Office of Senate Floor Analyses
(916) 651-1520 Fax: (916) 327-4478

THIRD READING

Bill No: AB 853
Author: Wicks (D), et al.
Amended: 9/5/25 in Senate
Vote: 21

SENATE JUDICIARY COMMITTEE: 11-0, 7/15/25
AYES: Umberg, Allen, Arreguín, Ashby, Caballero, Durazo, Laird, Stern, Wahab, Weber Pierson, Wiener
NO VOTE RECORDED: Niello, Valladares

SENATE APPROPRIATIONS COMMITTEE: 5-2, 8/29/25
AYES: Caballero, Cabaldon, Grayson, Richardson, Wahab
NOES: Seyarto, Dahle

ASSEMBLY FLOOR: 58-2, 6/2/25 - See last page for vote

SUBJECT: California AI Transparency Act

SOURCE: California Initiative on Technology and Democracy

DIGEST: This bill establishes requirements on large online platforms, capture device manufacturers, and generative AI (GenAI) system hosting platforms to embed and disclose provenance data in certain GenAI created or altered content.

Senate Floor amendments of 9/5/25 delay implementation of the existing statute, clarify its application, and narrow the scope by, in part, including exemptions and removing several of the requirements placed on large online platforms and capture device manufacturers.

ANALYSIS:

Existing law:

1) Establishes the California AI Transparency Act, which becomes operative on January 1, 2026. (Business and Professions Code (Bus. & Prof. Code) § 22757 et seq.)

2) Requires a "covered provider," a person 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 the state, to make an AI detection tool available at no cost by which a person can assess whether content was created or altered by the provider's GenAI system. (Bus. & Prof. Code § 22757.2(a).)

3) Prohibits a covered provider from doing any of the following in carrying out the duties above:

    a) Collect or retain personal information when a person utilizes the covered provider's AI detection tool, except that it may collect and retain the contact information of a person who submitted feedback.
    b) Retain any content submitted to the AI detection tool for longer than is necessary to comply with this law. (Bus. & Prof. Code § 22757.2(c).)

4) Requires a covered provider to offer users the option to include in AI-generated image, video, or audio content created by its own generative AI system a manifest disclosure that meets specified criteria, including that it identifies the content as AI-generated content. (Bus. & Prof. Code § 22757.3(a).)

5) Requires a covered provider to include in AI-generated image, audio, and video content created by its generative AI system a latent disclosure that is detectable by the tool specified above and is, to the extent technically feasible, permanent or extraordinarily difficult to remove. (Bus. & Prof. Code § 22757.3(b).)

6) Provides that a covered provider that violates the above provisions is liable for a civil penalty in the amount of $5,000 per violation to be collected in a civil action filed by the Attorney General, a city attorney, or a county counsel. Each day that a covered provider is in violation shall be deemed a discrete violation. (Bus. & Prof. Code § 22757.4.)

This bill:

1) Requires a large online platform to do both of the following:

   a) Detect whether any provenance data that is compliant with widely adopted specifications adopted by an established standards-setting body is embedded into or attached to content distributed on the large online platform.
   b) Provide a user interface to disclose the availability of system provenance data that reliably indicates that the content was generated or substantially altered by a GenAI system or captured by a capture device. The user interface shall make clearly and conspicuously available to users information sufficient to identify the content's authenticity, origin, or history of modification, including specified information such as whether provenance data is available.
   c) Allow a user to inspect all available system provenance data that is compliant with widely adopted specifications adopted by an established standards-setting body in an easily accessible manner by any of several specified means.

2) Provides that a large online platform shall not, to the extent technically feasible, knowingly strip any system provenance data or digital signature that is compliant with widely adopted specifications adopted by an established standards-setting body from content uploaded or distributed on the large online platform.

3) Makes the above provisions operative on January 1, 2027.

4) Defines "digital signature" as a cryptography-based method that identifies the user or entity that attests to the information provided in the signed section. "Large online platform" means a public-facing social media platform, file-sharing platform, mass messaging platform, or stand-alone search engine that distributes content to users who did not create or collaborate in creating the content that exceeded 2,000,000 unique monthly users during the preceding 12 months. It does not include a broadband internet access service or telecommunications service, as provided.

5) Requires a capture device manufacturer, with respect to any capture device the capture device manufacturer first produced for sale in the state on or after January 1, 2028, to do the following to the extent technically feasible and

compliant with widely adopted specifications adopted by an established standards-setting body:

a) Provide a user with the option to include a latent disclosure in content captured by the capture device that conveys specified information, including identifying information for the manufacturer and the device, as well as the time and date of the content's creation or alteration.
b) Embed latent disclosures in content captured by the device by default.

