CLIENT ALERTS & INSIGHTS

California Senator Scott Wiener Introduced New Bill that Would Require a Kill Switch for Applicable Artificial Intelligence Models

FEBRUARY 15, 2024

Authors: Daniel S. Marks

On February 7, 2024, California State Senator Scott Wiener introduced SB 1047, which is titled the "Safe and Secure Innovation for Frontier Artificial Intelligence Systems Act." The bill is designed to impose additional due diligence, testing, and determination findings for certain covered artificial intelligence (A.I.) models prior to their commercial, public, or widespread use.

While the definition of what constitutes an A.I. model under the bill is broad (i.e., more than generative A.I. models), the bill primarily focuses on only large A.I. models that represent the current state-of-the art (i.e. models that were trained using a quantity of computing power greater than 10^26 integer or floating-point operations in 2024).

The scope and threshold for "covered models" under SB 1047 is similar to criteria put forth in President Biden's Executive Order from October 2023. You can read more about our analysis of that Executive Order here.

SB 1047 would impose requirements on developers of covered A.I. models, which include implementation of:

- administrative, technical, and physical cybersecurity protections to prevent unauthorized access or use of the covered model;
- the capability to promptly enact a full shutdown of the covered model;
- all guidance issued by NIST and industry best practices;
- a written and separate safety and security protocol; and
- annual reviews of the above to account for changes in capabilities of the covered model and industry best practices.

In addition, SB 1047 would require the developer to make a positive safety determination, which means that the developer has done sufficient testing to reasonably exclude the possibility that a covered model has a "hazardous capability." Hazardous capabilities are defined as either types of harm (e.g., chemical, biological, radiological, or nuclear) or amounts of harm (e.g., $500,000,000 of damage through cyberattacks on critical infrastructure).

Lastly, SB 1047 would create the Frontier Model Division, under the Department of Technology, that would review annual certification reports from developers received pursuant to this bill. The Attorney General would have the exclusive authority to enforce violations. Such enforcement could include civil penalties of up to 10% of the cost to develop the covered model (excluding labor) for initial violations and up to 30% of the cost to develop the covered model (excluding labor) for subsequent violations.

SB 1047 was referred to the Committee on Judiciary and Committee on Government Organization, which isn't expected to act on it until after March 9, 2024.

For more information on this subject, please contact a member of Benesch's AI Commission or Intellectual Property Practice Group.

Daniel Marks at dmarks@beneschlaw.com or 312.517.9546.