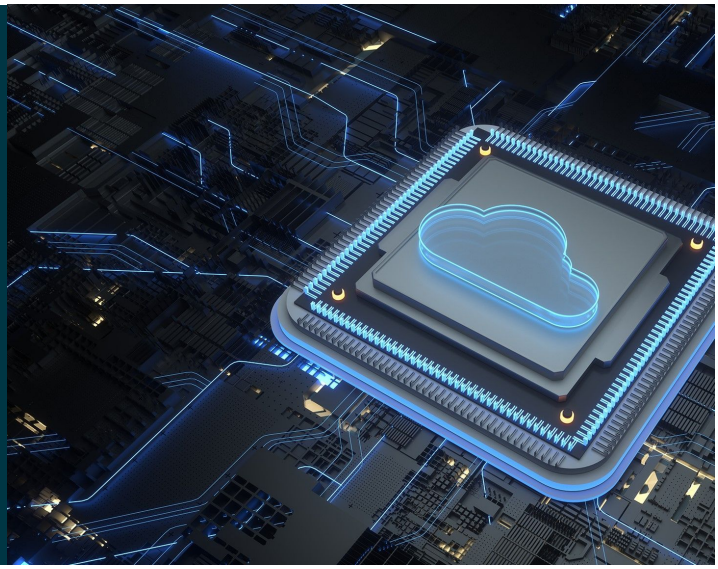
WILSON SONSINI

Commerce Proposes New Reporting Requirement for Advanced AI Developers and Cloud Computing Providers



CONTRIBUTORS


Josephine I. Aiello LeBeau


Joshua F. Gruenspecht


Anne E. Seymour


Kara D. Millard

ALERTS

September 16, 2024

Last week, the U.S. Department of Commerce's (Commerce) Bureau of Industry and Security (BIS) released a Notice of Proposed Rulemaking (NPRM) outlining a new mandatory reporting requirement for large-scale AI developers and cloud computing providers that provide compute to AI model developers. The NPRM stems from requirements issued under the Biden Administration's October 2023 Executive Order on "Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence" (Biden's EO, discussed here), which directed Commerce to collect certain information on dual-use foundation models and large-scale computing clusters. Commerce has already collected initial responses to a similar mandatory survey issued pursuant to Biden's EO from major AI developers and compute providers; this new NPRM will systematize that data collection process.

The proposed reporting requirements are expected to apply to companies with "dual-use foundation models" as defined by Biden's EO. A dual-use foundation model is defined as an AI model that is "trained on broad data; generally uses self-supervision; contains tens of billions of parameters; is applicable across various contexts; and that displays, or can be modified to display, high performance at tasks posing serious risks to security, national economic security, national public health or safety, or any combination of these matters…"

While the proposed reporting requirements would assess the capabilities of potential dual-use foundation models, the reporting requirements would be tied to specific technical criteria rather than attempting to identify those models *ex ante*. Under the NPRM, reporting would be required by "covered U.S. persons," meaning U.S. companies, individuals, or other organizations or entities, that engage or plan within six months to engage in "applicable activities." While the criteria are subject to future updates, the NPRM proposes that initially such activities would include:

- conducting any AI model training run using more than 1026 computational operations (e.g., integer or floating-point operations).[1]
- acquiring, developing, or coming into possession of a computing cluster that has a set of machines transitively connected by data center networking of over 300 Gbit/s and having a theoretical maximum performance greater than 1020 computational operations (e.g., integer or floating-point operations) per second (OP/s) for AI training, without sparsity.

Commerce estimates that at the time of publication, between zero and 15 companies have dual-use foundation models or computing clusters that may fit these criteria, all of which are well-resourced technology companies with powerful foundational AI models or training capabilities. With respect to the initial notification, a covered U.S. person (i.e., a U.S. company) that becomes subject to the reporting requirements must notify BIS of its engagement in the "applicable activities" via email at ai_reporting@bis.doc.gov. The company will then receive an initial questionnaire from BIS and must respond within 30 calendar days.

Based on the NPRM and the background information provided by BIS, as well as the surveys BIS already issued to select companies in January 2024 under Biden's EO, initial questions from BIS will likely address, but are not limited to, the following topics:

- Company Background
  - Current development or plans to develop dual-use foundation models
  - Current ownership or plans to own or acquire computing clusters
  - Computing hardware capacity for model development
- Security Issues
  - Company cybersecurity resources and practices
  - Physical and cybersecurity protections
  - Controls on ownership and possession of model weights
- Performance and Safety Issues
  - Results of red-team testing related to flaws and vulnerabilities in company models
  - Results of red-team testing related to lowering the barrier to entry for the development, acquisition, and use of chemical, biological, radiological, or nuclear weapons by non-state actors
  - Protective measures for dual-use foundation models
  - Safety measures implemented to address identified issues
  - Potentially dangerous capabilities identified by developers, including:
    - Discovery of software vulnerabilities and associated exploits
    - Potential to influence real or virtual events
    - Possibility of self-replication or propagation

Under the preliminary NPRM rules, once initial responses are provided, companies must file quarterly reports as long as they continue to engage in applicable activities, describing any changes or new covered activities. Companies that completed the January 2024 mandatory survey similarly will be responsible for quarterly updates. Even if a company ceases to engage in new applicable activities, it must continue filing for seven quarters, affirming no new covered business since its most recently filed change report.

Finally, the NPRM suggests potential for government funding based on the data gathered through these reporting requirements. Such funding could potentially stimulate the development of dual-use foundation models or support the creation of specific types of models.

Commerce Secretary Gina Raimondo remarked in her commentary on the NPRM that AI technology is "progressing rapidly," posing both "tremendous promise and risk."[2] By requiring ongoing reporting from leading AI developers and cloud providers, the new proposed rules are intended to inform and instruct the U.S. government's policy decisions about the AI industry. According to Commerce, this data will be used to facilitate the continual evaluation of the safety and defense capabilities of AI technology.[3] The NPRM also highlights the importance of integrating dual-use foundation models into the U.S. defense industrial base in order to maintain global competitiveness.[4]

BIS has solicited public comments on the proposed rules in the NPRM by October 11, 2024, expressing particular interest in feedback regarding: i) the quarterly notification schedule; ii) methods for collecting and storing sensitive data received as a result of the rule, and iii) the two technical parameters that trigger reporting requirements.

For more information or advice concerning your compliance efforts related to AI, please contact Josephine Aiello LeBeau, Joshua Gruenspecht, Anne Seymour, Kara Millard, or any member of the firm's national security practice or artificial intelligence and machine learning working group.

---

[1] Note: Models trained on primarily biological sequence data, but at a lower threshold of 1023 computational operations, will be addressed in a separate survey.

[2] BIS, "Commerce Proposes Reporting Requirements for Frontier AI Developers and Compute Providers," https://www.bis.gov/press-release/commerce-proposes-reporting-requirements-frontier-ai-developers-and-compute-providers (September 9, 2024).

[3] 89 FR 72326, 73614 (filed September 9, 2024, published September 11, 2024), https://www.federalregister.gov/d/2024-20529/page-73614.

[4] *Id.* at 73615.