Comparison of California's SB 53 and New York's RAISE Act: Frontier Model Frameworks

Created by: Justine Gluck, AI Policy Analyst

Overview: On December 19, Governor Hochul (D-NY) signed the "**Responsible AI Safety and Education (RAISE) Act,"** incorporating chapter amendments that revise the bill text adopted by the Legislature. Introduced by Asm. Bores (D), the bill is the second major legislative effort to address the safety and oversight of AI frontier models in the U.S., following **SB 53**, or the "Transparency in Frontier Artificial Intelligence Act (TFAIA)," signed by Governor Newsom on September 29. While Hochul's chapter amendments align the RAISE Act more closely to TFAIA, some areas of divergence may affect entities seeking to comply with the laws:

1. **Scope:** Both bills contain similar scopes of regulated technologies and entities, however the RAISE Act exempts universities and explicitly applies only to models developed or operated in whole or in part in New York (SB 53 does not contain this limitation and was explicitly framed to go beyond state borders).
2. **Requirements:** SB 53 includes employee whistleblower protections, while the RAISE Act introduces a frontier developer disclosure registration program requiring corporate information.
3. **Safety Incident Reporting:** SB 53 allows a longer reporting window (15 days), whereas RAISE requires disclosure within 72 hours based on a "reasonable belief" standard.
4. **Rulemaking Authority:** SB 53 relies on legislative recommendations to update definitions, while the RAISE Act grants regulators direct rulemaking authority to shape implementation.
5. **Enforcement:** SB 53 caps penalties at $1 million per violation; RAISE authorizes higher penalties ($3 million) for repeat violations.

This comparative analysis considers these similarities and differences between California's and New York's frameworks, covering (1) scope; (2) compliance requirements; (3) other requirements; and (4) enforcement.

Red text indicates key differences between SB 53 and the RAISE Act.

	California SB 53 (TFAIA)	New York A 9449 (RAISE Act)	Comparison
	Scope		
Scope	**Foundation Model** means an AI model that is: 1. Trained on a broad data set; 2. Designed for generality of output; and 3. Adaptable to a wide range of distinctive tasks (Sec. 22757.11 (f)). **Frontier Model** means a foundation model that was trained using a quantity of computing power greater than 10^26 computational operations (e.g., integer or floating-point operations) (Sec. 22757.11 (i)). - This quantity of computing power includes the original training run and subsequent modifications to the foundation model. **Frontier Developer:** A person who has trained or initiated the training of a frontier model which the person intends to use at	***Identical, except that the RAISE Act adds the following qualifiers:*** Frontier models must also be developed, deployed, or operated in whole or in part in New York State (Sec. 1425). Accredited colleges and universities are not included within scope to the extent they engage in academic research (Sec. 1426(1)).	***RAISE and SB 53 have identical scopes, however, the RAISE Act offers a few additional qualifiers.*** **Frontier Model:** Both bills use the same compute benchmark, targeting the most advanced and resource-intensive systems. The definition also includes cumulative compute used not only in initial training but *also* in any fine-tuning or modifications. **Cost Threshold:** Both laws use a $500M+ annual revenue threshold to define "large" frontier developers, narrowing the scope further. Neither bill is scoped to apply to small developers.

	least as 10^26 computing power (Sec. 22757.11 (h)). **Large Frontier Developer:** A frontier developer that together with its affiliates collectively had gross revenues in excess of $500 million (Sec. 22757.11 (j)).		**Developer Categories:** The laws distinguish between "frontier developers" and "large frontier developers," applying different requirements to each group. The cost threshold applies only to the latter, meaning some obligations extend to developers regardless of cost. **Extraterritorial Application:** SB 53 doesn't explicitly restrict scope to California-based developers and was framed by Newsom as a blueprint beyond state borders. However, RAISE applies only to models "developed, deployed, or operated in whole or in part in New York." **Academia Exemption:** RAISE explicitly exempts universities; SB 53 provides no similar exemption.
Key Terms	**Catastrophic Risk:** Foreseeable and material risk that a frontier developer's development, storage, use, or deployment of a frontier model will materially contribute to the death or serious injury of 50+ or at least $1 billion of damages to loss of property (tangible and intangible) arising from a single incident involving a frontier model doing any of the following: a. Provide expert-level assistance in the creation or release of a chemical, biological, radiological, or nuclear weapon; b. Engage in conduct with no meaningful human intervention that would, if committed by a human, constitute a crime or is a cyberattack; or c. Evade the control of its frontier developer or user (Sec. 22757.11 (c)(1)). The loss of value of equity does not count as damage to or loss of property (Sec. 1107.2). Catastrophic risk does not include a foreseeable and material risk from any of the following: a. Information that a frontier model outputs if it is publicly accessible in a substantially similar form from another source; b. Lawful activity of the federal government; and c. Harm caused by a frontier model in combination with	***Identical.***	***The RAISE Act and SB 53 carry identical definitions for catastrophic risk.*** **Scope of Risk and Harm:** Both bills set high thresholds for risk, (e.g. 50+ deaths) and define "catastrophic risk" to also include additional model behaviors, such as evading developer control. **Dangerous Capabilities:** Both bills include any model providing "expert-level assistance" in creating or using a weapon, a relatively high standard. **Liability Limitations:** Both laws use a broad standard, eliminating previous language in the RAISE Act that would have required that harm be a "probable consequence" of the developer's activities, and that the developer's actions be a "substantial factor."

