CA bills on AI auditing

Sarah Cen | CMU AIAH | April 24, 2026

Motivation

There has been a lot of interest in AI audits

Why do we need audits?

1. Without the ability to test compliance, a law has no bite
2. Some laws require testing or assessments

But this process is messy. What should audits test for? What prevents them from becoming check-boxing exercise? Who should do the audits? Who pays for audits? How does this affect small businesses?

Process (for CA)

1. **Idea**: by interested party, agency, lobbyist, activist group, even researchers
2. **Shopping**: find a member of legislature (assembly or senate) to author
3. **Drafting language**: create specific language, get private stakeholder input to test political viability
4. **Introduce bill**: bill is assignment to a committee ➡️ which committee affects future
5. **Committee(s)**: hold hearings, take expert & public testimony, debate bill, add amendments ➡️ may move to Appropriations (financial) committee. This is where most bills die (either silently or by vote)
6. **Chamber**: if passes, goes to Assembly/Senate ➡️ subject to debate and negotiation, then amended and passed to other chamber if passed. If both chambers pass, reconcile versions and re-pass. More predictable but may be subject to a lot of negotiation and amendments.
7. **Governor**: sign into law, veto, or allow to become law without signature. Very predictable.

This process is usually Dec through October

Note: CA has ballot initiative process where it can bypass the legislature.

SB-813 (2025-26)

California AI Standards and Safety Commission: independent verification organizations

Overall goals:

1. Establish CA AI Standards and Safety Commission
2. State-backed standards and certification system for AI safety with independent orgs to set/enforce standards
3. Company participation is voluntary

Designation of IVOs (independent verification organizations)

1. Commission oversees these certification orgs and their approval
2. IVOs do audits, certification, ongoing monitoring

Some obligations of commission

"The commission shall determine whether an applicant IVO's plan ensures acceptable mitigation of risk from any IVO-verified artificial intelligence model and artificial intelligence application by considering all of the following:

... (A) The viability and rigor of the applicant's evaluation methods, technologies, and administrative procedures.

(B) The adequacy of the applicant's plan to develop measurable standards for evaluating artificial intelligence developers' and deployers' mitigation of risks in the development and deployment of an artificial intelligence model or artificial intelligence application.

(C) The adequacy of the applicant's procedures for ongoing supervision of artificial intelligence models or artificial intelligence applications...

... The commission may revoke a designation if any of the following is true:

(1) The IVO's plan is materially misleading or inaccurate.

(2) The IVO systematically fails to adhere to its plan.

(3) A material change compromises the IVO's independence from the artificial intelligence industry.

(4) Evolution of technology renders the IVO's methods obsolete for ensuring acceptable levels of risk of personal injury, reasonable risk of foreseeable harm, and property damage."

IVOs submit applications for approval

"(a) An applicant to the commission for designation as an IVO shall submit with its application a plan that contains all of the following elements:

(1) The applicant's approach to auditing of artificial intelligence models and artificial intelligence applications to verify that an artificial intelligence developer or deployer has exercised heightened care and adhered to predeployment and postdeployment best practices and procedures to prevent personal injury, reasonably foreseeable harm, or property damage caused by the artificial intelligence model or artificial intelligence application…

(3) An approach to ensuring disclosure by developers and deployers to the IVO of risks detected, material changes to risk profiles, including risks detected before verification and risks resulting from fine-tuning or modifying an artificial intelligence model or artificial intelligence application after verification, incident reports, and risk mitigation efforts for a particular artificial intelligence model or artificial intelligence application.

(4) All of the following with respect to any risk the applicant intends to verify …

(A) A proposed definition of acceptable levels of risk.

(B) Metrics that are measurable and can be used to determine whether the acceptable level of risk…

(C) Target levels for the metrics, including data sources those levels are based on and methods for measurement.

(D) A description of the evaluation and reporting protocol to determine whether verified models of applications meet the outcomemetrics on an ongoing basis.

(5) 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.

(6) An approach to data collection for public reporting from audited developers, deployers, and vendors that addresses all of the following:

(A) Aggregating and tracking evaluation data from certified labs.

(B) Categories of metadata to be aggregated and tracked.

(C) Measures to protect trade secrets and mitigate antitrust risk from information sharing."

IVO responsibilities

"An IVO designated pursuant to this chapter shall do all of the following:

(a) Ensure developers', deployers', 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…

(d) Submit to the Legislature, pursuant to Section 9795, and to the commission an annual report that addresses all of the following:

(1) 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.

(2) The adequacy of existing evaluation resources and mitigation measures to mitigate observed and potential risks.

(3) Developer, deployer, and security vendor certifications.

(4) Aggregated results of certification assessments.

