SENATE COMMITTEE ON APPROPRIATIONS

Senator Anna Caballero, Chair
2025 - 2026 Regular Session

SB 813 (McNerney) - Multistakeholder regulatory organizations

Version: May 1, 2025	**Policy Vote:** JUD. 10 - 0
Urgency: No	**Mandate:** No
Hearing Date: May 12, 2025	**Consultant:** Liah Burnley

Bill Summary: SB 813 establishes a rebuttable presumption that a developer exercised reasonable care in civil actions for harms caused by artificial intelligence they are certified by a "multistakeholder regulatory organization" (MRO).

Fiscal Impact:

- **Department of Justice (DOJ)**: Unknown, potentially significant workload cost pressures (General Fund) to designate MROs as required by this bill.

- **Trail Courts:** Unknown, potentially cost pressures to the state funded trial court system (Trial Court Trust Fund, General Fund) to adjudicate civil actions affected by this bill. By creating a rebuttable presumption if certain requirements are met, this bill may encourage litigants to bring their claims that otherwise would not have, and could lead to more complex court proceedings with attendant workload and resource costs to the court. The fiscal impact of this bill to the courts will depend on many unknown factors, including the number of cases filed and the factors unique to each case. An eight-hour court day costs approximately $10,500 in staff in workload. If court days exceed 10, costs to the trial courts could reach hundreds of thousands of dollars. In 2023–24, over 4.8 million cases were filed statewide in the superior courts. Filings increased over the past year, driven mostly by misdemeanors and infractions, and civil limited cases. The increase in filings from the previous year is greater than 5% for civil limited and unlimited, appellate division appeals, juvenile delinquency, misdemeanors and infractions, and probate. While the courts are not funded on a workload basis, an increase in workload could result in delayed court services and would put pressure on the General Fund to fund additional staff and resources and to increase the amount appropriated to backfill for trial court operations. The Governor's 2025-26 budget proposes a $40 million ongoing increase in discretionary funding from the General Fund to help pay for increased trial court operation costs beginning in 2025-26.

Background: Artificial intelligence has introduced new safety and security risks. Automated systems can make critical errors in high-stakes situations like self-driving vehicles, medical diagnostics, or home security systems when they encounter edge cases or adversarial inputs. This bill creates a shield against liability caused by AI models and applications that are certified at the time of the injuries by a private entity designated by the AG, called an MRO. This bill also creates requirements for the establishment of MROs.

Proposed Law:

- "Multistakeholder regulatory organization (MRO)" means an entity designated as an MRO by the Attorney General that certifies developers' exercise of heightened care and compliance with standards based on best practices for the prevention of personal injury and property damage.

- Requires the Attorney General to designate one or more MROs.

- Requires the Attorney General to determine whether an applicant MRO's plan ensures acceptable mitigation of risk from any MRO-certified artificial intelligence models and artificial intelligence applications by considering all of the following:

    - The applicant's personnel and the qualifications of those personnel;

    - The quality of the applicant's plan with respect to ensuring that artificial intelligence model and application developers exercise heightened care and comply with best practice-based standards for the prevention of personal injury and property damage, considering factors including, but not limited to, both of the following:

        - The viability and rigor of the applicant's evaluation methods, technologies, and administrative procedures; and,

        - The adequacy of the applicant's plan to develop measurable standards for evaluating artificial intelligence developers' mitigation of risks;

    - The applicant's independence from the artificial intelligence industry; and,

    - Whether the applicant serves a particular existing or potential artificial intelligence industry segment.

- A designation as an MRO expires after three years, and the MRO may apply for a new designation.

- The Attorney General may revoke a designation if any of the following is true:

    - The MRO's plan is materially misleading or inaccurate;

    - The MRO systematically fails to adhere to its plan;

    - A material change compromises the MRO's independence from the artificial intelligence industry;

    - Evolution of technology renders the MRO's methods obsolete for ensuring acceptable levels of risk of personal injury and property damage; and/or,

    - An artificial intelligence model or artificial intelligence application certified by the MRO causes a significant harm.

- An applicant for designation as an MRO shall submit with its application a plan that contains all of the following elements:

  - The applicant's approach to auditing of artificial intelligence models and artificial intelligence applications to verify that an artificial intelligence developer has exercised heightened care and adhered to predeployment and postdeployment best practices and procedures to prevent personal injury or property damage caused by the artificial intelligence model or artificial intelligence application;

  - The applicant's approach to mitigating specific high-impact risks, including cybersecurity, chemical, biological, radiological, and nuclear threats, malign persuasion, and artificial intelligence model autonomy and exfiltration;

  - An approach to ensuring disclosure by developers to the MRO of risks detected, incident reports, and risk mitigation efforts;

  - An approach to specifying the scope and duration of certification of an artificial intelligence model or artificial intelligence application, including technical thresholds for updates requiring renewed certification;

  - An approach to data collection for public reporting from audited developers and vendors that addresses all of the following:

    - Aggregating and tracking evaluation data from certified labs;

    - Categories of metadata to be aggregated and tracked; and,

    - Measures to protect trade secrets and mitigate antitrust risk from information sharing.

  - The applicant's intended use, if any, of security vendors to evaluate artificial intelligence developers, models, or applications, including a method of certifying and training vendors to accurately evaluate an artificial intelligence model or developer exercising heightened care and complying with best practices;

  - Implementation and enforcement of whistleblower protections among certified developers;

  - Remediation of postcertification noncompliance;

  - An approach to reporting of societal risks and benefits identified through auditing; and,

  - An approach to interfacing effectively with federal and non-California state authorities.

- The MRO's plan may be tailored to a particular artificial intelligence market segment.