	other software if the frontier model did not materially contribute (Sec. 22757.11 (c)(2)). **Deploy:** to make a frontier model available to third-party for use, modification, and copying (except developing/evaluating frontier model) (Sec. 22757.11 (e)).		
Compliance Requirements			
Frontier AI Framework	**Content:** **Frontier AI framework** means documented technical and organizational protocols to manage, assess, and mitigate catastrophic risks (Sec. 22757.11 (g)). **A large frontier developer shall implement, comply with, and clearly and conspicuously publish a frontier AI framework** that describes how the large frontier developer approaches the following: 1. Incorporating national standards, international standards, and industry-consensus best practices into its frontier AI framework; 2. Defining and assessing thresholds used to assess whether frontier model has capabilities that could pose a catastrophic risk; 3. Applying mitigations to address potential for catastrophic risks; 4. Reviewing assessments and adequacy of mitigations as part of decision to deploy the frontier model; 5. Using third parties to assess the potential for catastrophic risks and effectiveness of mitigations; 6. Updating frontier AI framework, including criteria triggering updates and how the developer determines when its frontier models are substantially modified enough to require disclosures; 7. Cybersecurity practices and how they secure unreleased model weights; 8. Identifying and responding to critical safety incidents; 9. Instituting internal governance practices for implementation of these processes; and 10. Assessing and managing catastrophic risk from internal use (Sec. 22757.12 (a)).	***Largely Identical.*** **A large frontier developer shall implement, comply with, and clearly and conspicuously publish a frontier AI framework** that describes how the large frontier developer handles the following (Sec. 1421 (1)) …	***Both laws mandate identical written frontier AI frameworks and public disclosure.*** **Drafting Differences:** The laws use slightly different verbiage, SB 53 requires developers to describe how they "approach" required risk-management processes, while the RAISE Act asks how they "handle" them. In practice, the distinction may be non-substantive, as both provisions are followed by identical, detailed requirements; if anything, "handles" may modestly signal a greater expectation of operational implementation rather than high-level process description. **Protocol Content:** Both laws require detailed documentation of governance structures, mitigation processes, and alignment with national/international standards. They also explicitly cover catastrophic risk from *internal use* of models, raising the scope of compliance obligations. **Testing Requirements:** Neither bill explicitly mandates that specific tests be performed, leaving open questions about what level of testing is actually necessary for compliance, as there is a reference to reviewing "assessments." **Timing & Updates:** Both require annual reviews, including the framework to be re-published within 30 days of modifications, which may force