(5) Remedial measures prescribed by the IVO and whether the developer, deployer, or security vendor complied with those measures.

(6) Identified additional risks outside personal injury or property damage and the adequacy of existing mitigation measures to address those risks.

(e) An IVO shall annually audit all of the following to ensure independence from the artificial intelligence industry and report the findings …

(1) The applicant's board composition.

(2) The availability of resources to implement the applicant's plan.

(3) The applicant's funding sources.

(4) Representation of civil society representatives in evaluation and reporting functions.

(f) Retain for 10 years a document that is related to the IVO's activities under this chapter."

Definitions

"(a) “Artificial intelligence application” means a software program or system that uses artificial intelligence models to perform tasks that typically require human intelligence.

(b) “Artificial intelligence model” means an engineered or machine-based system that varies in its level of autonomy and that can, for explicit or implicit objectives, infer from the input it receives how to generate outputs that can influence physical or virtual environments.

(d) “Deployer” means a person or entity that implements, integrates, or makes operational an artificial intelligence model or artificial intelligence application within the state, including a person or entity that makes an artificial intelligence model or artificial intelligence application available for use by others within the state, whether directly or as part of a product or service.

(e) “Developer” means a person who develops an artificial intelligence model or artificial intelligence application that is deployed in the state.

(f) “Independent verification organization (IVO)” means a private entity, nonprofit organization, academic consortium, or multistakeholder partnership designated as an IVO by the commission pursuant to this chapter."

Thoughts on SB-813?

(Note: removed rebuttal presumption from earlier versions)

AB-1405 (2025-26)

Artificial intelligence: auditors: enrollment

Overall goal: Establish system to certify and register auditors

Main things:

1. GOA oversees enrollment/qualifications + enrollment fee
2. Seeks to prevent COIs and protect confidential information
3. Auditors only do what is required by (other) law

Definitions

"(b) “Artificial intelligence” or “AI” means an engineered or machine-based system that varies in its level of autonomy and that can, for explicit or implicit objectives, infer from the input it receives how to generate outputs that can influence physical or virtual environments.

(c) “Artificial intelligence auditor” or “AI auditor” means a person, partnership, or corporation that assesses an AI system or model on behalf of a third party.

(d) “Covered audit” means an audit conducted pursuant to any state statute that requires an audit of an AI system or model by an independent third party auditor."

Enrollment

"(a) Beginning January 1, 2027, prior to initially conducting a covered audit, an AI auditor shall do all of the following:

(1) Enroll with the agency using the mechanism established pursuant to paragraph (1) of subdivision (a) of Section 11549.82.

(2) Pay to the agency the enrollment fee set forth in paragraph (2) of subdivision (a) of Section 11549.82.

(3) Provide to the agency all of the following information:

(A) The legal name of the auditor.

(B) All of the following contact information:

(i) The primary physical address of the auditor, if applicable.

(ii) The primary internet website of the auditor, if applicable.

(iii) A telephone number enabling a natural person to communicate with the auditor.

(iv) An email address enabling a natural person to communicate with the auditor.

(C) The types of AI systems or models that the auditor is enrolling to audit.

(D) Any relevant certifications or accreditations held by the AI auditor, and the names of the certifying or accrediting entities.

(E) A written description of the auditor and the services they provide, not to exceed 200 words in length.

(F) A standard operating procedure that does both of the following:

(i) Describes the auditor's protocols in sufficient detail to enable a third party to assess whether audits are conducted according to widely recognized industry standards appropriate to the system or model being audited.

(ii) Includes documentation substantiating any claims made by the AI auditor regarding the accuracy, reliability, or validity of its protocols.

(b) In conducting a covered audit, an enrolled AI auditor shall abide by widely recognized industry standards ..."

Audit report

"(a) After conducting a covered audit, an enrolled AI auditor shall provide the auditee with an audit report that contains…

(1) The scope and objectives of the audit.

(2) The results of the audit and any documentation necessary to demonstrate the basis of those results.

(3) An explanation of any steps the auditee can take to meet widely recognized industry standards …

(4) An explanation of any steps the auditee can take to become compliant with state law.

(5) A statement that is signed and dated by each auditor that certifies that the covered audit was completed.

…

(c) An enrolled AI auditor shall retain any documentation that is provided to an auditee pursuant to this chapter, or that is necessary to demonstrate the basis of the result of a covered audit, for at least 10 years.

(d) An enrolled AI auditor shall not conduct a covered audit if it has a financial interest in the auditee other than financial compensation for performing an audit.

…

(1) Notwithstanding Chapter 1 (commencing with Section 16600) of Part 2 of Division 7 of the Business and Professions Code, an enrolled AI auditor shall not accept employment with an auditee within 12 months of completing a covered audit of the auditee.

