

1) **Purpose.** According to the author:

As AI technology advances, distinguishing between human and machine-generated content becomes increasingly challenging. This ambiguity poses significant risks to Californians, exacerbating problems of disinformation, harassment, and fraud while threatening the integrity of the information environment our democracy and economy depend on. In 2024, the legislature passed SB 942, the first bill in the nation to establish disclosure requirements for synthetic content. Since then, content provenance technology has developed in such a way that the leading technologies for embedding provenance are not accurately described by the law. Additionally, some methods of marking content are still in their nascency, and aren't yet robust enough to withstand adversarial exploitation or certain kinds of transformations (e.g. file compression, spoofing attacks, added noise to an image, or file type conversions). The original act also did not address the risk that visible labels create a binary signal implying that unlabeled content is authentic and labeled content is suspect - regardless of whether either inference is warranted. Furthermore, obligations for developers established SB 942 do not require that the information embedded in AI generated or modified content be consistent - or interoperable with - widely adopted standards. Last year, AB 853 (Wicks) established obligations for large online platforms to read provenance data, but only if it was compliant with such standards. This gap risks establishing a fragmented ecosystem where new methods of provenance disclosure that aren't interoperable with current standards are not read by large online platforms, and never subsequently relayed to users. This bill is critical to ensuring California's content labeling laws are effective at providing consumers with consistent information about where the content they see online comes from.

- 2) **Background.** The California AI Transparency Act, codified in the Business and Professions Code, requires a covered provider of a GenAI system to make available a free tool allowing users to assess whether content was generated or altered by the provider's system, to offer disclosures identifying AI-generated content, and to impose disclosure-related obligations on third parties licensing the provider's system. This bill refines those requirements. The most consequential change for the scope of the chapter is the removal of the 1,000,000-monthly-user threshold from the definition of "covered provider," which extends coverage from the largest systems to any GenAI system publicly accessible in the state. The bill also reworks the detection tool into a "disclosure verification tool" with revised functionality and privacy limits, revises the information a latent disclosure must convey and the standards with which a disclosure must be compatible, and shortens the window — from 96 to 72 hours — within which a covered provider must terminate a licensee's authorization after learning the licensee has modified a system out of compliance. Many of the new obligations are qualified as applying only "to the extent technically feasible."