- An applicant shall annually audit all of the following to ensure independence from the artificial intelligence industry and report the findings of its audit to the Attorney General:
  - The applicant's board composition;
  - The availability of resources to implement the applicant's plan;
  - The applicant's funding sources; and,
  - Representation of civil society representatives in evaluation and reporting functions.
- The Attorney General shall not modify a plan.
- The Attorney General shall adopt regulations, with input from stakeholders that establish both of the following:
  - Minimum requirements for plans; and,
  - Conflict of interest rules for MROs that include, but are not limited to, reporting requirements on boards of directors and donors funding the MRO to ensure adequate independence from the artificial intelligence industry and transparency on revenues streaming from certification services.
- The Attorney General may establish a fee structure for charging fees to applicants and designated MROs to offset the reasonable costs incurred by the Attorney General.
- The Attorney General may adopt regulations necessary to administer these provisions.
- An MRO shall do all of the following:
  - Ensure developers' and security vendors' exercise of heightened care and compliance with best practices for the prevention of personal injury and property damage and certify qualified artificial intelligence models or artificial intelligence applications that meet the requirements prescribed by the MRO;
  - Implement the plan submitted;
  - Decertify an artificial intelligence model or artificial intelligence application that does not meet the requirements prescribed by the MRO;
  - Submit to the Legislature, and to the Attorney General an annual report that addresses all of the following:
    - Aggregated information on capabilities of artificial intelligence models, the observed societal risks and benefits associated with

        those capabilities, and the potential societal risks and benefits associated with those capabilities;

        - The adequacy of existing evaluation resources and mitigation measures to mitigate observed and potential risks;

        - Developer and security vendor certifications;

        - Aggregated results of certification assessments;

        - Remedial measures prescribed by the MRO and whether the developer or security vendor complied with those measures; and,

        - Identified additional risks outside personal injury or property damage and the adequacy of existing mitigation measures to address those risks; and,

    - Retain for 10 years a document that is related to the MRO’s activities under this chapter.

- In a civil action asserting claims for personal injury or property damage caused by an artificial intelligence model or artificial intelligence application against a developer of the artificial intelligence model or artificial intelligence application, there shall be a rebuttable presumption that the developer exercised reasonable care if the artificial intelligence model or artificial intelligence application in question was certified by an MRO at the time of the plaintiff’s injuries. The rebuttable presumption provided for in this section may be overcome by the introduction of admissible evidence the court finds contrary to the presumption.

Related Legislation: This bill is one of a many of bills related to AI this Legislative Session:

- SB 53 (Weiner) establishes a consortium develop a framework for the creation of a public cloud computing cluster to advance the development of AI that is safe, ethical, equitable, and sustainable. SB 53 is pending on this Committee’s suspense file.

- SB 366 (Smallwood Cuevas) creates a study evaluating the impact of AI on worker well-being. SB 366 is pending in the Senate Committee on Labor.

- SB 503 (Weber Pierson) requires developers of patient care decision support tools and health facilities to make reasonable efforts to identify uses of patient care decision support tools in health programs. SB 503 is pending in Senate Judiciary Committee.

- SB 524 (Arreguin) requires law enforcement agencies to note when they use AI on official reports. SB 524 is pending on this Committee’s Suspense File.

- SB 579 (Padilla) establishes a mental health and AI working group. SB 579 is pending on this Committee’s Suspense File.

- SB 833 (McNerney) requires a state agency in charge of critical infrastructure that deploys AI to establish a human oversight mechanism. SB 833 is pending in this Committee.

- AB 222 (Bauer-Kahan) requires reporting about energy use related to AI. AB 222 is pending in the Assembly Committee on Privacy and Consumer Protection.

- AB 316 (Krell) prohibits a defendant that used AI from asserting a defense that the AI autonomously caused the harm to the plaintiff. AB 316 is pending in the Assembly Committee on Privacy and Consumer Protection.

- AB 410 (Wilson) requires bots using AI to disclose that they are bots. AB 410 is pending on the Assembly Appropriations Committee Suspense File.

- AB 412 (Bauer Kahan) requires a of a generative AI model to document any copyrighted materials used to train the model. AB 412 is pending in the Senate Judiciary Committee.

- SB 420 (Padilla) regulates high-risk automated decision systems. SB 420 is pending in this Committee.

- SB 468 (Becker) imposes a duty on business that deploy a high-risk AI systems that processes personal information to protect personal information. SB 468 is pending in this Committee.

- AB 489 (Bonta) makes provisions of law that prohibit the use of specified terms, letters, or phrases to falsely indicate or imply possession of a license or certificate to practice a health care profession enforceable against an entity who uses AI. AB 489 is pending in the Assembly Appropriations Committee.

- AB 853 (Wicks) requires a large online platform to retain any available provenance data in content posted on the large online platform. AB 853 is pending in the Senate Judiciary Committee.

- AB 979 (Irwin) develops a California AI Cybersecurity Collaboration Playbook to facilitate information sharing across the AI community. AB 979 is pending in the Assembly Committee on Privacy and Consumer Protection.

- AB 1018 (Bauer-Kahan) regulates automated decision systems. AB 1018 is pending in the Assembly Judiciary Committee.

- AB 1064 (Bauer-Kahan) adopts criteria for determining the level of estimated risk of an AI system on children. AB 1064 is pending in the Assembly Judiciary Committee.

- AB 1159 (Addis) prohibits using student personal information to train AI. AB 1159 is pending in the Assembly Committee on Privacy and Consumer Protection.

- AB 1405 (Bauer-Kahan) establishes a mechanism allowing natural persons to report misconduct by AI auditors. AB 1405 is pending on the Assembly Appropriations Committee Suspense File.

-- END --