(2) An enrolled AI auditor shall not conduct a covered audit if the auditee had employed the auditor during the 12-month period preceding the audit."

Exceptions and extra tidbits/protections

"(a) An enrolled AI auditor may disclose confidential information concerning an auditee only if the auditee provides written authorization or if the disclosure is any of the following:

(1) Made in compliance with a subpoena or a summons enforceable by order of a court.

(2) Reasonably necessary to maintain or defend the auditor in a legal proceeding initiated by the auditee.

(3) Made in response to an official inquiry from a federal or state government regulatory agency.

(4) Made to another enrolled AI auditor or person in connection with a proposed sale or merger of the auditor's professional practice, provided the parties enter into a written nondisclosure agreement with regard to all auditee information shared between the parties.

(5) Made to either of the following:

(A) Another enrolled AI auditor to the extent necessary for purposes of professional consultation.

(B) Organizations that provide professional standards review and ethics or quality control peer review.

(6) Specifically permitted by state or federal law.

(b) An enrolled AI auditor shall not do either of the following:

(1) Prevent an employee from disclosing information to the Attorney General or the Labor Commissioner … if the employee has reasonable cause to believe the information indicates that the auditor is out of compliance with the if the employee has reasonable cause to believe the information indicates that the auditor is out of compliance with the requirements of this chapter.

(2) Retaliate against an employee for disclosing information pursuant to paragraph (1).

(a) Nothing in this chapter shall be construed to impede, delay, or otherwise affect the conduct of any audit required under any other statute or regulation that becomes operative prior to the effective dates of this chapter."

Thoughts?

AB-1018 (2025-26)

Automated decision systems

Overall goal: Regulate automated systems used to make consequential decisions

Main things:

1. GOA oversees enrollment/qualifications + enrollment fee
2. Seeks to prevent COIs and protect confidential information
3. Auditors only do what is required by (other) law

Definitions

Too many ... see here

Impact assessment (some requirements)

(a) (1) With respect to a covered ADS that was first deployed, or made available to potential deployers, before January 1, 2026, the developer of the covered ADS shall conduct an initial impact assessment on the covered ADS before January 1, 2027, and annually thereafter.

(2) With respect to a covered ADS that is first deployed or made available to potential deployers on or after January 1, 2026, the developer of the covered ADS shall conduct an impact assessment on the covered ADS before initially deploying the covered ADS, or making the covered ADS available to potential deployers, and annually thereafter.

(b) In conducting an impact assessment on a covered ADS, a developer shall do all of the following:

(1) Describe the purpose of the covered ADS.

(2) List and describe all developer-approved uses of the covered ADS.

(3) For each developer-approved use, document all of the following:

(A) The expected accuracy and reliability of the covered ADS.

(B) Whether any disparate treatment is intended to occur and, if so, all of the following:

(i) The conditions under which each category of disparate treatment is intended to occur.

(ii) Whether each category of disparate treatment is necessary for the developer-approved use.

(iii) Whether any alternatives not involving disparate treatment were considered.

(C) Whether any disparate impacts are reasonably likely to occur and, if so, all of the following:

(i) The conditions under which each category of disparate impact is reasonably likely to occur.

(ii) Whether each category of disparate impact is necessary for the developer-approved use.

(iii) Whether any alternatives not involving disparate impacts were considered.

The "bite"

(a) Any of the following public entities may bring a civil action against a developer, deployer, or auditor who violates this chapter:

(1) The Attorney General.

(2) A district attorney, county counsel of any county within which a city has a population in excess of 750,000, city attorney in a city and county, or city attorney of a city having a population in excess of 750,000.

(3) A city prosecutor in any city having a full-time city prosecutor with the consent of the district attorney.

(4) The Civil Rights Department.

(5) The Labor Commissioner with respect to employment-related decisions only.

(b) A court may award a prevailing plaintiff who brings an action pursuant to subdivision (a) all of the following:

(1) Injunctive relief.

(2) Declaratory relief.

(3) Reasonable attorney's fees and litigation costs.

(4) (A) A civil penalty of up to twenty-five thousand dollars ($25,000) per violation.

(B) In assessing the amount of the civil penalty pursuant to this paragraph, the court may consider relevant circumstances presented by the parties to the action, including, but not limited to, all of the following:

(i) The nature and severity of the misconduct.

(ii) The number of violations.

(iii) The length of time over which the misconduct occurred and the persistence of the misconduct.

(iv) The willfulness of the misconduct.

(v) The defendant's assets, liabilities, and net worth.

(c) A developer or deployer who contracts with a third party to perform the developer's or deployer's duties under this chapter, other than those duties related to auditing, is subject to liability under this chapter for the third party's failure to perform those duties.

Thoughts?