AI Regulation is Coming- What is the Likely Outcome?

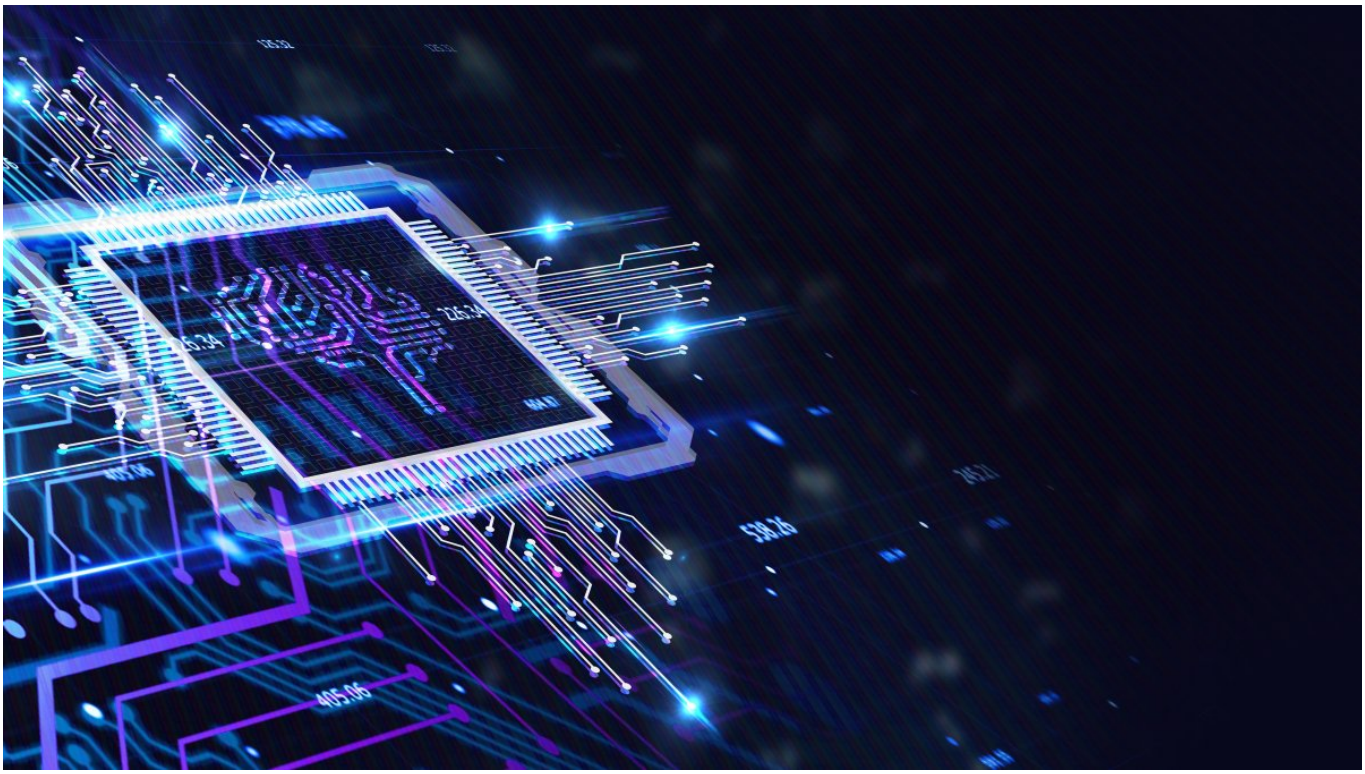

Photo: putilov_denis/Adobe Stock

Blog Post by **Bill Whyman**
Published October 10, 2023

"AI is too important not to regulate—and too important not to regulate well," says Google. It is highly likely that Artificial Intelligence will be regulated. In many ways, it already is. The EU's AI Act has extensive top-down prescriptive rules including prohibiting uses of AI that it says pose unacceptable risk; it's in the final stages of passing. China has ruled that algorithms must be reviewed in advance by the state and "should adhere to the core socialist values". The United States is taking its typical decentralized approach.

The most likely outcome for the United States is a bottom-up patchwork quilt of executive branch actions. Unlike Europe, the United States is not likely to pass a broad national AI law over the next few years. Successful legislation is likely focused on less controversial and targeted measures like funding AI research and AI child safety. This likely disappoints proponents of strong national AI regulation.