6) Defines "capture device" as a device that can record photographs, audio, or video content, including video and still photography cameras, mobile phones with built-in cameras or microphones, and voice recorders.

7) Prohibits, as of January 1, 2027, a GenAI system hosting platform from knowingly making available a GenAI system that does not place disclosures pursuant to existing Section 22757.3.

8) Defines "GenAI hosting platform" as a website or application that makes available for download the source code or model weights for a GenAI system by a resident of the state, regardless of whether the terms of that use include compensation.

9) Subjects those in violation of these provisions to the liability currently imposed by the California AI Transparency Act.

10) Includes a severability clause.

11) Delays operation of the California AI Transparency Act until August 2, 2026.

Background

Certain forms of media – audio recordings, video recordings, and still images – can be powerful evidence of the truth. While such media have always been susceptible to some degree of manipulation, fakes were relatively easy to detect. The rapid advancement of AI technology, specifically the wide-scale introduction of GenAI models, has made it drastically cheaper and easier to produce synthetic content created, audio, images, text, and video recordings that are not real, but that are so realistic that they are virtually impossible to distinguish from authentic content, including so-called "deepfakes."

This bill builds on, and delays the operative date of, the California AI Transparency Act (CAIT Act), passed last year. It places obligations on large online platforms, manufacturers of devices that can record photographs, audio, or video content, like cameras and phones, and GenAI system hosting platforms.

This bill is sponsored by the California Initiative on Technology and Democracy (CITED). It is supported by several groups, including Consumer Reports. This bill is opposed by several industry associations, including the Consumer Technology Association.

Comments

Last year, the Legislature responded to issues concerning the proliferation of AI-generated content by passing the CAIT Act, SB 942 (Becker, Chapter 291, Statsutes 2024), which is set to become operative on January 1, 2026. The CAIT Act requires providers to make an AI detection tool available at no cost by which a person can assess whether content was created or altered by the provider's GenAI system. The CAIT Act also regulates AI-generated images, video, or audio that are created by a GenAI system. Covered providers are required to include a latent disclosure in such content that is detectable using the above tool, and that is, to the extent technically feasible, permanent or extraordinarily difficult to remove. This latent disclosure must identify the provider, the tool, and the time and date of the content's creation or alteration. Covered providers are also required to provide users making such content with their system with the option to include a manifest disclosure that identifies it as AI-generated content.

A covered provider that violates the CAIT Act is liable for a civil penalty in the amount of $5,000 per violation to be collected in a civil action filed by the Attorney General, a city attorney, or a county counsel. Each day that a covered provider is in violation is a discrete violation.

This bill seeks to bolster the CAIT Act by establishing similar transparency requirements on large online platforms, capture device manufacturers, and GenAI system hosting platforms. However, it delays the operation of the CAIT Act until August 2, 2026.

According to the author:

> New and emerging developments of generative AI (GenAI) tools have made it easier to create, edit, and doctor images, video, and audio. AI technologies can create and manipulate content to look realistic and

convincing, which allows bad actors to create harmful content and spread disinformation. AB 853 will help mitigate some of the harmful impacts of AI-generated content and provide more transparency of content in the digital information ecosystem by ensuring that large online platforms and capture devices provide more information for users to understand the source of content.

FISCAL EFFECT: Appropriation: No Fiscal Com.: Yes Local: No

According to the Senate Appropriations Committee:

- The Department of Justice (DOJ) reports a fiscal impact of approximately $1 million or less (Unfair Competition Law Fund). DOJ notes that implementation of this bill will be dependent upon the appropriation of funds. The DOJ will be unable to absorb the costs to comply with or implement the requirements of the bill within existing budgeted resources. DOJ reports that the Consumer Protection Section (CPS) within the Public Rights Division anticipates increased workloads in investigating and enforcing violations of this bill, beginning on January 1, 2026, and ongoing. CPS will require additional resources consisting of two Deputy Attorneys General (DAGs), one Associate Governmental Program Analyst (AGPA), and one Legal Secretary.

- Unknown, potentially significant costs to the state funded trial court system (Trial Court Trust Fund, General Fund) to adjudicate civil actions. By expanding a civil penalty with statutory damages, this bill may lead to additional case filings that otherwise would not have been commenced. Expanding civil penalties could lead to lengthier and 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 unknowns, 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. 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.

