INSTITUTE
FOR LAW & AI

Legal Considerations for Defining "Frontier Model"

LawAI Working Paper Series, No. 2-2024

Charlie Bullock, Suzanne Van Arsdale, Mackenzie Arnold, Cullen O'Keefe, and Christoph Winter

September 2024

law-ai.org

LawAI Working Paper No. 2-2024
Last update: September 2024
Comments welcome: charlie.bullock@law-ai.org

Legal Considerations for Defining "Frontier Model"

Charlie Bullock,[*] Suzanne Van Arsdale,[†] Mackenzie Arnold,[‡] Cullen O'Keefe,[**] and Christoph Winter[††]

Abstract

Many proposed laws and rules for the regulation of artificial intelligence would distinguish between a category consisting of the most advanced models—often called "frontier models"—and all other AI systems. Legal rules that make this distinction will typically need to include or reference a definition of "frontier model" or whatever analogous term is used. The task of creating this definition implicates several important legal considerations. The role of statutory and regulatory definitions in the overall definitional scheme should be considered, as should the advantages and disadvantages of incorporating elements such as technical inputs, capability metrics, epistemic elements, and deployment context into a definition. Additionally, existing legal obstacles to the rapid updating of regulatory definitions should be taken into account—including recent doctrinal developments in administrative law such as the elimination of *Chevron* deference and the introduction of the major questions doctrine.

Keywords

law & artificial intelligence; legal definitions; frontier models; regulatory updating

[*] Institute for Law and AI, Cambridge, MA, USA. Email: charlie.bullock@law-ai.org
[†] Institute for Law and AI, Cambridge, MA, USA. Email: suzanne.vanarsdale@law-ai.org
[‡] Institute for Law and AI, Cambridge, MA, USA. Email: mackenzie.arnold@law-ai.org
[**] Institute for Law and AI, Cambridge, MA, USA. Email: cullen.okeefe@law-ai.org
[††] Instituto Tecnológico Autónomo de México, Mexico City, Mexico / Harvard University, Cambridge, MA, USA / Institute for Law and AI, Cambridge, MA, USA. Email: christoph_winter@fas.harvard.edu.

I. Introduction

One of the few concrete proposals on which AI governance stakeholders in industry[1] and government[2] have mostly[3] been able to agree is that AI legislation and regulation should recognize a distinct category consisting of the most advanced AI systems. The executive branch of the U.S. federal government refers to these systems, in Executive Order 14110 and related regulations, as "dual-use foundation models."[4] The European Union's AI Act refers to a similar class of models as "general-purpose AI models with systemic risk."[5] And many researchers, as well as leading AI labs and some legislators, use the term "frontier models" or some variation thereon.[6]

These phrases are not synonymous, but they are all attempts to address the same issue—namely that the most advanced AI systems present additional regulatory challenges distinct from those posed by less sophisticated models. Frontier models are expected to be highly capable across a broad variety of tasks and are also expected to

---

[1] *See* Microsoft, Governing AI: A Blueprint for the Future (2023), https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RW14Gtw; Chris Meserole, *Year in Review: Building a Safer Future Together*, Frontier Model Forum (Dec. 21, 2023), https://www.frontiermodelforum.org/updates/year-in-review/; Markus Anderljung et al., *Frontier AI Regulation: Managing Emerging Risks to Public Safety* (2023), https://arxiv.org/abs/2307.03718.

[2] *See* Exec. Order No. 14110, 3 C.F.R. 14110 (2024) § 4.2; Senator Richard Blumenthal & Senator Josh Hawley, Bipartisan Framework for U.S. AI Act (2023), https://www.blumenthal.senate.gov/imo/media/doc/09072023bipartisanaiframework.pdf; EU Artificial Intelligence Act, 2024 O.J. (L 1689) § 52a; Frontier AI Taskforce, First Progress Report (2023), https://www.gov.uk/government/publications/frontier-ai-taskforce-first-progress-report/frontier-ai-taskforce-first-progress-report; Safe and Secure Innovation for Frontier Artificial Intelligence Models Act, SB-1047 (CA 2024), https://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=202320240SB1047.

[3] *But see, e.g.,* Jeremy Howard, AI Safety and the Age of Dislightenment (2023), https://www.fast.ai/posts/2023-11-07-dislightenment.html; Gina Helfrich, *The Harms of Terminology: Why We Should Reject So-Called "Frontier AI,"* 4 AI Ethics 699 (2024); Adam Thierer, Flexible, Pro-Innovation Governance Strategies for Artificial Intelligence (2023), https://www.rstreet.org/research/flexible-pro-innovation-governance-strategies-for-artificial-intelligence/.

[4] 3 C.F.R. 14110 §3(k); Taking Additional Steps To Address the National Emergency With Respect to Significant Malicious Cyber-Enabled Activities, 89 Fed. Reg. 5698 (proposed Jan. 29, 2024).

[5] EU AI Act, 2024 O.J. (L 1689) art. 51.

[6] *See* Markus Anderljung & Anton Korinek, *Frontier AI Regulation: Safeguards Amid Rapid Progress*, Lawfare (Jan. 4, 2024), https://www.lawfaremedia.org/article/frontier-ai-regulation-safeguards-amid-rapid-progress; Paul Scharre, Future-Proofing Frontier AI Regulation (2024), https://www.cnas.org/publications/reports/future-proofing-frontier-ai-regulation; Helen Toner & Timothy Fist, Regulating the AI Frontier: Design Choices and Constraints (2023) https://cset.georgetown.edu/article/regulating-the-ai-frontier-design-choices-and-constraints/; AI Safety Summit, The Bletchley Declaration by Countries Attending the AI Safety Summit, 1-2 November 2023 (2023), https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023; Frontier Model Forum: Advancing Frontier AI Safety, Frontier Model Forum, https://www.frontiermodelforum.org/; Anderljung et al. *supra* note 1; Senator Romney et al., Framework for Mitigating Extreme AI Risks (2024), https://www.romney.senate.gov/wp-content/uploads/2024/04/AI-Framework_2pager.pdf; Cal. SB-1047.

have applications and capabilities that are not readily predictable prior to development, nor even immediately known or knowable after development.[7] It is likely that not all of these applications will be socially desirable; some may even create significant risks for users or for the general public.

The question of precisely how frontier models should be regulated is contentious and beyond the scope of this paper. But any law or regulation that distinguishes between "frontier models" (or "dual-use foundation models," or "general-purpose AI models with systemic risk") and other AI systems will first need to define the chosen term. A legal rule that applies to a certain category of product cannot be effectively enforced or complied with unless there is some way to determine whether a given product falls within the regulated category. Laws that fail to carefully define ambiguous technical terms often fail in their intended purposes, sometimes with disastrous results.[8] Because the precise meaning of the phrase "frontier model" is not self-evident,[9] the scope of a law or regulation that targeted frontier models without defining that term would be unacceptably uncertain. This uncertainty would impose unnecessary costs on regulated companies (who might overcomply out of an excess of caution or unintentionally undercomply and be punished for it) and on the public (from, e.g., decreased compliance, increased enforcement costs, less risk protection, and more litigation over the scope of the rule).

The task of defining "frontier model" implicates both legal and policy considerations. This paper provides a brief overview of some of the most relevant legal considerations for the benefit of researchers, policymakers, and anyone else with an interest in the topic.

II. Statutory and Regulatory Definitions

Two related types of legal definition—statutory and regulatory—are relevant to the task of defining "frontier model." A statutory definition is a definition that appears in a statute enacted by a legislative body such as the U.S. Congress or one of the 50 state legislatures. A regulatory definition, on the other hand, appears in a regulation promulgated by a government agency such as the U.S. Department of Commerce or the California Department of Technology (or, less commonly, in an executive order).

---

[7] *See* Anderljung et al., *supra* note 1 at 9–13, 34–36; Tom Davidson et al., *AI Capabilities Can Be Significantly Improved without Expensive Retraining* (2023), https://arxiv.org/abs/2312.07413.

[8] *See* Christoph Winter & Charlie Bullock, The Governance Misspecification Problem (on file with authors) (discussing historical examples of this phenomenon in U.S. regulation); David Nimmer, *Codifying Copyright Comprehensibly*, 51 UCLA L. Rev. 1235, 1331–34 (2004) (discussing the Audio Home Recording Act of 1992, which was rendered obsolete a few years after its enactment because its definition of "digital musical recording" did not include digital files stored on computer hard drives); Matthew Chung et al., *"Title Zero": Ending the Infinite Loop of Classifications for Broadband via a Technology-Agnostic Definition*, 110 Cal. L. Rev 1375, 1386–95 (2022) (discussing the negative consequences of the lack of a statutory definition of the term "broadband" in the Telecommunications Act of 1996).

[9] *See* Anderljung et al. *supra* note 1 at 34–37.

Regulatory definitions have both advantages and disadvantages relative to statutory definitions. Legislation is generally a more difficult and resource-intensive process than agency rulemaking, with additional veto points and failure modes.[10] Agencies are therefore capable of putting into effect more numerous and detailed legal rules than Congress can,[11] and can update those rules more quickly and easily than Congress can amend laws.[12] Additionally, executive agencies are often more capable of acquiring deep subject-matter expertise in highly specific fields than are congressional offices due to Congress's varied responsibilities and resource constraints.[13] This means that regulatory definitions can benefit from agency subject-matter expertise to a greater extent than can statutory definitions, and can also be updated far more easily and often.