This outcome will be messier and have gaps, but no broad national law does not mean no regulation. We're likely to see domain-specific agency actions especially in health care, financial services, housing, workforce, and child safety, plus multiple executive orders. This patchwork quilt of rules, if implemented well, could be grounded in the expertise of specific agencies and more tailored to innovation. The U.S. government will likely boost spending on AI and AI research, especially in defense and intelligence, and use its buying power to shape the market. AI trade friction with Europe is likely to emerge, and private companies will advance their own "responsible AI" initiatives and face a fragmented global AI regulatory landscape. Looming China competition will escalate a "don't fall behind" debate. The Federal Trade Commission (especially) and Department of Justice likely fire warning shots to forestall AI concentration in big tech. There is a real but less likely chance that a key U.S. state (e.g., California) passes major AI legislation, or that a big AI-related disaster leads to a strong national response.

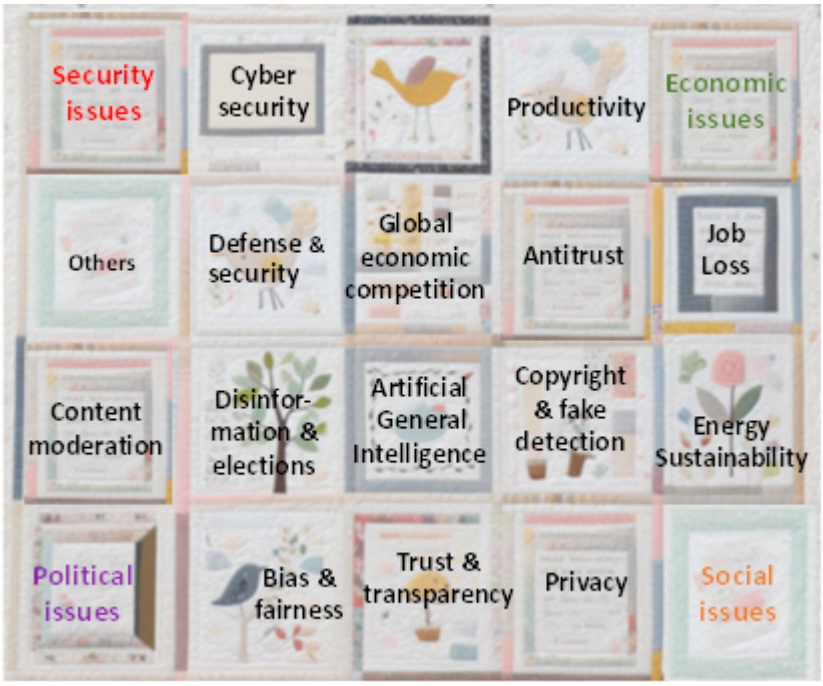

Photo: Bill Whyman

This viewpoint proposes ten key parameters of successful AI regulatory design and the likely outcomes. AI raises many profound questions. Nearly all agree that AI promises both huge benefits for society and yet poses major risks. The challenge is getting the balance right between innovation and societal risks - which governments don't have a good record of achieving. The devil is in the details, making rules adaptable for a technology that is likely to change rapidly and be pervasive. The existential question of Artificial General Intelligence (AGI) is

generating much debate. However, this is more distant and tough to forecast accurately. We focus here on issues like model safety, bias, transparency, privacy, security, trust, copyright, content regulation, education, and economic impacts (job loss, work force adjustment, productivity). For commercial actors, the key question is regulation's impact on innovation, and does it address society's concerns so that adoption is broad and sustainable.

I. The AI Regulation Debate: A Wide Spectrum of Views. The AI regulation debate has become heated and politicized, reflecting the immense hopes and fears that we have invested in AI.

Hysterical fear and paranoia. On one end, leading technologist and venture capitalist Marc Andreessen says "we have a full-blown moral panic about AI right now." He views this as "hysterical fear" that is "irrational" and compares it to earlier introduction of new technologies that were similarly feared but were overblown and subsequently managed. He believes that AI is simply a computer program that is owned, controlled and used by people, and that AI does not have its own goals or its own wants. Moreover, he emphasizes, AI can be a force for tremendous good.

Risk of Extinction. At the same time, AI academic computer scientists like Yoshua Bengio and Geoffrey Hinton, who are the Turing prize-winning "godfathers" of generative AI, warn of dramatic risks that threaten humanity's very existence. Bengio, Hinton, Bill Gates, top executives from Google, Microsoft, OpenAI and many AI luminaries including from China and Russia have signed a statement stating "Mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war." There are also proposals to pause developing more powerful AI models than ChatGPT 4. Elon Musk has said, "With AI we're summoning the demon." The precautionary principle suggests low-probability, high-impact outcomes should be taken seriously.