SUPPORT: (Verified 9/8/25)

California Initiative on Technology and Democracy (Source)
Consumer Reports
Tech Oversight California

TechEquity Action
Transparency Coalition.ai
Truepic

OPPOSITION: (Verified 9/8/25)

California Civil Liberties Advocacy
Computer & Communications Industry Association
Consumer Technology Association
Recording Industry Association of America
Technet

ARGUMENTS IN SUPPORT: CITED, the sponsor of the bill, writes:

> Generative artificial intelligence (GenAI) technologies are powerful tools capable of creating all manners of images, audio, video, and text content from simple prompts. The breakneck speed at which these tools have evolved has meant human beings are increasingly unable to tell the difference between authentic, human-generated content, and synthetic content generated by AI.[1]
>
> The impact of this increasingly blurry line between authentic and synthetic digital media is already being felt by our society. From supercharging online scams,[2] to using child sexual abuse material to generate non-consensual intimate imagery,[3] to the proliferation of public safety[4] and political disinformation,[5] GenAI tools have contributed to the steady erosion of trust in our information ecosystem. Without adequate tools to help differentiate between human-generated authentic content and AI-generated synthetic

---

[1] Nightingale & Farid, *AI-synthesized faces are indistinguishable from real faces and more trustworthy, Proceedings of the National Academy of Sciences* (Feb. 14, 2022), **https://www.pnas.org/doi/full/10.1073/pnas.2120481119**.
[2] Bob Violino, *AI tools such as ChatGPT are generating a mammoth increase in malicious phishing emails*, CNBC (Nov. 28, 2023), **https://www.cnbc.com/2023/11/28/ai-like-chatgpt-is-creating-huge-increase-in-malicious-phishing-email.html**.
[3] Thiel, David, *Identifying and Eliminating CSAM in Generative ML Training Data and Models*, Stanford Digital Repository (Dec. 20, 2023), **https://purl.stanford.edu/kh752sm9123**.
[4] Shannon Bond, *Fake viral images of an explosion at the Pentagon were probably created by AI*, NPR (May 22, 2023), https://www.npr.org/2023/05/22/1177590231/fake-viral-images-of-an-explosion-at-the-pentagon-were-probably-created-by-ai.
[5] Alex Isenstadt, *Desantis PAC uses AI-generated Trump voice in ad attacking ex-president*, Politico (Jul. 17, 2023), **https://www.politico.com/news/2023/07/17/desantis-pac-ai-generated-trump-in-ad-00106695**.

content, the truth decay[6] already happening in our society will only accelerate.

ARGUMENTS IN OPPOSITION: The Consumer Technology Association writes:

> AI is not static—it is dynamic infrastructure that evolves at exponential velocity. Regulating at the application layer, before consensus has formed on the tooling, provenance standards, or technological feasibility, risks freezing innovation in place. That is the essential flaw of AB 853: it attempts to legislate before the tools for compliance are widely available, widely adopted, or even fully developed.
>
> CTA supports meaningful transparency and accountability in AI-generated content. In fact, we've published ANSI/CTA-2125, a standard which can address content provenance and assurance by providing a foundation for detecting and labeling AI-generated media in a consistent way. But like all standards, it requires time, industry convergence, and implementation capacity. AB 853 ignores this timeline, imposing requirements without the supporting ecosystem.

ASSEMBLY FLOOR: 58-2, 6/2/25
AYES: Addis, Aguiar-Curry, Ahrens, Alvarez, Arambula, Ávila Farías, Bauer-Kahan, Bennett, Berman, Boerner, Bonta, Bryan, Calderon, Caloza, Carrillo, Connolly, Elhawary, Fong, Gabriel, Garcia, Gipson, Mark González, Haney, Harabedian, Hart, Irwin, Jackson, Kalra, Krell, Lee, Lowenthal, McKinnor, Muratsuchi, Nguyen, Ortega, Pacheco, Papan, Patel, Pellerin, Petrie-Norris, Quirk-Silva, Ramos, Ransom, Celeste Rodriguez, Michelle Rodriguez, Rogers, Blanca Rubio, Schiavo, Schultz, Sharp-Collins, Solache, Soria, Stefani, Ward, Wicks, Wilson, Zbur, Rivas
NOES: DeMaio, Patterson

[6] J. Kavanagh & M. Rich, *Truth Decay: An Initial Exploration*, Rand, (Jan 16, 2018), **https://www.rand.org/pubs/research_reports/RR2314.html**.

NO VOTE RECORDED: Alanis, Bains, Castillo, Chen, Davies, Dixon, Ellis, Flora, Gallagher, Jeff Gonzalez, Hadwick, Hoover, Lackey, Macedo, Sanchez, Ta, Tangipa, Valencia, Wallis

Prepared by: Christian Kurpiewski / JUD. / (916) 651-4113
9/8/25 21:26:38

****** END ******