The immense procedural and political costs associated with enacting a statute do, however, purchase a greater degree of democratic legitimacy and legal resiliency than a comparable regulation would enjoy. A number of legal challenges that might persuade a court to invalidate a regulatory definition would not be available for the purpose of challenging a statute.[14] And since the rulemaking power exercised by regulatory agencies is generally delegated to them by Congress, most regulations must be authorized by an existing statute. A regulatory definition generally cannot eliminate or override a statutory definition[15] but can clarify or interpret.

Often, a regulatory regime will include both a statutory definition and a more detailed regulatory definition for the same term.[16] This can allow Congress to choose the

---

[10] *See* Michael J Gerhardt, *Why Gridlock Matters*, 88 NOTRE DAME LAW REVIEW 2107 (2013) (discussing structural features of the U.S. Constitution that make federal lawmaking "difficult but not impossible" by design).

[11] Federal administrative agencies generally publish between 3500 and 4000 final rules in a given year. *See* MAEVE P CAREY, CONG. RSCH. SERV., R43056, COUNTING REGULATIONS: AN OVERVIEW OF RULEMAKING, TYPES OF FEDERAL REGULATIONS, AND PAGES IN THE FEDERAL REGISTER (2019). The U.S. Congress, on the other hand, passed only 27 bills that became law in the year 2023. Andrew Solender, *Capitol Hill Stunner: 2023 Led to Fewest Laws in Decades*, AXIOS (2023), https://www.axios.com/2023/12/19/118-congress-bills-least-unproductive-chart.

[12] *See* Brad A Greenberg, *Rethinking Technology Neutrality*, 1495 MINN. L. REV. 1557 (2016) ("[A]n agency reduces legal uncertainty and delay both by streamlining the process of updating the law and unifying rulemaking in a central body. Congress, by empowering an agency, would avoid the pitfalls of locking the law in place for decades without deliberately tailoring it to technology that did not exist at the time the law was passed.").

[13] *See* Stephen Breyer, *The Executive Branch, Administrative Action, and Comparative Expertise*, 32 CARDOZO L. REV. 2189 (2011); Kathy Goldschmidt, *Congress Lacks the Capacity to Meet the Demands of a 21st Century Constituency*, CONGRESSIONAL MANAGEMENT FOUNDATION (2022), https://www.congressfoundation.org/news/blog/1921-congress-lacks-the-capacity-to-meet-the-demands-of-a-21st-century-constituency.

[14] For example, regulations can be challenged under the major questions doctrine, described in Section V below, while statutes cannot. Additionally, a court may find that a regulation is invalid if the agency that promulgated it did so in a manner inconsistent with the Administrative Procedure Act.

[15] *See, e.g.,* United States v. Maes, 546 F.3d 1068 (9th Cir. 2008) ("[A] regulation does not trump an otherwise applicable statute unless the regulation's enabling statute so provides.").

[16] For instance, the Family and Medical Leave Act ("FMLA") contains a brief statutory definition of the term "serious health condition" at 29 U.S.C. § 2611(11), while the U.S. Department of Labor

best of both worlds, establishing a threshold definition with the legitimacy and clarity of an act of Congress while empowering an agency to issue and subsequently update a more specific and technically informed regulatory definition.

III. Existing Definitions

This section discusses five noteworthy attempts to define phrases analogous to "frontier model" from three different existing measures. Executive Order 14110 ("EO 14110"), which President Biden issued in October 2023, includes two complementary definitions of the term "dual-use foundation model." Two definitions of "covered model" from different versions of the Safe and Secure Innovation for Frontier Artificial Intelligence Models Act, a California bill that was recently vetoed by Governor Newsom, are also discussed, along with the EU AI Act's definition of "general-purpose AI model with systemic risk."

A. Executive Order 14110

EO 14110 defines "dual-use foundation model" as:

> an AI model that is trained on broad data; generally uses self-supervision; contains at least tens of billions of parameters; is applicable across a wide range of contexts; and that exhibits, or could be easily modified to exhibit, high levels of performance at tasks that pose a serious risk to security, national economic security, national public health or safety, or any combination of those matters, such as by:
>
> (i) substantially lowering the barrier of entry for non-experts to design, synthesize, acquire, or use chemical, biological, radiological, or nuclear (CBRN) weapons;
>
> (ii) enabling powerful offensive cyber operations through automated vulnerability discovery and exploitation against a wide range of potential targets of cyber attacks; or
>
> (iii) permitting the evasion of human control or oversight through means of deception or obfuscation.

---

regulations implementing the FMLA contain a much more detailed definition of the same term at 29 C.F.R. 825.113. Compare also 7 U.S.C. § 136(t) (statutory definition of "pest") with 40 C.F.R. 152.5 (more detailed regulatory definition of "pest").

> Models meet this definition even if they are provided to end users with technical safeguards that attempt to prevent users from taking advantage of the relevant unsafe capabilities.[17]

The executive order imposes certain reporting requirements on companies "developing or demonstrating an intent to develop" dual-use foundation models,[18] and for purposes of these requirements it instructs the Department of Commerce to "define, and thereafter update as needed on a regular basis, the set of technical conditions for models and computing clusters that would be subject to the reporting requirements."[19] In other words, EO 14110 contains both a high-level quasi-statutory[20] definition and a directive to an agency to promulgate a more detailed regulatory definition. The EO also provides a second definition that acts as a placeholder until the agency's regulatory definition is promulgated:

> any model that was trained using a quantity of computing power greater than 10^{26} integer or floating-point operations, or using primarily biological sequence data and using a quantity of computing power greater than 10^{23} integer or floating-point operations[21]

Unlike the first definition, which relies on subjective evaluations of model characteristics,[22] this placeholder definition provides a simple set of objective technical criteria that labs can consult to determine whether the reporting requirements apply. For general-purpose models, the sole test is whether the model was trained on computing power greater than 10^{26} integer or floating-point operations (FLOP); only models that exceed this compute threshold[23] are deemed "dual-use foundation models" for purposes of the reporting requirements mandated by EO 14110.

---

[17] 3 C.F.R. 14110 § 3(k).

[18] *Id.* at § 4.2(a).

[19] *Id.* at § 4.2(b)

[20] An executive order is not a statute, and a definition in an executive order does not enjoy the legitimacy and resiliency advantages that the legislative process confers upon an enacted law. The EO's definition is "quasi-statutory" in that it is phrased at a high level of generality and directs the enactment of a more specific regulatory definition, as statutory definitions often do.

[21] 3 C.F.R. 14110 § 4.2(b).

[22] For example, there is no simple, objective test for determining whether a model poses a "serious" risk to national economic security, and the EO does not provide criteria for determining what constitutes "broad data" or "high levels of performance."

[23] *See generally* Matteo Pistillo et al., *The Role of Compute Thresholds for AI Governance*, GEO. WASH. J. L & TECH (forthcoming 2025).

B. California's "Safe and Secure Innovation for Frontier Artificial Intelligence Act" (SB 1047)

California's recently vetoed "Safe and Secure Innovation for Frontier Artificial Intelligence Models Act" ("SB 1047") focused on a category that it referred to as "covered models."[24] The version of SB 1047 passed by the California Senate in May 2024 defined "covered model" to include models meeting either of the following criteria:

> (1) The artificial intelligence model was trained using a quantity of computing power greater than 10^{26} integer or floating-point operations.
>
> (2) The artificial intelligence model was trained using a quantity of computing power sufficiently large that it could reasonably be expected to have similar or greater performance as an artificial intelligence model trained using a quantity of computing power greater than 10^{26} integer or floating-point operations in 2024 as assessed using benchmarks commonly used to quantify the general performance of state-of-the-art foundation models.[25]

This definition resembles the placeholder definition in EO 14110 in that it primarily consists of a training compute threshold of 10^{26} FLOP. However, SB 1047 added an alternative capabilities-based threshold to capture future models which "could reasonably be expected" to be as capable as models trained on 10^{26} FLOP in 2024. This addition was intended to "future-proof"[26] SB 1047 by addressing one of the main disadvantages of training compute thresholds—their tendency to become obsolete over time as advances in algorithmic efficiency produce highly capable models trained on relatively small amounts of compute.[27]

Following pushback from stakeholders who argued that SB 1047 would stifle innovation,[28] the bill was amended repeatedly in the California State Assembly. The final version defined "covered model" in the following way:

> (A) Before January 1, 2027, "covered model" means either of the following:

---

[24] *See* Cal. SB-1047 § 3(f) (as amended in California Senate on May 16, 2024).

[25] *Id.*

[26] *See* Rebecca Crootof & B. J. Ard, *Structuring Techlaw*, 34 Harv. J. L. & Tech. 347, 400–01 (2021).

[27] *See generally* Konstantin Pilz et al., *Increased Compute Efficiency and the Diffusion of AI Capabilities* (2024), https://arxiv.org/abs/2311.15377.

[28] *See* Lauren Wagner, *California's Effort to Regulate AI Would Stifle Innovation*, The Information (2024), https://www.theinformation.com/articles/californias-flawed-effort-to-regulate-ai; Anjney Midha & Derrick Harris, *California's Senate Bill 1047: What You Need to Know*, a16z Podcast (2024), https://a16z.com/podcast/californias-senate-bill-1047-what-you-need-to-know/.