International context. geopolitical competition, national security, and economic competitiveness. The AI debate is taking place globally. AI has direct implications for national security, military capabilities, and global economic competitiveness. This leads to pressures to stake out parochial positions, even as many of the issues are global in nature. The EU's AI Act will apply to non-EU companies providing AI

services in Europe and will set a precedent (like its privacy rules) that other countries will likely follow. 31 countries have passed AI legislation and 13 more are debating AI laws.

The EU's AI Act intends to be the "world's first comprehensive AI law". Central to the EU's approach, AI systems are classified into four tiers of risk, and different tiers are subject to different regulations. Implementation will be a challenge, for example even defining AI systems and AI risks is problematic. EU businesses have released an open letter stating it "would jeopardize Europe's competitiveness and technological sovereignty without effectively tackling the challenges we are and will be facing". A new EU AI office would be created to monitor enforcement. Penalties include fines of up to 6% of total worldwide revenue. Citizens also have the right to file complaints against the AI provider.

1. AI that falls in the EU's highest risk category of "unacceptable risk" is prohibited, with certain limited exceptions. Examples include social scoring that classifies people based on behavior and socio-economic status, or real-time biometric identification like face recognition.

2. The EU's second category is "high risk" AI, which is permitted but requires assessments before AI is released in the market and afterwards. This includes rigorous testing, documentation of data quality, and an accountability framework including human oversight. High risk services include autonomous vehicles, medical devices, critical infrastructure, education, and government services. Providers of high-risk AI must register their AI in an EU database managed by the Commission before introducing them in the market. Non-EU providers will require an authorized representative in the EU showing that they comply and have post-market monitoring.

3. Third, "limited risk" AI systems have minimal transparency requirements so that users can make informed decisions. This includes generating or manipulating images, audio or video.

4. Fourth, "low/minimal risk" AI systems have no new obligations, but likely voluntary codes of conduct.

Generative AI like Chat-GPT has additional requirements including disclosing that the content was generated by AI, designing the model to prevent it from generating

illegal content, and publishing summaries of copyrighted data used for training.

Venture Capitalists- Andreessen Horowitz: AI will save the world. Influential venture capitalist Marc Andreessen (co-creator of the Mozilla Internet browser) writes, "The development and proliferation of AI – far from a risk that we should fear – is a moral obligation…" He sees the drive to regulate AI as hysterical fear that is irrational. He calls out big companies for selling fear and pushing AI regulation to protect their economic interests. Big companies have the resources to meet extensive AI rules, while small companies and start-ups (that venture capitalists invest in) mostly do not. Andreessen argues:

- The risks are overblown. There are few legitimate concerns or negative externalities from AI.
- Companies (and open source) should be allowed to build AI as fast and aggressively as they can.
- Big AI companies "should not (be) allowed to achieve regulatory capture, not allowed to establish a government-protect cartel that is insulated from market competition…"
- "To offset the risk of bad people doing bad things with AI, governments working in partnership with the private sector should…use AI to maximize society's defensive capabilities."
- "…Prevent the risk of China achieving global AI dominance…" and ensure "We win, they lose."

In between the EU's more interventionist approach and venture capitalists' more hands-off approach, many businesses and civil society groups are advocating for varying degrees of regulation. For example, Microsoft has proposed a comprehensive approach in its, "Governing AI: A Blueprint for the Future". It would create a new government AI agency, a new AI-legal framework, require safety brakes for AI systems that control critical infrastructure, and would license AI datacenters that run critical AI.

AI rules create winners and losers, as some are constrained, and others have more freedom and resources. Governments—leery of being criticized for heavy handed regulation—are quick to speak of innovation but are also responding to public concerns about AI.

II. Top Ten Parameters Driving Regulatory Design and Outcomes

To help policy makers and regulators, we define 10 key parameters that must be addressed in some fashion. This does not propose solutions for each domain or issue.

1. **Transparency, fairness, explainability, security and trust are over-arching goals.** Unless these goals can be achieved, AI regulation is unlikely to succeed, and AI adoption will be stunted. Many specific issues, e.g., informing users of risks, disclosing model characteristics, model-bias, and independent model safety testing, support these broader goals.

2. **Risk-based approach.** AI is a general-purpose technology that is likely to be ubiquitous and apply to a wide range of applications. This makes it harder to regulate well. Many propose a risk-based approach where higher-risk AI applications that can cause greater damage have proportionately greater regulation and lower-risk applications have less. This has merits, but there are many implementation obstacles including who and how risk is defined and measured, what are the minimum requirements to offer AI services, and what AI uses are deemed illegal. Further, Google argues that focusing regulation on the highest risks may deter innovation that also has the highest value to society.

