



California's AI Transparency Act (CAITA) May be Amended to Regulate Social Media Platforms

What You Need to Know

Key takeaway #1

Governor Newsom has until October 12 to sign AB 853 into law.

Key takeaway #2

If the governor signs AB 853, then the amended CAITA goes into effect in August 2026. If he does not sign AB 853, then the current version of CAITA goes into effect on January 1, 2026.

Key takeaway #3

AB 853 includes additional obligations on “large online platforms” and manufacturers of “capture devices”.

Client Alert | 5 min read | 10.08.25

Last year, the California General Assembly passed the California AI Transparency Act (CAITA), which Governor Gavin Newsom signed into law on September 19, 2024, and goes into effect on January 1, 2026. This may change because this year, the same General Assembly passed AB 853, an amendment to CAITA with potentially far-reaching implications.

CAITA currently applies only to certain providers of publicly accessible generative artificial intelligence (GenAI) systems. However, the amended version of CAITA states that “large online platforms” and manufacturers of so-called “capture devices” must also comply with certain data preservation and disclosure requirements.

AB 853 defines large online platforms as social media platforms with at least two million monthly users during the preceding twelve months; capture devices are “device[s] that can record photographs, audio, or video content.” The definition of “covered provider” under the original CAITA arguably already included these platforms and devices if they were engaged in the provision or licensing of a covered GenAI system, but now their coverage under the amended CAITA would be explicit and their obligations more specifically tailored.

Governor Newsom has until October 12, 2025, to sign AB 853 into law, after which the effective date of CAITA's general requirements will be delayed until August 2, 2026. If the governor does not sign AB 853 into law, then the current version of CAITA will take effect on January 1, 2026.

Overview of the Current Version of CAITA

The current version of CAITA is, at its core, a notice law. It requires certain businesses that produce GenAI outputs to include “latent” disclosures within these outputs. A “latent” disclosure is defined under the law as “present but not manifest.” Covered businesses must offer users the option to include “manifest” disclosures on their GenAI outputs, with “manifest” defined as “easily perceived, understood, or recognized by a natural person.”

CAITA currently applies to any person or business “that creates, codes, or otherwise produces a [GenAI] system that has over 1,000,000 monthly visitors or users and is publicly accessible within the geographic boundaries of” California. As noted above, CAITA requires these covered providers to:

- Include a latent disclosure on all AI-generated content that is (a) consistent with industry standards and (b) includes the name of the covered provider, the name and version of the GenAI system, the date and time of creation, and a unique identifier;
- Provide an AI detection tool for users to assess whether image, video, or audio content was created or altered via GenAI;
- Offer users the option to include a manifest disclosure on AI-generated content that is conspicuous, appropriate to the medium, and identifies the origin of the content; and
- Monitor the conduct of any third-party licensees of a covered provider's GenAI system.

Importantly, AB 853 would not change these provisions, only extends their effective date.

The Amendment to CAITA

The Amendment Clarifies How Large Online Platforms and Capture Device Manufacturers Must Make Provenance Data Available to Users

The proposed amended CAITA would require large online platforms to detect and maintain provenance data—akin to a chain of custody record for training data—in compliance with widely accepted industry standards and to provide users with an interface to inspect that provenance data. The proposed amendments would also require capture device manufacturers, to the extent technically feasible, to provide users with the option to include specified latent disclosure information in the content they capture or alter. These requirements are more detailed and thus arguably more burdensome than the requirements in the existing law.

However, like the original CAITA, the proposed amendments would still require covered entities to make AI detection tools available at no cost to users. And like the original CAITA, the proposed amendments would leave intact the original law's requirements regarding latent and manifest disclosures for AI-generated content.

The Potential Impact of Provenance Data Preservation Requirements on Large Online Platforms

Unlike the current version of CAITA, the proposed amendment would require large online platforms to store provenance data in a manner that can be inspected easily by users. It would also prohibit such platforms from knowingly stripping that provenance data. Preserving and making this provenance data available will likely come with a cost, such as operational costs for data storage, management, and security.

Regulated entities, including platforms, must also make sure that these latent disclosures in AI-generated outputs are preserved in a manner that can be identified by an AI-detection tool. There will be a cost associated with this requirement as well. Depending on which standards and technologies are used to create latent disclosures, these AI detection tools could be bespoke to particular platforms, or could be more universal such that they can also detect GenAI content from social media platforms and capture devices.

Moreover, although CAITA requires the maintenance of specific provenance data, it also prohibits covered providers from collecting or retaining certain personal information from users and prohibits these providers from retaining any content submitted to an AI detection tool “for longer than is necessary to comply” with this law. Covered entities must therefore be careful to maintain only the provenance data required for content identification, while otherwise complying with California’s data privacy regulations.

Impact of CAITA Whether or Not It is Amended

Regardless of whether Governor Newsom signs AB 853 into law, the most significant impact of CAITA will stem from the penalties it imposes on companies for removing or failing to maintain provenance data, and from the requirement that covered providers offer AI detection tools to users. These penalties include liability in the amount of \$5,000 per violation and attorneys’ fees for a prevailing plaintiff.

These requirements could help combat disinformation, but they could also create a false sense of security that enables more sophisticated bad actors to exploit user reliance on these tools as a source of truth. It also remains unclear whether these detection tools must be able to identify content generated by any system, or only content generated by the covered provider offering that specific tool.

Implementation Timeline for AB 853

Provisions of CAITA	Effective Date
General Provisions	August 2, 2026
Large Online Platforms Provisions	January 1, 2027
Capture Device Manufacturers Provisions	January 1, 2028

Contacts

Jennie Wang VonCannon

Partner

She/Her/Hers

Los Angeles D | +1.213.310.7984

jvoncannon@crowell.com

Kate M. Growley

Partner, Crowell Global Advisors Senior Director

Washington, D.C. D | +1.202.624.2698

Washington, D.C. (CGA) D | +1 202.624.2500

kgrowley@crowell.com

Rafi Azim-Khan

Partner

London D | +44.20.7413.1307

San Francisco D | +1.415.365.7282

rafi@crowell.com

Meaghan Katz

Associate

They/Them/Theirs

San Francisco D | +1.415.365.7287

mkatz@crowell.com

Jacob Canter

Counsel

He/Him/His

San Francisco D | +1.415.365.7210

jcanter@crowell.com