> (i) An artificial intelligence model trained using a quantity of computing power greater than 10^{26} integer or floating-point operations, the cost of which exceeds one hundred million dollars[29] ($100,000,000) when calculated using the average market prices of cloud compute at the start of training as reasonably assessed by the developer.
>
> (ii) An artificial intelligence model created by fine-tuning a covered model using a quantity of computing power equal to or greater than three times 10^{25} integer or floating-point operations, the cost of which, as reasonably assessed by the developer, exceeds ten million dollars ($10,000,000) if calculated using the average market price of cloud compute at the start of fine-tuning.
>
> (B) (i) Except as provided in clause (ii), on and after January 1, 2027, "covered model" means any of the following:
>
> (I) An artificial intelligence model trained using a quantity of computing power determined by the Government Operations Agency pursuant to Section 11547.6 of the Government Code, the cost of which exceeds one hundred million dollars ($100,000,000) when calculated using the average market price of cloud compute at the start of training as reasonably assessed by the developer.
>
> (II) An artificial intelligence model created by fine-tuning a covered model using a quantity of computing power that exceeds a threshold determined by the Government Operations Agency, the cost of which, as reasonably assessed by the developer, exceeds ten million dollars ($10,000,000) if calculated using the average market price of cloud compute at the start of fine-tuning.
>
> (ii) If the Government Operations Agency does not adopt a regulation governing subclauses (I) and (II) of clause (i) before January 1, 2027, the definition of "covered model" in subparagraph (A) shall be operative until the regulation is adopted.

This new definition was more complex than its predecessor. Subsection (A) introduced an initial definition slated to apply until at least 2027, which relied on a training compute threshold of 10^{26} FLOP paired with a training cost floor of $100,000,000.[30] Subsection (B), in turn, provided for the eventual replacement of the

[29] A subsequent subsection of SB 1047 provides that the specified dollar amounts are to be adjusted annually for inflation. *See* Cal. SB-1047 § 3(e)(2).

[30] The new definition also includes certain models created via compute-intensive fine-tuning of a covered model.

training compute thresholds used in the initial definition with new thresholds to be determined (and presumably updated) by a regulatory agency.

The most significant change in the final version of SB 1047's definition was the replacement of the capability threshold with a $100,000,000 cost threshold. Because it would currently cost more than $100,000,000 to train a model using $>10^{26}$ FLOP, the addition of the cost threshold did not change the scope of the definition in the short term. However, the cost of compute has historically fallen precipitously over time in accordance with Moore's law.[31] This may mean that models trained using significantly more than 10^{26} FLOP will cost significantly less than the inflation-adjusted equivalent of 100 million 2024 dollars to create at some point in the future.

The old capability threshold expanded the definition of "covered model" because it was an alternative to the compute threshold—models that exceeded either of the two thresholds would have been "covered." The newer cost threshold, on the other hand, restricted the scope of the definition because it was linked conjunctively to the compute threshold, meaning that only models that exceed both thresholds were covered. In other words, where the May 2024 definition of "covered model" future-proofed itself against the risk of becoming *underinclusive* by including highly capable low-compute models, the final definition instead guarded against the risk of becoming *overinclusive* by excluding low-cost models trained on large amounts of compute. Furthermore, the final cost threshold was baked into the bill text and could only have been changed by passing a new statute—unlike the compute threshold, which could have been specified and updated by a regulator.

Compared with the overall definitional scheme in EO 14110, SB 1047's definition was simpler, easier to operationalize, and less flexible. SB 1047 lacked a broad, high-level risk-based definition like the first definition in EO 14110. SB 1047 did resemble EO 14110 in its use of a "placeholder" definition, but where EO 14110 confers broad discretion on the regulator to choose the "set of technical conditions" that will comprise the regulatory definition, SB 1047 only authorized the regulator to set and adjust the numerical value of the compute thresholds in an otherwise rigid statutory definition.

C. EU Artificial Intelligence Act

The EU AI Act classifies AI systems according to the risks they pose. It prohibits systems that do certain things, such as exploiting the vulnerabilities of elderly or disabled people,[32] and regulates but does not ban so-called "high-risk" systems.[33] While this classification system does not map neatly onto U.S. regulatory efforts, the EU AI Act

[31] *See* Matteo Pistillo & Suzanne Van Arsdale, *The Role of Compute Thresholds for AI Governance*, GEO. WASH. J. L & TECH (forthcoming 2025) at 10–13.

[32] 2024 O.J. (L 1689) art. 5.

[33] *Id.* at art. 6.

does include a category conceptually similar to the EO's "dual-use foundation model": the "general-purpose AI model with systemic risk."[34] The statutory definition for this category includes a given general-purpose model[35] if:

> a. it has high impact capabilities[36] evaluated on the basis of appropriate technical tools and methodologies, including indicators and benchmarks; [or]
>
> b. based on a decision of the Commission,[37] ex officio or following a qualified alert from the scientific panel, it has capabilities or an impact equivalent to those set out in point (a) having regard to the criteria set out in Annex XIII.

Additionally, models are presumed to have "high impact capabilities" if they were trained on $>10^{25}$ FLOP.[38] The seven "criteria set out in Annex XIII" to be considered in evaluating model capabilities include a variety of technical inputs (such as the model's number of parameters and the size or quality of the dataset used in training the model), the model's performance on benchmarks and other capabilities evaluations, and other considerations such as the number of users the model has.[39] When necessary, the European Commission is authorized to amend the compute threshold and "supplement benchmarks and indicators" in response to technological developments, such as "algorithmic improvements or increased hardware efficiency."[40]

The EU Act definition resembles the initial, broad definition in the EO in that they both take diverse factors like the size and quality of the dataset used to train the model,

---

[34] *Id.* at art. 51.

[35] "General-purpose AI model" is defined elsewhere in the Act to mean "an AI model, including where such an AI model is trained with a large amount of data using self-supervision at scale, that displays significant generality and is capable of competently performing a wide range of distinct tasks regardless of the way the model is placed on the market and that can be integrated into a variety of downstream systems or applications, except AI models that are used for research, development or prototyping activities before they are placed on the market." *Id.* at art. 3(63).

[36] "High impact capabilities" is defined elsewhere in the Act to mean "capabilities that match or exceed the capabilities recorded in the most advanced general-purpose AI models."

[37] The "Commission" is the European Commission, i.e., the European Union's executive body.

[38] 2024 O.J. (L 1689) art. 51(2). Note that this threshold is one order of magnitude below the threshold specified in EO 14110 and SB 1047. This difference is potentially politically significant because state-of-the-art models built by leading U.S. labs, such as OpenAI's GPT-4 and Google's Gemini, are thought to have been trained on between 10^{25} and 10^{26} FLOP and therefore already fall within the scope of the Act, whereas the latest model released by Europe's leading AI lab, Mistral, is thought to have been trained on less than 10^{25} FLOP. *See* Matthew Barnett, *A Compute-Based Framework for Thinking About the Future of AI*, EPOCH AI (last updated Aug. 10, 2023), https://epochai.org/blog/a-compute-based-framework-for-thinking-about-the-future-of-ai; *see also Notable AI Models*, EPOCH AI (last updated Sept. 9, 2024), https://epochai.org/data/notable-ai-models.

[39] 2024 O.J. (L 1689) annex XIII.

[40] 2024 O.J. (L 1689) art. 51(3).

the number of parameters, and the model's capabilities into account. However, the EU Act definition is likely much broader than either EO definition. The training compute threshold in the EU Act is sufficient, but not necessary, to classify models as systemically risky, whereas the (much higher) threshold in the EO's placeholder definition is both necessary and sufficient. And the first EO definition includes only models that exhibit a high level of performance on tasks that pose serious risks to national security, while the EU Act includes all general-purpose models with "high impact capabilities," which it defines as including any model trained on more than 10^{25} FLOP.

The EU Act definition resembles the final SB 1047 definition of "covered model" in that both definitions authorize a regulator to update their thresholds in response to changing circumstances. It also resembles SB 1047's May 2024 definition in that both definitions incorporate a training compute threshold and a capabilities-based element.

IV. Elements of Existing Definitions

As the examples discussed above demonstrate, legal definitions of "frontier model" can consist of one or more of a number of criteria. This section discusses a few of the most promising definitional elements.

A. Technical inputs and characteristics

A definition may classify AI models according to their technical characteristics or the technical inputs used in training the model, such as training compute, parameter count, and dataset size and type. These elements can be used in either statutory or regulatory definitions.

Training compute thresholds are a particularly attractive option for policymakers,[41] as evidenced by the three examples discussed above. "Training compute" refers to the computational power used to train a model, often measured in integer or floating-point operations (OP or FLOP).[42] Training compute thresholds function as a useful proxy for

[41] *See* Pistillo et al. *supra* note 23 at 19–33; Anderljung et al., *supra* note 1 at 35–36; Toner & Fist, *supra* note 6.

[42] Compute usage is usually calculated based on the number of FLOP used in a model's final pre-training run, and compute usage from previous test runs or from fine-tuning the model after it has been pre-trained is typically not counted. *See* Pistillo et al. *supra* note 23 at 6. However, the EU Act specifically considers the total, cumulative amount of compute used during the entire training process, including "pre-training, synthetic data generation, and fine-tuning." 2024 O.J. (L 1689) recital 111. Additionally, the language of EO 14110 is arguably ambiguous with respect to whether compute used during the process of training a model for purposes other than the final pre-training run counts towards the included compute threshold. *See* 3 C.F.R. 14110 § 4.2(b) ("any model that was trained using a quantity of computing power greater than…").

model capabilities because capabilities tend to increase as computational resources used to train the model increase.[43]

One advantage of using a compute threshold is that training compute is a straightforward metric that is quantifiable and can be readily measured, monitored, and verified.[44] Because of these characteristics, determining with high certainty whether a given model exceeds a compute threshold is relatively easy. This, in turn, facilitates enforcement of and compliance with regulations that rely on a compute-based definition. Since the amount of training compute (and other technical inputs) can be estimated prior to the training run,[45] developers can predict whether a model will be covered earlier in development.

