI technologies and should focus on outputs and consequences in their domains, not on inputs and methods—an approach that would minimize regulatory duplication and ensure specialized oversight where it is needed. This may require more AI expertise and flexibility within agencies. For instance, the FDA has come up with procedures to approve AI-based devices that might fall foul of its standard rules given the lack of a fixed algorithm with such devices (FDA 2024).

5. Avoid setting up regulations that become a moat protecting incumbents. History shows that well-intentioned rules can entrench existing powers, from medieval guilds to hospital certificate-of-need laws (Mitchell 2021). It is possible that regulations in the AI space could follow the same pattern. Centralized licensing bodies could easily become gatekeepers stifling competition. A super-regulator could be captured by big companies. When tech giants enthusiastically promote regulation, it should raise red flags. A regulatory framework that aims to nurture a competitive AI landscape instead of solidifying the dominance of a few early movers should avoid these types of policies.

6. Not every problem caused by AI can be solved by regulating AI. I hope AI will raise wages without hurting employment, with especially large increases for workers with lower-paying skills as some expect (Autor 2023). Studies provide some evidence for this, for example finding that less-able writers benefit most from AI-based writing suggestions (Dhillon et al. 2024). But bleak scenarios of swift technological change displacing workers or causing inequality are possible (Acemoglu 2019). It is virtually impossible for regulators to figure out in advance whether each technological advance is job-replacing or inequality-increasing. Moreover, attempting to shift to

a system of permissioned innovation would slow progress and the benefits that AI could bring. Instead, a more effective approach to address any side effects for labor markets would be in more conventional economic policies like training programs that connect people to jobs, wage subsidies, and a more progressive tax and transfer system to ensure that AI's benefits are shared broadly.

To the degree policymakers want to maximize the net benefits of AI they should proceed cautiously. Well-intentioned efforts could inadvertently slow progress while falling short of their goals. These six principles would result in an outcome that would help ensure AI lives up to its potential while addressing legitimate concerns around safety and other issues.

References:

Acemoglu, D., & Restrepo, P. (2019). Automation and new tasks: How technology displaces and reinstates labor. *Journal of economic perspectives*, *33*(2), 3-30. https://pubs.aeaweb.org/doi/pdfplus/10.1257/jep.33.2.3

Acemoglu, D. (2024). The simple macroeconomics of AI. Massachusetts Institute of Technology. Paper prepared for *Economic Policy*. https://economics.mit.edu/sites/default/files/2024-04/The%20Simple%20Macroeconomics%20of%20AI.pdf

Arnold, M., Fleming, S., & Jones, C. (2024, May 13). Can Europe's economy ever hope to rival the US again? *Financial Times*. https://www.ft.com/content/93f88255-787b-4c06-849c-f7722c83e8b6

Autor, D. (2023, April 19). How AI could help rebuild the middle class. *Noema Magazine*. https://www.noemamag.com/how-ai-could-help-rebuild-the-middle-class/

Blandin, A., Bick, A., & Deming, D. (2024). The Rapid Adoption of Generative AI. *Federal Reserve Bank of St. Louis Working Paper 2024-027.*

Coglianese, C. (2023). *Regulating machine learning: The challenge of heterogeneity* (Public Law and Legal Theory Research Paper No. 23-06). University of Pennsylvania Carey Law School. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4368604

Dhillon, P. S., Molaei, S., Li, J., Golub, M., Zheng, S., & Robert, L. P. (2024). Shaping Human-AI Collaboration: Varied Scaffolding Levels in Co-writing with Language Models. *In Proceedings of the CHI Conference on Human Factors in Computing Systems* (pp. 1-18). https://arxiv.org/abs/2402.11723

Filippucci, F., Gal, P., Jona-Lasinio, C., Leandro, A., & Nicoletti, G. (2024). The impact of artificial intelligence on productivity, distribution and growth: Key mechanisms, initial evidence and policy challenges. *Organisation for Economic Co-operation and Development (OECD).*

https://www.oecd.org/content/dam/oecd/en/publications/reports/2024/04/the-impact-of-artificial-intelligence-on-productivity-distribution-and-growth_d54e2842/8d900037-en.pdf

Guha, N., Lawrence, C. M., Gailmard, L. A., Rodolfa, K. T., Surani, F., Bommasani, R., Raji, I. D., Cuéllar, M.-T., Honigsberg, C., Liang, P., & Ho, D. E. (2023). *The AI regulatory alignment problem*. Stanford Institute for Human-Centered Artificial Intelligence (HAI). https://hai.stanford.edu/sites/default/files/2023-11/AI-Regulatory-Alignment.pdf

Jin, G., Wagman, L., & Jia, J. (2019). The short-run effects of GDPR on technology venture investment. *VoxEU*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3278912

Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J., & Mullainathan, S. (2018). Human decisions and machine predictions. *The Quarterly Journal of Economics, 133*(1), 237–293. Retrieved from https://sendhil.org/wp-content/uploads/2019/08/Publication-5.pdf

Kusano, K. D., Scanlon, J. M., Chen, Y. H., McMurry, T. L., Chen, R., Gode, T., & Victor, T. (2024). Comparison of Waymo rider-only crash data to human benchmarks at 7.1 million miles. *Traffic Injury Prevention*, 1-12. https://arxiv.org/pdf/2312.12675

Manyika, J., Chui, M., Miremadi, M., Bughin, J., George, K., Willmott, P., & Dewhurst, M. (2017). A future that works: AI, automation, employment, and productivity. *McKinsey Global Institute Research, Tech. Rep*, 60, 1-135. https://www.mckinsey.com/~/media/mckinsey/featured%20insights/Digital%20Disruption/Harnessing%20automation%20for%20a%20future%20that%20works/MGI-A-future-that-works-Executive-summary.ashx

McKinsey & Company. (2024, October 29). *AI power: Expanding data center capacity to meet growing demand*. https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/ai-power-expanding-data-center-capacity-to-meet-growing-demand#/

Mitchell, M. D. (2021). *Certificate-of-need laws: How they affect healthcare access, quality, and cost*. Mercatus Center at George Mason University. https://www.mercatus.org/economic-insights/features/certificate-need-laws-how-they-affect-healthcare-access-quality-and-cost

National Conference of State Legislatures. (2024, September 9). *Artificial intelligence 2024 legislation*. https://www.ncsl.org/technology-and-communication/artificial-intelligence-2024-legislation

The White House. (2023). *Voluntary AI commitments*. Retrieved from https://www.whitehouse.gov/wp-content/uploads/2023/09/Voluntary-AI-Commitments-September-2023.pdf

US Department of Education. (2023). *Artificial intelligence and the future of teaching and learning: Insights and recommendations*. https://www.ed.gov/sites/ed/files/documents/ai-report/ai-report.pdf

US Food and Drug Administration. (2024). Marketing submission recommendations for a predetermined change control plan for artificial intelligence. *Machine Learning (AI/ML) Enabled Device Software Functions*. www.fda.gov/media/166704/download

World Economic Forum. (2024). *How AI can combat climate change*. Retrieved from https://www.weforum.org/stories/2024/02/ai-combat-climate-change/