	Administration: Large frontier developer shall review and, if appropriate, update its frontier AI framework at least annually (Sec. 22757.12 (b)(1)). If a large frontier developer makes a material modification to its frontier AI framework, they must clearly and conspicuously publish the framework and justification within 30 days (Sec. 22757.12 (b)(2)). **Redactions:** Frontier developers may make redactions to the framework *(and transparency report)* to protect trade secrets, cybersecurity, national security. To the extent permitted, must publicly describe redactions and retain unredacted information for 5 years (Sec. 22757.12 (f)).		companies to account for quick revision cycles.
Transparency Report	Before, or concurrently with, deploying a frontier model a frontier developer shall clearly and conspicuously publish on its website a transparency report containing all of the following: a. Website of the frontier developer; b. Mechanism that allows a natural person to communicate with the frontier developer; c. Release date of the frontier model; d. Languages supported by the frontier model; e. Modalities of output supported by frontier model; f. Intended uses of frontier model; g. Restrictions or conditions on uses of the frontier model (Sec. 22757.12 (c)(1)). Before, or concurrently with, deploying a frontier model, a frontier developer shall include in the transparency report summaries of all of the following: a. Assessments of catastrophic risks conducted pursuant the frontier AI framework; b. Assessment results; c. Involvement of third-party evaluators; d. Other steps to fulfill requirements of the frontier AI framework (Sec. 22757.12 (c)(2)). Frontier developers can publish this information as part of a larger document, like a system or model card (Sec. 22757.12	***Identical, except the RAISE Act does not include the provision in SB 53 that encourages (but does not require) alignment with industry best practices.***	***Both laws require frontier developers to publish detailed transparency reports.*** **Pre-Deployment Transparency:** The laws mandate that frontier developers publish a transparency report before or concurrently with deployment. **Scope:** Unlike many other obligations that apply only to "large frontier developers," the transparency report requirement applies to all frontier developers, meaning even smaller firms that meet the "frontier developer" definition face compliance requirements. **Internal Use:** The transparency report requires publication of a summary of any catastrophic risk assessment stemming from **internal use** of a foundation model, broadening the scope of required transparency. **Integration with Practices:** Firms can incorporate disclosures into existing documents like system or model cards, which may help ease compliance. **Alignment with Best Practices:** SB 53 also encourages alignment with industry best practices,

	(c)(3)). Frontier developers encouraged, but not required, to make disclosures that are consistent or superior to industry best practices (Sec. 22757.12 (c)(4)). Large frontier developers shall transmit to the Office of Emergency Services (OES) a summary of any assessment of catastrophic risk resulting from internal use of its frontier models every three months (Sec. 22757.12 (d)).		though the lack of clear benchmarks may create compliance uncertainty. The RAISE Act omits this provision.
Disclosure of Safety Incidents	The OES will establish a mechanism for the frontier developer or member of the public to report a critical safety incident that includes: 1. The date of the safety incident; 2. Reasons the incident qualifies as a safety incident; 3. A short and plain statement describing the safety incident; and 4. Whether the incident was associated with internal model use (Sec. 22757.13 (a)). A frontier developer shall report any critical safety incident within 15 days of discovery (Sec. 22757.13 (c)(1)). If a frontier developer discovers a critical safety incident poses an imminent risk of danger of death or serious injury, they shall disclose that incident no later than 24 hours to an authority (Sec. 22757.13 (c)(2)). A frontier developer is encouraged, but not required, to report critical safety incidents pertaining to foundation models that are not frontier models (Sec. 22757.13 (c)(4)). **Critical Safety Incident** means any of the following: 1. Unauthorized access to, modification of the model weights of a foundation model that results in death or bodily injury; 2. Harm resulting from the materialization of a catastrophic risk; 3. Loss of control of a frontier model causing death or bodily injury, or loss of property; or 4. A foundation model that employs deceptive	***Identical requirements to establish mechanisms to report critical safety incidents (and identical definitions), but RAISE requires the Department of Financial Services.*** A frontier developer shall report any critical safety incident pertaining to one or more of its frontier models to the Department of Financial Services (DFS) within 72 hours from a determination that a critical safety incident has occurred OR within 72 hours of the developer learning facts sufficient to establish a reasonable belief that a safety incident has occurred (Sec. 1422(3)(a)). If a frontier developer discovers a critical safety incident poses an imminent risk of danger of death or serious injury, they shall disclose that incident no later than 24 hours to an authority (Sec. 1422(3)(b)). DFS may transmit reports of critical safety incidents or summaries of any assessments of catastrophic risk from internal use of frontier models to other governmental entities at their discretion, considering for example: incident severity, potential ongoing risks, legal or regulatory obligations, the need for coordinating with other entities, and the availability of information. The	***SB 53 offers developers more time to disclose non-imminent risks, while RAISE imposes a shorter 72-hour window and uses specific legal qualifiers for a reportable incident.*** **Reporting Timeline:** RAISE requires disclosure within 72 hours or "reasonable belief," while SB 53 allows 15 days, unless there is an imminent risk of "danger of death or serious physical injury" and the timeline shortens to 24 hours for both laws. **Threshold:** RAISE uses a "reasonable belief" standard and requires "demonstrable evidence" of increased risk, raising the bar for what qualifies as incident reporting and requiring action even in the absence of confirmed harm. **Incident Scope:** Both include comparable incidents involving unauthorized access, misuse, or loss of control and define "critical safety incident" identically. **Public Reporting:** Both laws require the establishment of a mechanism for the public to report safety incidents, expanding oversight beyond developers. While this could enhance transparency, it may also raise concerns for developers about unverified public claims. **Information Sharing:** Both laws permit interagency