One disadvantage of a compute-based definition is that compute thresholds are a proxy for model capabilities, which are in turn a proxy for risk. Definitions that make use of multiple nested layers of proxy terms in this manner are particularly prone to becoming untethered from their original purpose.[46] This can be caused, for example, by the operation of Goodhart's Law, which suggests that "when a measure becomes a target, it ceases to be a good measure."[47] Particularly problematic, especially for statutory definitions that are more difficult to update, is the possibility that a compute threshold may become underinclusive over time as improvements in algorithmic efficiency allow for the development of highly capable models trained on below-threshold levels of compute.[48] This possibility is one reason why SB 1047 and the EU AI Act both supplement their compute thresholds with alternative, capabilities-based elements.

---

[43] *See* Anderljung et al., *supra* note 1 at 37 ("Empirically, scaling training compute has reliably led to better performance on many of the tasks AI models are trained to solve, and many similar downstream tasks."); Deep Ganguli et al., *Predictability and Surprise in Large Generative Models*, *in* 2022 ACM Conference on Fairness, Accountability, and Transparency 1747 (2022), https://arxiv.org/abs/2202.07785 at 2–6 ("In most cases, these scaling laws predict a continued increase in certain capabilities as models get larger. . . . More precisely, by general capability scaling we mean two things. First, the training (and test) loss improves predictably with scale on a broad data distribution. Second, this improvement in loss tends to correlate on average with increased performance on a number of downstream tasks."); Gwern Branwen, *The Scaling Hypothesis*, Gwern.net (May 28, 2020), https://gwern.net/scaling-hypothesis.

[44] *See* Anderljung et al., supra note 1 at 35–36; Toner & Fist, *supra* note 6; Girish Sastry et al., *Computing Power and the Governance of Artificial Intelligence* (2024), https://arxiv.org/abs/2402.08797 at 19–33.

[45] *See, e.g.,* 2024 O.J. (L 1689), at Foreword (112) ("[T]raining of general purpose AI models takes considerable planning which includes the upfront allocation of compute resources and, therefore, providers of general purpose AI models are able to know if their model would meet the threshold before the training is completed.").

[46] *See* Winter & Bullock, *supra* note 8 (discussing historical examples of this phenomenon in U.S. regulation).

[47] *See* Jacob Hilton & Leo Gao, *Measuring Goodhart's Law*, OpenAI (2022), https://openai.com/index/measuring-goodharts-law/.

[48] *See* Pilz et al., *supra* note 28.

In addition to training compute, two other model characteristics correlated with capabilities are the number of model parameters[49] and the size of the dataset on which the model was trained.[50] Either or both of these characteristics can be used as an element of a definition. A definition can also rely on training data characteristics other than size, such as the quality or type of the data used; the placeholder definition in EO 14110, for example, contains a lower compute threshold for models "trained… using primarily biological sequence data."[51] EO 14110 requires a dual-use foundation model to contain "at least tens of billions of parameters,"[52] and the "number of parameters of the model" is a criteria to be considered under the EU AI Act.[53] EO 14110 specified that only models "trained on broad data" could be dual-use foundation models,[54] and the EU AI Act includes "the quality or size of the data set, for example measured through tokens" as one criterion for determining whether an AI model poses systemic risks.[55]

Dataset size and parameter count share many of the pros and cons of training compute. Like training compute, they are objective metrics that can be measured and verified, and they serve as proxies for model capabilities.[56] Training compute is often considered the best and most reliable proxy of the three, in part because it is the most closely correlated with performance and is difficult to manipulate.[57] However, partially redundant backup metrics can still be useful.[58] Dataset characteristics other than size are typically less quantifiable and harder to measure but are also capable of capturing information that the quantifiable metrics cannot.

B. Capabilities

Frontier models can also be defined in terms of their capabilities. A capabilities-based definition element typically sets a threshold level of competence that a model must achieve to be considered "frontier," either in one or more specific domains or

---

[49] Parameters are the learnable variables or weights of a model. Generally, more complex models have more parameters.

[50] In large language models, "[l]anguage modeling performance improves smoothly as we increase the model size, dataset size, and amount of compute used for training… Empirical performance has a power-law relationship with each individual factor when not bottlenecked by the other two." Jared Kaplan et al., *Scaling Laws for Neural Language Models* (2020), https://arxiv.org/abs/2001.08361 at 3.

[51] 3 C.F.R. 14110 § 4.2(b).

[52] 3 C.F.R. 14110 § 3(k).

[53] 2024 O.J. (L 1689) annex XIII(b).

[54] 3 C.F.R. 14110 § 3(k).

[55] 2024 O.J. (L 1689) annex XIII(a).

[56] *See* Kaplan et al., *supra* note 51, at 3.

[57] *See* Lennart Heim & Leonie Koessler, *Training Compute Thresholds: Features and Functions in AI Regulation* (2024), https://arxiv.org/abs/2405.10799. Parameter count, for example, is "manipulable through techniques like model pruning." *Id.*

[58] For example, a model trained on large amounts of compute but on a very small dataset would not be highly capable; therefore, a definition that included both training compute and dataset size thresholds would properly exclude it.

across a broad range of domains. A capabilities-based definition can provide specific, objective criteria for measuring a model's capabilities,[59] or it can describe the capabilities required in more general terms and leave the task of evaluation to the discretion of future interpreters.[60] The former approach might be better suited to a regulatory definition, especially if the criteria used will have to be updated frequently, whereas the latter approach would be more typical of a high-level statutory definition.

Basing a definition on capabilities, rather than relying on a proxy for capabilities like training compute, eliminates the risk that the chosen proxy will cease to be a good measure of capabilities over time. Therefore, a capabilities-based definition is more likely than, e.g., a compute threshold to remain robust over time in the face of improvements in algorithmic efficiency. This was the point of the May 2024 version of SB 1047's use of a capabilities element tethered to a compute threshold ("similar or greater performance as an artificial intelligence model trained using a quantity of computing power greater than 10^{26} integer or floating-point operations in 2024")—it was an attempt to capture some of the benefits of an input-based definition while also guarding against the possibility that models trained on less than 10^{26} FLOP may become far more capable in the future than they are in 2024.

However, capabilities are far more difficult than compute to accurately measure. Whether a model has demonstrated "high levels of performance at tasks that pose a serious risk to security" under the EO's broad capabilities-based definition is not something that can be determined objectively and to a high degree of certainty like the size of a dataset in tokens or the total FLOP used in a training run. Model capabilities are often measured using benchmarks (standardized sets of tasks or questions),[61] but creating

---

[59] For example, capabilities could be evaluated based on model performance on a specified set of benchmark tests. *See, e.g.,* Google et al., *Frontier Model Forum: A New Partnership to Promote Responsible AI*, GOOGLE (July 26, 2023), https://blog.google/outreach-initiatives/public-policy/google-microsoft-openai-anthropic-frontier-model-forum/ ("There will be a strong focus initially on developing and sharing a public library of technical evaluations and benchmarks for frontier AI models."). SB 1047 and the EU AI Act both mention performance on commonly-used or standard benchmarks, but neither specifies the precise benchmarks to be used.

[60] This is the approach favored by the first EO definition, which requires "high levels of performance" at tasks that threaten national security, but does not elaborate on what constitutes a "high level" of performance except by offering a non-exhaustive list of examples, e.g., "enabling powerful offensive cyber operations through automated vulnerability discovery and exploitation." 3 C.F.R. 14110 § 3(k).

[61] *See, e.g.,* Aarohi Srivastava et al., *Beyond the Imitation Game: Quantifying and Extrapolating the Capabilities of Language Models* (2022), https://arxiv.org/abs/2206.04615v3; Melanie Mitchell, Alessandro B. Palmarini & Arseny Moskvichev, *Comparing Humans, GPT-4, and GPT-4V On Abstraction and Reasoning Tasks* (2023), https://arxiv.org/abs/2311.09247v3.

benchmarks that accurately measure the complex and diverse capabilities of general-purpose foundation models[62] is notoriously difficult.[63]

Additionally, model capabilities (unlike the technical inputs discussed above) are generally not measurable until after the model has been trained.[64] This makes it difficult to regulate the development of frontier models using capabilities-based definitions, although post-development, pre-release regulation is still possible.

C. Risk

Some researchers have suggested the possibility of defining frontier AI systems on the basis of the risks they pose to users or to public safety instead of or in addition to relying on a proxy metric, like capabilities, or a proxy for a proxy, such as compute.[65] The principal advantage of this direct approach is that it can, in theory, allow for better-targeted regulations—for instance, by allowing a definition to exclude highly capable but demonstrably low-risk models. The principal disadvantage is that measuring risk is even more difficult than measuring capabilities.[66] The science of designing rigorous safety evaluations for foundation models is still in its infancy.[67]

Of the three real-world measures discussed in Section III, only EO 14110 mentions risk directly. The broad initial definition of "dual-use foundation model" includes models that exhibit "high levels of performance at tasks that pose a serious risk to security," such as "enabling powerful offensive cyber operations through automated vulnerability discovery" or making it easier for non-experts to design chemical weapons. This is a capability threshold combined with a risk threshold; the tasks at which a dual-use foundation model must be highly capable are those that pose a "serious risk" to security, national economic security, and/or national public health or safety. As EO 14110

---

[62] "A "foundation model" is a large general-purpose AI model that can be used for a wide variety of tasks. State-of-the-art LLMs like GPT-4, Claude Opus, and Gemini Ultra are foundation models. *See* Rishi Bommasani et al., On the Opportunities and Risks of Foundation Models (2021), https://crfm.stanford.edu/assets/report.pdf. The term "foundation" is used because the initial training run provides a versatile base model onto which additional capabilities can be added via fine-tuning. *See* Helen Toner, *What Are Generative AI, Large Language Models, and Foundation Models?*, Center for Security and Emerging Technology (May 12, 2023), https://cset.georgetown.edu/article/what-are-generative-ai-large-language-models-and-foundation-models/.