3. **Mitigating risk and addressing malicious actors.** The existence of risk may not justify full regulation if those risks can be mitigated, as in other domains (e.g., mitigating environmental impacts or mitigating financial risk through lending standards and insurance). Rules also need to reflect the difference between legitimate businesses seeking to comply with rules and malicious actors with bad intent. Malicious actions (cybersecurity, cyber-crime, terrorism) pursuing dangerous goals (theft, physical safety, child safety) need stronger rules and more punitive deterrents.

4. **Innovation and pre-approval.** The EU, Microsoft, and OpenAI approaches require government pre-approval or licensing of AI models in high-risk areas. Regulatory pre-approval has historically yielded much less innovation (e.g., the telephone network pre-“Carterphone decision”). This is especially important given how rapidly AI is changing. The Internet has produced rapid innovation in part because there is no pre-approval; anyone can put any lawful application on the Web without permission. All say they want innovation, but pre-approval or

licensing deliberately creates barriers to market entry and hence works against innovation and open competition. Outside of high-risk areas, a lighter-touch pre-release notification system may achieve similar goals.

5. **Data flows, data privacy, data security.** AI crucially depends on large volumes of high-quality data. Model accuracy and outcomes directly reflect the data it was trained on. Hence, data regulation is central to AI regulation. AI likely magnifies data privacy concerns. The United States has no national data privacy law; the EU builds on GDPR. Data security and cross-border data flows are also connected to AI regulation. These are not new issues, and AI regulation should build on prior efforts.

6. **New Laws? Sector specific?** The U.S. government has ~50 independent regulatory bodies, and many AI risks can be addressed via existing authorities. For example, the FTC has authority over "false and deceptive" practices such as AI deepfakes. The Equal Employment Opportunity Commission is addressing potential bias of AI models in hiring processes. Yet, it's not clear how existing legal authority applies to AI in many cases, and in some cases, there are gaps in authority. Microsoft proposes "developing a broad legal and regulatory framework based on the technology architecture of AI." Others oppose creating a new legal regime and think AI is best regulated under existing laws.

7. **Institutional approach.** New government agency? What should private businesses do? To answer this, we must develop a view on how unique the challenges posed by AI are and whether they require an AI-specific institution. Microsoft and others propose creating a new government AI agency, at least to address "highly capable AI foundation models." A new AI agency could bring resources and expertise. However, history shows a new agency can lead to "gaming" of regulation, regulatory "capture" by incumbents, agency "mission creep", and slow, bureaucratic decision making.

- IBM and Google don't propose a new AI agency. An alternative approach is to strengthen existing agencies' competencies to address AI while leveraging their domain expertise. This could be supplemented with a multi-stakeholder approach that has been used for Internet governance. Expert advisory committees are already employed in trade (ISAC) and cybersecurity (SCC).

- A new White House coordinating office (such as the Office of the National Cyber Director) could help bring focus, consistency across agencies, and high-level political leadership. On a technical level, agencies like NIST can also provide a common technical approach to defining and measuring key concepts like data quality, bias, explainability, and auditability.
- Private businesses are better positioned to address certain tasks such as safe model design, testing, publishing model capabilities, internal AI ethics boards, etc. Other market-based approaches include third party auditors to test AI (there are not yet full AI standards to test to) or insurance-based approaches (only mixed success in cybersecurity).

8. **"Carrots" encouraging "good AI" outcomes and "sticks" for enforcement and penalties.** Regulation often works best when it uses carrots (incentives, safe harbors) as well as sticks (penalties). Google and others focus on "good AI" outcomes such as an AI-ready workforce, investing in AI innovation and competitiveness, and supporting broad AI adoption. The EU AI Act takes a strong enforcement approach and has big financial penalties. U.S. AI efforts have been mostly voluntary, such as the White House July 21, 2023 commitments from AI companies.

9. **Scope.** AI supply chain including cloud data centers. For high risk uses, some want regulation to apply not just to the model developer, but to applications and the IT infrastructure it runs on, even including components. For example, AI systems that support critical infrastructure would be deployed only in licensed AI data centers. This potentially expands regulation to new areas.