	techniques to evade the controls or monitoring of its frontier developer in a manner that demonstrates materially increased risk (Sec. 22757.11(d)). The AG/ OES may transmit reports of safety incidents to the Legislature, Gov., federal government, or agencies. Risks related to trade secrets, public safety, cybersecurity, or national security shall be strongly considered when transmitting reports (Sec. 22757.13 (e)(1)). The OES shall produce an anonymized annual report with information on critical safety incidents and transmit the report to the Legislature and Governor (Sec. 22757.13 (g)).	office shall consider transmitting such reports to the AG as appropriate Sec. 1422(5(a))). Risks related to trade secrets, public safety, cybersecurity, or national security shall be considered by DSF, at its discretion, when transmitting reports (Sec. 1422(5(b))).	sharing of safety incident reports. SB 53 directs regulators to *"strongly consider"* risks related to trade secrets, public safety, cybersecurity, and national security, while the RAISE Act affords broader agency "*discretion*" and specifies illustrative factors, such as incident severity and coordination needs. This added detail may provide developers clearer expectations about how and when incident information could circulate, even as ultimate decisions remain discretionary. **Foundation Models:** SB 53 encourages (but does not require) reporting of safety incidents involving foundation models that fall below the frontier threshold, which may incentivize voluntary disclosure. The RAISE Act omits this provision.
Other Requirements			
Frontier Developer Disclosure	**N/A**	No large frontier developer may develop, deploy, or operate a frontier model without a current disclosure statement filed with DFS and paying the pro rata share (Sec. 1428(1)). The disclosure statement shall be renewed every two years, whenever ownership is transferred, or whenever there is a material change to the information reported in the previous disclosure (Sec. 1428(2)). Disclosure statement includes: (a) the identity of the large frontier developer and all names under which it conducts business; (b) the address of the principal place of business and New York offices; (c) list all persons holding at least a 5% interest in a privately held developer (or its parent) over the past five years and any person holding a 50% or greater interest in a publicly held developer, at the time of registration; and (d) the name and contact information for three points of contact ((Sec. 1428(3)).	***The RAISE Act establishes a registration-style disclosure program not present in SB 53.*** The RAISE Act requires large frontier developers to maintain current filings with DFS covering ownership, business information, and points of contact. Disclosures must be updated on a recurring basis, introducing an ongoing administrative obligation distinct from those required under SB 53. Daily penalties reinforce incentives for compliance.

		Large frontier developers shall be assessed in pro rata shares by the department to defray the operating expenses, including all direct and indirect costs, of administering the program (Sec. 1428(4)). DFS may level civil penalties and fees if a person develops, deploys, or operates a large frontier model without a current disclosure filed, or submits false information in its disclosure or fails to pay: (a) a civil penalty of $1,000 per day for failing to file a disclosure or false information; and (b) an amount equal to assessments owed (Sec. 1428(5)). 6. The office shall publish a list of large frontier developers who have filed disclosure statements, but without point of contact information (Sec. 1428(6)).	
Whistleblower Protections	A frontier developer shall not adopt a policy or contract that retaliates against or prevents a covered employee from disclosing information to the AG, or other authority, if the covered employee has reasonable cause to believe developer's activities pose a specific and substantial danger to the public health or safety resulting from a catastrophic risk or have violated the Act. **Covered Employee:** an employee responsible for assessing, managing, or addressing risk of critical safety incidents (Sec. 1107(b)). Frontier developer shall provide a clear notice to all covered employees of their rights/ responsibilities, by: 1. Displaying at all times within the workplace a notice to all covered employees of their rights; or 2. Annually, providing written notice to covered employees of their rights (Sec. 1107.1 (e)). The frontier developer shall provide a reasonable internal process for an employee to anonymously disclose	**N/A**	***SB 53 establishes whistleblower protections; the RAISE Act does not address whistleblowers.*** SB 53 prohibits retaliation against employees or contractors who report activity from a catastrophic risk, mandates notice of employee rights, and requires anonymous internal reporting channels.