[63] *See* Matthew Stewart, *The Olympics of AI: Benchmarking Machine Learning Systems*, Towards Data Science (Sep. 22, 2023), https://towardsdatascience.com/the-olympics-of-ai-benchmarking-machine-learning-systems-c4b2051fbd2b.

[64] Heim & Kossler, *supra* note 58, at 25–26.

[65] *See, e.g.,* Leonie Koessler, Jonas Schuett & Markus Anderljung, *Risk Thresholds for Frontier AI* (2024), https://arxiv.org/abs/2406.14713v1.

[66] *See id.* at 14 ("The key argument against using risk thresholds is that risk estimation is extremely hard for risks from frontier AI development and deployment.").

[67] *See* Laura Weidinger et al., *Holistic Safety and Responsibility Evaluations of Advanced AI Models* (2024), https://arxiv.org/abs/2404.14068v1.

shows, risk-based definition elements can specify the type of risk that a frontier model must create instead of addressing the severity of the risks created.

D. Epistemic elements

One of the primary justifications for recognizing a category of "frontier models" is the likelihood that broadly capable AI models that are more advanced than previous generations of models will have capabilities and applications that are not readily predictable *ex ante*.[68] As the word "frontier" implies, lawmakers and regulators focusing on frontier models are interested in targeting models that break new ground and push into the unknown.[69] This was, at least in part, the reason for the inclusion of training compute thresholds of 10^{26} FLOP in EO 14110 and SB 1047—since the most capable current models were trained on $5x10^{25}$ or fewer FLOP,[70] a model trained on 10^{26} FLOP would represent a significant step forward into uncharted territory.

While it is possible to target models that advance the state of the art by setting and adjusting capability or compute thresholds, a more direct alternative approach would be to include an epistemic element in a statutory definition of "frontier model." An epistemic element would distinguish between "known" and "unknown" models, i.e., between well-understood models that pose only known risks and poorly understood models that may pose unfamiliar and unpredictable risks.[71]

This kind of distinction between known and unknown risks has a long history in U.S. regulation.[72] For instance, the Toxic Substances Control Act (TSCA) prohibits the manufacturing of any "new chemical substance" without a license.[73] The EPA keeps and regularly updates a list of chemical substances which are or have been manufactured in

---

[68] *See* Toby Shevlane et al., *Model Evaluation for Extreme Risks* (2023), https://arxiv.org/abs/2305.15324 ("[F]rontier models are uniquely risky because (a) more capable models can excel at a wider range of tasks, which will unlock more opportunities to cause harm; and (b) novel models are less well-understood by the research community."); Anderljung et al., *supra* note 1 at 10–13.

[69] *See* Anderljung et al., *supra* note 1 at 9–13, 34–36; Margot E. Kaminski, *Regulating the Risks of AI*, 103 B.U. L. Rev. 1347, 1368 (2022); Sella Nevo et al., Securing AI Model Weights: Preventing Theft and Misuse of Frontier Models (2024), https://www.rand.org/pubs/research_reports/RRA2849-1.html ("Frontier models are those that match or exceed the capabilities of the most advanced AI models at the time of their development." (emphasis omitted)).

[70] *See Notable AI Models*, *supra* note 39.

[71] *Cf.* Toner & Fist, *supra* note 6 ("Firstly, it's important to distinguish between two different categories of systems that might be meant by frontier AI: (1) Models that are at or beyond the current cutting edge, which 'expand the frontier' or 'push into the unknown,' in that they are pushing the limits of what AI can do. (2) Any model that is considered to be above a certain threshold of riskiness. Definition (1), which would move over time as the field advances, is a more intuitive use of the word 'frontier.' Using this definition, one might distinguish between a 'thin' frontier—i.e., models with literally unprecedented capabilities or scale—and a 'thicker' frontier that would move over time but still include some systems behind the absolute bleeding edge.").

[72] *See* Peter Huber, *The Old-New Division in Risk Regulation*, 69 Va. L. Rev. 1025 (1983).

[73] 15 U.S.C. § 2604 ¶ (a); *see id.* at 1038.

the U.S., and any substance not included on this list is "new" by definition.[74] In other words, the TSCA distinguishes between chemicals (including potentially dangerous chemicals) that are familiar to regulators and unfamiliar chemicals that pose unknown risks.

One advantage of an epistemic element is that it allows a regulator to address "unknown unknowns" separately from better-understood risks that can be evaluated and mitigated more precisely.[75] Additionally, the scope of an epistemic definition, unlike that of most input- and capability-based definitions, would change over time as regulators became familiar with the capabilities of and risks posed by new models.[76] Models would drop out of the "frontier" category once regulators became sufficiently familiar with their capabilities and risks.[77] Like a capabilities- or risk-based definition, however, an epistemic definition might be difficult to operationalize.[78] To determine whether a given model was "frontier" under an epistemic definition, it would probably be necessary to either rely on a proxy for unknown capabilities or authorize a regulator to categorize eligible models according to a specified process.[79]

E. Deployment context

The context in which an AI system is deployed can serve as an element in a definition. The EU AI Act, for example, takes the number of registered end users and the number of registered EU business users a model has into account as factors to be

[74] *See* 15 U.S.C. §§ 2602 ¶ (11), 2607.

[75] *Cf.* Anderljung et al., *supra* note 1 at 26–27 (recommending that regulations of model deployment be tailored to the assessed risk of the model).

[76] *See* Anderljung et al., *supra* note 1 at 34 ("[Frontier AI] is a dynamic category. In particular, what may initially qualify as a frontier AI model could change over time due to improvements in society's defenses against advanced AI models and an improved understanding of the nature of the risks posed."), 36 ("A related approach [to defining 'frontier AI model'] could be to define the regulatory target by reference to the most capable broad models that have been shown not to have sufficiently dangerous capabilities. The idea here is that, if a model has been shown not to have sufficiently dangerous capabilities, then every model that can be expected to perform worse than it should also not be expected to have sufficiently dangerous capabilities. Regulation would then apply only to those models that exceed the capabilities of models known to lack sufficiently dangerous capabilities. This approach has the benefit of updating quickly based on observations from newer models. It would also narrow the space of regulated models over time, as regulators learn more about which models have sufficiently dangerous capabilities."); Shevlane et al., *supra* note 69 at 3 (presenting an illustrative model of frontier AI development in which AI capabilities progress towards models that carry extreme risks).

[77] *See* Toner & Fist, *supra* note 6 ("Using this [epistemic] definition, one might distinguish between a 'thin' frontier—i.e., models with literally unprecedented capabilities or scale—and a 'thicker' frontier that would move over time but still include some systems behind the absolute bleeding edge.")

[78] *Cf.* Daniel Carpenter & Carson Ezell, *An FDA for AI? Pitfalls and Plausibility of Approval Regulation for Frontier Artificial Intelligence* (2024), https://arxiv.org/abs/2408.00821 (discussing challenges to adequately defining regulated units in FDA-like regulation of AI systems).

[79] *Cf.* Toner & Fist, *supra* note 6 ("[T]he regulator would, of course, need to be authorized to update the criteria [used to define 'frontier AI'] regularly to keep pace with changes in the field in consultation with relevant experts.").

considered in determining whether a model is a "general-purpose AI model with systemic risk."[80] Deployment context typically does not in and of itself provide enough information about the risks posed by a model to function as a stand-alone definitional element, but it can be a useful proxy for the kind of risk posed by a given model. Some models may cause harms in proportion to their number of users, and the justification for aggressively regulating these models grows stronger the more users they have. A model that will only be used by government agencies, or by the military, creates a different set of risks than a model that is made available to the general public.

V. Updating Regulatory Definitions

A recurring theme in the scholarly literature on the regulation of emerging technologies is the importance of regulatory flexibility.[81] Because of the rapid pace of technological progress, legal rules designed to govern emerging technologies like AI tend to quickly become outdated and ineffective if they cannot be rapidly and frequently updated in response to changing circumstances.[82] For this reason, it may be desirable to authorize an executive agency to promulgate and update a regulatory definition of "frontier model," since regulatory definitions can typically be updated more frequently and more easily than statutory definitions under U.S. law.[83]

Historically, failing to quickly update regulatory definitions in the context of emerging technologies has often led to the definitions becoming obsolete or counterproductive. For example, U.S. export controls on supercomputers in the 1990s and early 2000s defined "supercomputer" in terms of the number of millions of theoretical operations per second (MTOPS) the computer could perform.[84] Rapid advances in the processing power of commercially available computers soon rendered the initial definition obsolete, however, and the Clinton administration was forced to revise the MTOPS threshold repeatedly to avoid harming the competitiveness of the American computer industry.[85] Eventually, the MTOPS metric itself was rendered obsolete, leading

---

[80] 2024 O.J. (L 1689) annex XIII.