10. **International harmonization.** Nearly every country is having its own debate about AI's dramatic impact. Different national rules will inevitably conflict, including intellectual property rights and antitrust. The EU AI Act is almost certain to be more regulatory than the U.S. approach, and the U.S.-EU Trade and Technology Council is coordinating their differences. The G7 has launched the "Hiroshima AI process". The OECD has developed AI principles, and the United Nations has proposed a new UN AI advisory body to better include views of developing countries. International technical standards, such as developed via ISO and IEEE, are a promising approach that should be considered. Shared R&D and education assistance can also encourage AI harmonization.

III. Likely Outcome: A Bottom-up Patchwork Quilt of AI Rules.

The United States is likely to take a decentralized bottom-up approach that is messy and may have gaps and inconsistencies. This will likely disappoint proponents of strong AI regulation. Yet, over time this may produce a patchwork quilt of AI rules that if implemented well is grounded in the expertise of specific agencies and more tailored to innovation. The tremendous risks and opportunities of AI have made it a presidential-level issue. The White House is coordinating executive agencies, as each moves ahead with actions in its own domain. The lack of meaningful regulation of social media is widely seen as a failure in Washington, and bolsters efforts to be more proactive with AI. Yet the decentralized structure of the U.S. government, political differences, and the complexities of AI forestall big, quick actions. There is also a real but less-likely probability that a big AI-related failure sparks a strong national government action.

- A broad-based national AI law like the EU Act is unlikely over the next few years. The administration likely builds on its AI "bill of rights" that spans different sectors, voluntary commitments, and executive orders. Voluntary rules are seen by many as a stop-gap measure, but a divided Congress is unlikely to pass a major law with new mandatory rules.
- Actions addressing sensitive areas such as health care, financial services (lending, insurance, housing), work force practices (discrimination), and child safety are more likely. Executive branch agencies are likely to move forward with existing authorities in the absence of new laws.
- Private tech companies advance their own responsible AI initiatives to serve their customers and avoid tougher government action. These are largely voluntary and seen as insufficient.
- Federal spending on AI and AI research expands including for non-profits and universities. Funds come with market-shaping rules, as the US uses its buying power in key areas such as health, education, national security, and public safety. The government responds to problems by spending money. The National Science Foundation is funding 25 AI research institutes in part to boost national economic competitiveness.
- Executive Orders to limit AI bias and risks in federal agency programs are in process. Additional orders also likely focus on enhancing adoption of AI in

federal IT to improve citizen services and strengthen AI security, with uneven success.

- Risk Management, Standards, Audits and Assessments. The Commerce Department's NIST plays a key role in defining AI standards and risk management practices. They can have a wide impact, but they are mostly voluntary. Commerce's NTIA AI inquiry is focused on "build(ing) an ecosystem of AI audits, assessments and other mechanisms," that are likely part of the US response. They are key to enforcement. These efforts can play an important role globally.
- U.S. states and cities likely pass targeted actions. There is a real but lower chance that a big state like California passes a major AI law, substantially altering the U.S. regulatory environment.

Domain Specific Outcomes

- Trade conflict emerges with the EU and others, creating a fragmented global regulatory environment. The United States likely globally advances its AI principles and voluntary business commitments to influence international rules (vs. China and EU approaches). IPR and copyright will be especially important for creative industries. The EU's AI Act will likely influence other major nations, frustrating U.S. leadership. Leading companies like Amazon, Apple, Google, Meta, Microsoft, and Nvidia likely face multiple AI regimes around the world.
- Antitrust agencies lead the effort to forestall "big tech" companies dominating AI, false and deceptive practices, and AI-driven fraud. The high cost and scale of AI foundation models likely leads to market concentration. We expect FTC actions designed as warnings shots to industry.
- The Department of Defense incorporates AI in weapon systems, command and control, and intelligence. Humans will stay "in the loop". DoD is likely criticized for moving too slow. The intelligence community and law enforcement agencies are also likely major adopters of AI, sparking civil rights concerns.
- Growing competition with China shadows the AI regulatory effort, escalating a "don't fall behind China" debate. Controls on exports and investments in AI-related technologies such as advanced GPUs are likely to expand over time and broaden, e.g., cloud computing, quantum.

Tags

Artificial Intelligence and Technology

Center for Strategic and International Studies
1616 Rhode Island Avenue, NW
Washington, DC 20036

Tel: 202.887.0200
Fax: 202.775.3199

MEDIA INQUIRIES

H. Andrew Schwartz

Chief Communications Officer

202.775.3242

aschwartz@csis.org

Samuel Cestari

Media Relations Coordinator, External Relations

202.775.7317

scestari@csis.org

See Media Page for more interview, contact, and citation details.

©2024 Center for Strategic & International Studies. All Rights Reserved.