	information and provide monthly update to the discloser. The large developer has the burden of proof. Courts must consider direct harm and potential chilling effect on other employees (Sec. 1107.1 (h)).		
	Enforcement		
Enforcement	The Attorney General (AG) may bring a civil action against a large frontier developer that fails to publish a document, report an incident, or comply with its own frontier AI framework. Civil penalty up to $1 million per violation, dependent on the severity of the violation (Sec. 22757.15(a)(b)). **Definitional Recommendations:** Before January 1, 2027, the Department of Technology may make recommendations about updating the definitions of "frontier model," "frontier developer," and "large frontier developer" to ensure it reflects technological developments, submitting a report to the Legislature (Sec. 22757.15 (a)). The AG shall produce a report about reports from employees responsible for addressing critical safety incidents and submit the report to the Legislature (Sec. 22757.14 (d)).	The AG may bring a civil action for a violation, determined based on the severity of the violation: a. A civil penalty in an amount not exceeding $1 million for a first violation and $3 million for any subsequent violation; b. Injunctive or declaratory relief (Sec. 1427(1)). No private right of action (Sec. 1427(2)). Nothing shall be construed to prevent a large frontier developer from asserting that another person, entity, or factor may be responsible for any alleged harm, injury, or damage resulting from a catastrophic risk or critical safety incident (Sec. 1427(3)). **Rulemaking Authority:** DFS is authorized to adopt rules and regulations to implement the provisions as needed. To the extent that doing so will facilitate safety and transparency consistent with the underlying purpose of the law, DFS may consider additional reporting or publication requirements, like post-critical safety incident information (Sec. 1429).	***SB 53 sets slightly lower penalties compared to the RAISE Act, offers definitional adaptability, and AG reporting requirements. However, the RAISE Act offers rulemaking authority.*** **Enforcement Mechanism:** Both bills authorize the AG to bring civil actions for violations. Neither bill includes a private right of action, though RAISE explicitly prohibits one, unlike SB 53. **Penalties:** RAISE sets significantly higher penalties, up to $1 million for a first violation and $3 million for subsequent ones, based on severity. SB 53 also considers severity, but with lower penalties capped at $1 million per violation, signaling a more modest enforcement. Neither bill offers any affirmative defense or safe harbor, although RAISE expressly clarifies that developers may assert that alleged harm was caused by another person, entity, or factor. **Definitional Adjustment:** SB 53 uniquely empowers the Department of Technology to recommend updates to statutory definitions to keep pace with technological change. While these recommendations require legislative adoption, this mechanism builds in definitional adaptability absent in RAISE. Earlier drafts of SB 53 granted the AG direct rulemaking authority to revise definitions, but was narrowed in the final bill. **Rulemaking Authority:** The RAISE Act offers direct rulemaking authority to DFS, such as considering additional reporting or publication requirements. It

			is possible that this difference may affect how flexibly each law evolves over time, with New York relying more on regulatory implementation and California on legislative updates.

Key Differences Between SB 53 and RAISE vs. California's SB 1047 (2024, vetoed)

1. **Pre-Training Requirements:** SB 1047 would have required developers to implement safety protocols, cybersecurity protections, and full shutdown capabilities before beginning initial training of a covered model.
    a. *Neither RAISE nor SB 53 imposes pre-training obligations; both focus on deployment-stage requirements.*

2. **Full Shutdown:** SB 1047 would have mandated that covered models include a full shutdown capability as a safety mechanism, developed pre-training.
    a. This requirement is not present in either RAISE or SB 53.

3. **Prohibition on Deployment:** SB 1047 would have prohibited developers from deploying a frontier model if doing so would create an unreasonable risk of critical harm.
    a. Neither RAISE nor SB 53 includes a strict prohibition on deployment, although the RAISE Act initially contained a similar prohibition that was later removed through the Governor's chapter amendments.

4. **Third-Party Audits:** SB 1047 would have required developers to retain an independent third-party auditor annually to assess internal controls and compliance.
    a. *RAISE and SB 53 do not contain any third-party audit requirements.*

5. **72-Hour Safety Incident Reporting:** SB 1047 would have required reporting a safety incident to the Attorney General within 72 hours of forming a "reasonable belief" that it occurred.
    a. RAISE mirrors this standard; SB 53 provides a longer 15-day window and lacks the "reasonable belief" trigger, but limits to 24 hours for imminent risks.

6. **Civil Penalties:** SB 1047 would have set penalties up to 10% of compute cost used to train the model (30% for subsequent violations), scaled to the harm's severity.
    a. SB 53 caps penalties at $1 million per violation, while RAISE caps penalties at $1 million for first and $2 million for subsequent violations, although both still use severity of harm as a metric.