[81] *See* Crootof & Ard, *supra* note 27 at 347, 400–408; Matthijs Maas, Artificial Intelligence Governance Under Change: Foundations, Facets, Frameworks (2021) (Ph.D. dissertation, University of Copenhagen). (discussing the need for flexible and dynamic regulatory regimes to effectively govern emerging technologies).

[82] *See* Winter & Charlie, *supra* note 8; Hadassah Drukarch, et al., *An Iterative Regulatory Process for Robot Governance*, 5 Data & Policy e8 (2023); Gary E Marchant & Yvonne A Stevens, *Resilience: A New Tool in the Risk Governance Toolbox for Emerging Technologies*, 51 UC Davis L. Rev 233, 252 (2017).

[83] *See* Section II, *supra*.

[84] Charles J. Holland et al., Export Control of High Performance Computing: Analysis and Alternative Strategies: Analysis and Alternative Strategies (2001), https://apps.dtic.mil/sti/tr/pdf/ADA397730.pdf.

[85] *Id.* at 2.

to a period of several years in which supercomputer export controls were ineffective at best.[86]

There are a number of legal considerations that may prevent an agency from quickly updating a regulatory definition and a number of measures that can be taken to streamline the process. One important aspect of the rulemaking process is the Administrative Procedure Act's "notice and comment" requirement.[87] In order to satisfy this requirement, agencies are generally obligated to publish notice of any proposed amendment to an existing regulation in the Federal Register, allow time for the public to comment on the proposal, respond to public comments, publish a final version of the new rule, and then allow at least 30–60 days before the rule goes into effect.[88] From the beginning of the notice-and-comment process to the publication of a final rule, this process can take anywhere from several months to several years.[89] However, an agency can waive the 30–60 day publication period or even the entire notice-and-comment requirement for "good cause" if observing the standard procedures would be "impracticable, unnecessary, or contrary to the public interest."[90] Of course, the notice-and-comment process has benefits as well as costs; public input can be substantively valuable and informative for agencies, and also increases the democratic accountability of agencies and the transparency of the rulemaking process. In certain circumstances, however, the costs of delay can outweigh the benefits. U.S. agencies have occasionally demonstrated a willingness to waive procedural rulemaking requirements in order to respond to emergency AI-related developments. The Bureau of Industry and Security ("BIS"), for example, waived the normal 30-day waiting period for an interim rule prohibiting the sale of certain advanced AI-relevant chips to China in October 2023.[91]

Another way to encourage quick updating for regulatory definitions is for Congress to statutorily authorize agencies to eschew or limit the length of notice and comment, or to compel agencies to promulgate a final rule by a specified deadline.[92] Because notice and comment is a statutory requirement, it can be adjusted as necessary by statute.

---

[86] Colin B. Picker, *A View from 40,000 Feet: International Law and the Invisible Hand of Technology*, 232 Cardozo L. Rev. 149, 209–210 (2001).

[87] 5 U.S.C. § 553.

[88] 5 U.S.C. § 553 ¶ (d).

[89] U.S. Gov't Accountability Off., GAO-09-205, Improvements Needed to Monitoring and Evaluation of Rules Development as Well as to the Transparency of OMB Regulatory Reviews (2009).

[90] 5 U.S.C. § 553 ¶ (b)(3)(b).

[91] NVIDIA Corporation, SEC Form 8-K Filing (Oct. 24, 2023), https://www.sec.gov/Archives/edgar/data/1045810/000104581023000221/nvda-20231023.htm; *see* Martin & Bloomberg, *'We Cannot Let China Get These Chips': Commerce Secretary Raimondo Says More Funding Needed for AI Export Controls*, Fortune (Dec. 2, 2023), https://fortune.com/2023/12/02/ai-chip-export-controls-china-nvidia-raimondo/.

[92] *See* Kevin J Hickey, Cong. Rsch. Serv., R45336, Agency Delay: Congressional and Judicial Means to Expedite Agency Rulemaking (2018); Jacob E Gersen & Anne Joseph O'Connell, *Deadlines in Administrative Law*, 156 U. PA. L. Rev 923, 945–949 (2008).

For regulations exceeding a certain threshold of economic significance, another substantial source of delay is OIRA review. OIRA, the Office of Information and Regulatory Affairs, is an office within the White House that oversees interagency coordination and undertakes centralized cost-benefit analysis of important regulations.[93] Like notice and comment, OIRA review can have significant benefits—such as improving the quality of regulations and facilitating interagency cooperation—but it also delays the implementation of significant rules, typically by several months.[94] OIRA review can be waived either by statutory mandate or by OIRA itself.[95]

VI. Deference, Delegation, and Regulatory Definitions

Recent developments in U.S. administrative law may make it more difficult for Congress to effectively delegate the task of defining "frontier model" to a regulatory agency. A number of recent Supreme Court cases signal an ongoing shift in U.S. administrative law doctrine intended to limit congressional delegations of rulemaking authority.[96] Whether this development is good or bad on net is a matter of perspective; libertarian-minded observers who believe that the U.S. has too many legal rules already[97] and that overregulation is a bigger problem than underregulation have welcomed the change,[98] while pro-regulation observers predict that it will significantly reduce the regulatory capacity of agencies in a number of important areas.[99]

Regardless of where one falls on that spectrum of opinion, the relevant takeaway for efforts to define "frontier model" is that it will likely become somewhat more difficult for agencies to promulgate and update regulatory definitions without a clear statutory

---

[93] Cass R. Sunstein, *The Office of Information and Regulatory Affairs: Myths and Realities*, 126 HARV. L. REV 1838 (2013).

[94] Curtis W Copeland, *Length of Rule Reviews by the Office of Information and Regulatory Affairs*, *in* ADMINISTRATIVE CONFERENCE OF THE UNITED STATES (2013) at 4.

[95] Regulatory Planning and Review, 58 Fed. Reg. 51735 (Sept. 30 1993) ("The Administrator of OIRA may waive review of any planned regulatory action designated by the agency as significant.").

[96] *See* Penn Program on Regulation, , *What Are the Implications of the End of Chevron for U.S. Administrative Law?*, YOUTUBE (JULY 3, 2024), https://youtu.be/pszALge4TYo ("The current Court has one idea… it's that Congress over-delegates. Almost every major doctrine it's developed or abandoned can be explained by an underlying concern with congressional delegation.").

[97] *See* Neil Gorsuch & Janie Nitze, *Neil Gorsuch: America Has Too Many Laws*, THE ATLANTIC (Aug. 5, 2024), https://www.theatlantic.com/ideas/archive/2024/08/america-has-too-many-laws-neil-gorsuch/679237/.

[98] *See* Adam J. White, *Loper Bright and the End of Administrative Exceptionalism*, AMERICAN ENTERPRISE INSTITUTE (July 10, 2024), https://www.aei.org/op-eds/loper-bright-and-the-end-of-administrative-exceptionalism/.

[99] *See* Cary Coglianese, *The Supreme Court's Judicial Earthquake Will Shake the Administrative State*, BARRON'S (July 2, 2024), https://www.barrons.com/articles/supreme-court-decision-chevron-administrative-state-a3fcb801; Coral Davenport et al., *Here's What the Court's Chevron Ruling Could Mean in Everyday Terms*, THE NEW YORK TIMES (June 28, 2024), https://www.nytimes.com/2024/06/28/us/politics/chevron-deference-decision-meaning.html.

authorization to do so. If Congress still wishes to authorize the creation of regulatory definitions, however, it can protect agency definitions from legal challenges by clearly and explicitly authorizing agencies to exercise discretion in promulgating and updating definitions of specific terms.

A. Loper Bright *and deference to agency interpretations*

In a recent decision in the combined cases of *Loper Bright Enterprises v. Raimondo* and *Relentless v. Department of Commerce*, the Supreme Court repealed a longstanding legal doctrine known as *Chevron* deference.[100] Under *Chevron*, federal courts were required to defer to certain agency interpretations of federal statutes when (1) the relevant part of the statute being interpreted was genuinely ambiguous and (2) the agency's interpretation was reasonable. After *Loper Bright*, courts are no longer required to defer to these interpretations—instead, under a doctrine known as *Skidmore* deference,[101] agency interpretations will prevail in court only to the extent that courts are persuaded by them.[102]

Justice Elena Kagan's dissenting opinion in *Loper Bright* argues that the decision will harm the regulatory capacity of agencies by reducing the ability of agency subject-matter experts to promulgate regulatory definitions of ambiguous statutory phrases in "scientific or technical" areas.[103] The dissent specifically warns that, after *Loper Bright*, courts will "play a commanding role" in resolving questions like "[w]hat rules are going to constrain the development of A.I.?"[104]

Justice Kagan's dissent probably somewhat overstates the significance of *Loper Bright* to AI governance for rhetorical effect.[105] The end of *Chevron* deference does not mean that Congress has completely lost the ability to authorize regulatory definitions; where Congress has explicitly directed an agency to define a specific statutory term, *Loper Bright* will not prevent the agency from doing so.[106] An agency's authority to

---

[100] So named after the 1984 Supreme Court case in which the doctrine was first articulated, Chevron U.S.A., v. Natural Resources Defense Council, 467 U.S. 837 (1984).

[101] *See* Skidmore v. Swift & Co., 323 U.S. 134 (1944).

[102] *See* Gonzales v. Oregon, 546 U.S. 243, 256 (2006).

[103] *See* Loper Bright Enterprises v. Raimondo, 603 U.S. __ (2024) at 5–11.

[104] *Id.* at 32

[105] Estimates of the impact that *Loper Bright* will have on the regulatory capacity of the administrative state vary dramatically. Some scholars predict that *Loper Bright* "will [not] make any major difference, legally or substantively," because courts will simply find implicit delegations of authority in most of the cases that would have been decided under *Chevron* prior to *Loper Bright*. Adrian Vermeule, *Chevron By Any Other Name*, The New Digest (June 28, 2024), https://thenewdigest.substack.com/p/chevron-by-any-other-name. Others foresee a significant increase in the likelihood that rules will be challenged, an increase in the likelihood of a given challenge succeeding, and a decrease in the willingness of litigation-averse agencies to regulate aggressively. *See* Cass R. Sunstein, *The Consequences of Loper Bright* (2024), https://doi.org/10.2139/ssrn.4881501.

[106] *See Loper Bright* at 17–18.

promulgate a regulatory definition under a statute resembling EO 14110, which explicitly directs the Department of Commerce to define "dual-use foundation model," would likely be unaffected. However, *Loper Bright* has created a great deal of uncertainty regarding the extent to which courts will accept agency claims that Congress has *implicitly* authorized the creation of regulatory definitions.[107]

To better understand how this uncertainty might affect efforts to define "frontier model," consider the following real-life example. The Energy Policy and Conservation Act ("EPCA") includes a statutory definition of the term "small electric motor."[108] Like many statutory definitions, however, this definition is not detailed enough to resolve all disputes about whether a given product is or is not a "small electric motor" for purposes of EPCA. In 2010, the Department of Energy ("DOE"), which is authorized under EPCA to promulgate energy efficiency standards governing "small electric motors,"[109] issued a regulatory definition of "small electric motor" specifying that the term referred to motors with power outputs between 0.25 and 3 horsepower.[110] The National Electrical Manufacturers Association ("NEMA"), a trade association of electronics manufacturers, sued to challenge the rule, arguing that motors with between 1 and 3 horsepower were too powerful to be "small electric motors" and that the DOE was exceeding its statutory authority by attempting to regulate them.[111]

In a 2011 opinion that utilized the *Chevron* framework, the federal court that decided NEMA's lawsuit considered the language of EPCA's statutory definition and concluded that EPCA was ambiguous as to whether motors with between 1 and 3 horsepower could be "small electric motors."[112] The court then found that the DOE's regulatory definition was a reasonable interpretation of EPCA's statutory definition, deferred to the DOE under *Chevron*, and upheld the challenged regulation.[113]

Under *Chevron*, federal courts were required to assume that Congress had implicitly authorized agencies like the DOE to resolve ambiguities in a statute, as the DOE did in 2010 by promulgating its regulatory definition of "small electric motor." After *Loper Bright*, courts will recognize fewer implicit delegations of definition-making authority. For instance, while EPCA requires the DOE to prescribe "testing requirements" and "energy conservation standards" for small electric motors, it does not explicitly authorize the DOE to promulgate a regulatory definition of "small electric motor." If a

---

[107] *See* Adrian Vermeule, *Implied Delegations After Loper*, YALE JOURNAL ON REGULATION: NOTICE & COMMENT (Jul. 9, 2024), https://www.yalejreg.com/nc/implied-delegations-after-loper-by-adrian-vermeule/.

[108] *See* 42 U.S.C. § 6311 ¶ (13)(g) ("The term 'small electric motor' means a NEMA general purpose alternating current single-speed induction motor, built in a two-digit frame number series in accordance with NEMA Standards Publication MG1-1987.").

[109] *See* 42 U.S.C. § 6317 ¶(b)

[110] *See* 75 Fed. Reg. 10874 (Mar. 9, 2010)

[111] *See* Nat'l Elec. Mfrs. Ass'n v. U.S. Dep't of Energy, 654 F.3d 496, 497 (4th Cir. 2011)

[112] *See id.* at 507.

[113] *Id.* at 512.

rule like the one challenged by NEMA were challenged today, the DOE could still argue that Congress implicitly authorized the creation of such a rule by giving the DOE authority to prescribe standards and testing requirements—but such an argument would probably be less likely to succeed than the *Chevron* argument that saved the rule in 2011.

Today, a court that did not find an implicit delegation of rulemaking authority in EPCA would not defer to the DOE's interpretation. Instead, the court would simply compare the DOE's regulatory definition of "small electric motor" with NEMA's proposed definition and decide which of the two was a more faithful interpretation of EPCA's statutory definition.[114] Similarly, when or if some future federal statute uses the phrase "frontier model" or any analogous term, agency attempts to operationalize the statute by enacting detailed regulatory definitions that are not explicitly authorized by the statute will be easier to challenge after *Loper Bright* than they would have been under *Chevron*.

Congress can avoid *Loper Bright* issues by using clear and explicit statutory language to authorize agencies to promulgate and update regulatory definitions of "frontier model" or analogous phrases. However, it is often difficult to predict in advance whether or how a statutory definition will become ambiguous over time. This is especially true in the context of emerging technologies like AI, where the rapid pace of technological development and the poorly understood nature of the technology often eventually render carefully crafted definitions obsolete.[115]

Suppose, for example, that a federal statute resembling the May 2024 draft of SB 1047 was enacted. The statutory definition would include future models trained on a quantity of compute such that they "could reasonably be expected to have similar or greater performance as an artificial intelligence model trained using [$>10^{26}$ FLOP] in 2024." If the statute did not contain an explicit authorization for some agency to determine the quantity of compute that qualified in a given year, any attempt to set and enforce updated regulatory compute thresholds could be challenged in court.

The enforcing agency could argue that the statute included an implied authorization for the agency to promulgate and update the regulatory definitions at issue. This argument might succeed or fail, depending on the language of the statute, the nature of the challenged regulatory definitions, and the judicial philosophy of the deciding court. But regardless of the outcome of any individual case, challenges to impliedly authorized regulatory definitions will probably be more likely to succeed after *Loper Bright* than they would have been under *Chevron*. Perhaps more importantly, agencies will be aware that regulatory definitions will no longer receive the benefit of *Chevron* deference and may regulate more cautiously in order to avoid being sued.[116] Moreover, even if the statute

[114] *See Loper Bright* at 35.

[115] *See* Winter & Bullock, *supra* note 8.

[116] Regulatory agencies are generally litigation-averse, and all else being equal tend to avoid taking actions that will increase their likelihood of being sued. *See* Connor Raso, *Agency Avoidance of Rulemaking*

did explicitly authorize an agency to issue updated compute thresholds, such an authorization might not allow the agency to respond to future technological breakthroughs by considering some factor other than the quantity of training compute used.

In other words, a narrow congressional authorization to regulatorily define "frontier model" may prove insufficiently flexible after *Loper Bright*. Congress could attempt to address this possibility by instead enacting a very broad authorization.[117] An overly broad definition, however, may be undesirable for reasons of democratic accountability, as it would give unelected agency officials discretionary control over which models to regulate as "frontier." Moreover, an overly broad definition might risk running afoul of two related constitutional doctrines that limit the ability of Congress to delegate rulemaking authority to agencies—the major questions doctrine and the nondelegation doctrine.

B. The nondelegation doctrine

Under the nondelegation doctrine, which arises from the constitutional principle of separation of powers, Congress may not constitutionally delegate legislative power to executive branch agencies. In its current form, this doctrine has little relevance to efforts to define "frontier model." Under current law, Congress can validly delegate rulemaking authority to an agency as long as the statute in which the delegation occurs includes an "intelligible principle" that provides adequate guidance for the exercise of that authority.[118] In practice, this is an easy standard to satisfy—even vague and general legislative guidance, such as directing agencies to regulate in a way that "will be generally fair and equitable and will effectuate the purposes of the Act," has been held to contain an intelligible principle.[119] The Supreme Court has used the nondelegation doctrine to strike down statutes only twice, in two 1935 decisions invalidating sweeping New Deal laws.[120]

Procedures, 67 ADMIN. L. REV. 65 (2015) (presenting empirical evidence showing that agencies are more likely to avoid procedural rulemaking requirements when the risk of being sued for doing so is low).

[117] *See, e.g.,* 15 U.S.C. § 78l ¶ (g)(5) (authorizing the SEC to regulatorily define the terms "total assets" and "held of record" "as it deems necessary or appropriate"); *see Loper Bright Enterprises v. Raimondo*, 603 U.S. __ at 17 (noting with approval statutes that empower an agency to "regulate subject to the limits imposed by a term or phrase that leaves agencies with flexibility, such as 'appropriate' or 'reasonable.'").

[118] J.W. Hampton, Jr., & Co. v. United States, 276 U.S. 394, 409 (1928); *see* Touby v. United States, 500 U.S. 160, 165 (1991).

[119] *See* Whitman v. Am. Trucking Associations Inc., 531 U.S. 457, 474 (2001) (quoting Yakus v. United States, 321 U.S. 414, 423–426 (1944)).

[120] *See id.* ("In the history of the Court we have found the requisite 'intelligible principle' lacking in only two statutes, one of which provided literally no guidance for the exercise of discretion, and the other of which conferred authority to regulate the entire economy on the basis of no more precise a standard than stimulating the economy by assuring 'fair competition.'") (citing A.L.A. Schechter Poultry Corp. v United States, 295 U.S. 495 (1935) and Panama Ref. Co. v. Ryan, 293 U.S. 388 (1935)).

However, some commentators have suggested that the Supreme Court may revisit the nondelegation doctrine in the near future,[121] perhaps by discarding the "intelligible principle" test in favor of something like the standard suggested by Justice Gorsuch in his 2019 dissent in *Gundy v. United States*.[122] In *Gundy*, Justice Gorsuch suggested that the nondelegation doctrine, properly understood, requires Congress to make "all the relevant policy decisions" and delegate to agencies only the task of "filling up the details" via regulation.[123]

Therefore, if the Supreme Court does significantly strengthen the nondelegation doctrine, it is possible that a statute authorizing an agency to create a regulatory definition of "frontier model" would need to include meaningful guidance as to what the definition should look like. This is most likely to be the case if the regulatory definition in question is a key part of an extremely significant regulatory scheme, because "the degree of agency discretion that is acceptable varies according to the power congressionally conferred."[124] Congress generally "need not provide any direction" to agencies regarding the manner in which it defines specific and relatively unimportant technical terms,[125] but must provide "substantial guidance" for extremely important and complex regulatory tasks that could significantly impact the national economy.[126]

C. The major questions doctrine

Like the nondelegation doctrine, the major questions doctrine is a constitutional limitation on Congress's ability to delegate rulemaking power to agencies. Like the nondelegation doctrine, it addresses concerns about the separation of powers and the increasingly prominent role executive branch agencies have taken on in the creation of important legal rules. Unlike the nondelegation doctrine, however, the major questions doctrine is a recent innovation. The Supreme Court acknowledged it by name for the first time in the 2022 case *West Virginia v. Environmental Protection Agency*,[127] where it was

---

[121] *See* THE ADMINISTRATIVE STATE BEFORE THE SUPREME COURT: PERSPECTIVES ON THE NONDELEGATION DOCTRINE (Peter J. Wallison & John Yoo eds., 2022); Paul J. Larkin, *Revitalizing the Nondelegation Doctrine*, 23 FEDERALIST SOC. REV. 238, 245 (2022); Richard J. Pierce, *Are Most Federal Statutes Unconstitutional?*, THE REGULATORY REVIEW (Aug. 28, 2023), https://www.theregreview.org/2023/08/28/pierce-are-most-federal-statutes-unconstitutional/.

[122] Gundy v. United States, 139 S. Ct. 2116, 2131 (2019).

[123] *Id.* at 2136.

[124] *Whitman* at 457, 475.

[125] *Id.* ("Congress need not provide any direction to the EPA regarding the manner in which it is to define 'country elevators,' which are to be exempt from new-stationary-source regulations governing grain elevators, *see* 42 U.S.C. § 7411(i)").

[126] *Id.*

[127] West Virginia v. EPA, 597 U.S. 697. The Supreme Court had previously relied on the major questions doctrine without identifying it as such in a number of cases. *See, e.g.,* FDA v. Brown & Williamson Tobacco Corp., 529 U.S. 120 (2000); Whitman v. American Trucking Associations, Inc., 531

used to strike down an EPA rule regulating power plant carbon dioxide emissions. Essentially, the major questions doctrine provides that courts will not accept an interpretation of a statute that grants an agency authority over a matter of great "economic or political significance" unless there is a "clear congressional authorization" for the claimed authority.[128] Whereas the nondelegation doctrine provides a way to strike down statutes as unconstitutional, the major questions doctrine only affects the way that statutes are interpreted.

Supporters of the major questions doctrine argue that it helps to rein in excessively broad delegations of legislative power to the administrative state and serves a useful separation-of-powers function. The doctrine's critics, however, have argued that it limits Congress's ability to set up flexible regulatory regimes that allow agencies to respond quickly and decisively to changing circumstances.[129] According to this school of thought, requiring a clear statement authorizing each economically significant agency action inhibits Congress's ability to communicate broad discretion in handling problems that are difficult to foresee in advance.

This difficulty is particularly salient in the context of regulatory regimes for the governance of emerging technologies.[130] Justice Kagan made this point in her dissent from the majority opinion in *West Virginia*, where she argued that the statute at issue was broadly worded because Congress had known that "without regulatory flexibility, changing circumstances and scientific developments would soon render the Clean Air Act obsolete."[131] Because advanced AI systems are likely to have a significant impact on the U.S. economy in the coming years,[132] it is plausible that the task of choosing which systems should be categorized as "frontier" and subject to increased regulatory scrutiny will be an issue of great "economic and political significance." If it is, then the major questions doctrine could be invoked to invalidate agency efforts to promulgate or amend a definition of "frontier model" to address previously unforeseen unsafe capabilities.

For example, consider a hypothetical federal statute instituting a licensing regime for frontier models that includes a definition similar to the placeholder in EO 14110 (empowering the Bureau of Industry and Security to "define, and thereafter update as

---

U.S. 457 (2001); Alabama Assn. of Realtors v. Department of Health and Human Servs., 141 S.Ct. 2485 (2021).

[128] *Id.* at 721; *see* Util. Air Regul. Grp. v. E.P.A., 573 U.S. 302, 324 (2014) ("We expect Congress to speak clearly if it wishes to assign to an agency decisions of vast economic and political significance.")

[129] *See, e.g.,* Jonas J. Monast, *Emerging Technology Governance in the Shadow of the Major Questions Doctrine*, 24 N.C. J.L. & TECH. 1 (2023); Daniel T. Deacon & Leah M. Litman, *The New Major Questions Doctrine*, 109 VA. L. REV. 1009 (2023).

[130] *See id.*

[131] *West Virginia* at 2632.

[132] *See, e.g.,* Annual Global Corporate Investment in Artificial Intelligence, by Type, OUR WORLD IN DATA (2023), https://ourworldindata.org/grapher/corporate-investment-in-artificial-intelligence-by-type; Generative AI to Become a $1.3 Trillion Market by 2032, BLOOMBERG, June 1, 2023, https://www.bloomberg.com/company/press/generative-ai-to-become-a-1-3-trillion-market-by-2032-research-finds/.

needed on a regular basis, the set of technical conditions [that determine whether a model is a frontier model]."). Suppose that BIS initially defined "dual-use foundation model" under this statute using a regularly updated compute threshold, but that ten years after the statute's enactment a new kind of AI system was developed that could be trained to exhibit cutting-edge capabilities using a relatively small quantity of training compute. If BIS attempted to amend its regulatory definition of "frontier model" to include a capabilities threshold that would cover this newly developed and economically significant category of AI system, that new regulatory definition might be challenged under the major questions doctrine. In that situation, a court with deregulatory inclinations might not view the broad congressional authorization for BIS to define "frontier model" as a sufficiently clear statement of congressional intent to allow BIS to later institute a new and expanded licensing regime based on less objective technical criteria.[133]

VI. Conclusion

One of the most common mistakes that nonlawyers make when reading a statute or regulation is to assume that each word of the text carries its ordinary English meaning. This error occurs because legal rules, unlike most writing encountered in everyday life, are often written in a sort of simple code where a number of the terms in a given sentence are actually stand-ins for much longer phrases catalogued elsewhere in a "definitions" section.

This tendency to overlook the role that definitions play in legal rules has an analogue in a widespread tendency to overlook the importance of well-crafted definitions to a regulatory scheme. The object of this paper, therefore, has been to explain some of the key legal considerations relevant to the task of defining "frontier model" or any of the analogous phrases used in existing laws and regulations.

---

[133] This hypothetical is loosely based on *West Virginia v. EPA*, 597 U.S. 697 (2022). That case involved a challenge to an EPA regulation that relied on a provision of the Clean Air Act of 1970 that contained a broad authorization for the EPA to "determine the best system of emission reduction which… has been adequately demonstrated" for an identified pollutant and set a limit for emissions of the identified pollutant based on the "best system of emission reduction" identified. 42 U.S.C. § 7411 ¶ (a)(1). Between 1970 and 2015, the EPA exercised this authority rarely, and the "best systems of emission reduction" it determined typically involved specific emission-limiting technologies or processes such as specialized filtration systems. In 2015, the EPA identified carbon dioxide as an air pollutant under the Clean Air Act and identified a "best system of emission reduction" that involved requiring coal-burning power plants to shift some of their energy production to wind, solar, and natural gas plants. *See* Standards of Performance for Greenhouse Gas Emissions From New, Modified, and Reconstructed Stationary Sources: Electric Utility Generating Units, 80 Fed. Reg. 64510-0180 (Oct. 23, 2015). The Supreme Court struck down the EPA's rule under the major questions doctrine, finding that the broad language of the authorization in the Clean Air Act did not constitute a "clear congressional authorization" for the kind of sweeping and economically significant policy change that the EPA rule would have instituted. *See West Virginia* at 722–724. Justice Kagan wrote a dissenting opinion noting that the new doctrine would make it more difficult for Congress to delegate broad authority to agencies in order to ensure agency access to "the tools needed to confront emerging dangers" in rapidly developing "scientific and technical areas." *Id.* at 781–782.

One such consideration is the role that should be played by statutory and regulatory definitions, which can be used independently or in conjunction with each other to create a definition that is both technically sound and democratically legitimate. Another is the selection and combination of potential definitional elements, including technical inputs, capabilities metrics, risk, deployment context, and familiarity, that can be used independently or in conjunction with each other to create a single statutory or regulatory definition. Legal mechanisms for facilitating rapid and frequent updating for regulations targeting emerging technologies also merit attention. Finally, the nondelegation and major questions doctrines and the recent elimination of *Chevron* deference may affect the scope of discretion that can be conferred for the creation and updating of regulatory definitions.