Date of Hearing: April 30, 2024

ASSEMBLY COMMITTEE ON PRIVACY AND CONSUMER PROTECTION
Rebecca Bauer-Kahan, Chair
AB 1791 (Weber) – As Amended April 25, 2024

AS PROPOSED TO BE AMENDED

SUBJECT: Digital content provenance

SYNOPSIS

As generative artificial intelligence (GenAI) has become more accessible, a large amount of misleading and fake content has begun to flood social media. One way to contend with this onslaught of misinformation is to embed provenance data – information related to the origin of the content – into the products of GenAI systems. Digital cameras, such as those found in modern smartphones, already do something similar: the photographs they generate contain "metadata" that can be used to demonstrate the photo's authenticity. If GenAI developers were required to embed provenance data into their systems, and social media platforms were required to retain provenance data in user-uploaded content, then users might be better able to discern GenAI products from authentic content.

However, provenance data relays more than just the hardware or software used to generate a piece of content. Camera metadata can contain deeply personal information: for example, the name of the camera's owner, the time-of-day, down to the millisecond that a photo is taken, and the exact location of the camera's owner. When a photo containing this metadata is uploaded to social media, a user's privacy can be protected by stripping provenance data from uploaded content.

There is a tension here: retaining provenance data in user-uploaded content helps combat misinformation, but deleting provenance data from content helps protect privacy. This bill attempts to strike a balance between these priorities by defining two types of provenance data: personal provenance data, which relates to a users' identity; and system provenance data, which relates to a piece of content's system-of-origin. This bill would require social media platforms to delete personal provenance data and retain system provenance data. If it is not possible to do both, the bill would require a platform to delete both types of provenance data and to append a label to the content in order to preserve the system provenance data. This label would be retained when content is shared or reposted within a platform, and embedded back into content or its metadata when content is shared outside the platform or downloaded. This bill would make platforms liable for violations of these provisions by granting users a private right of action.

This bill is author-sponsored and has no registered support. The bill was recently amended to flesh out the provision above. With regard to the prior version of the bill, Electric Frontier Foundation adopted an "oppose unless amended" position, while a coalition of industry associations, including TechNet, had a "concerns" position.

Additional proposed Committee amendments would further clarify and refine the scope of the bill.

SUMMARY: Requires social media platforms to delete provenance data related to a user's identity from content uploaded to the platform, while retaining provenance data related to the system or service used to generate the content. Specifically, **this bill**:

1) Defines "provenance data" to mean data that is embedded into digital content, or that is included in the digital content's metadata, for the purpose of verifying the digital content's authenticity, origin, or history of modification, including, but not limited to, personal provenance data and system provenance data.

2) Defines "personal provenance data" to mean provenance data that contains personal information, as defined in Section 1798.140 of the Civil Code.

3) Defines "system provenance data" to mean provenance data that does either of the following:

   a) Identifies the device, system, or service that was used to generate a piece of digital content.

   b) Authenticates the digital content.

4) Requires platforms to redact personal provenance data from content uploaded to the platform by a user.

5) Prohibits platforms from redacting system provenance data from content uploaded to a platform by a user.

6) Requires a platform to redact both types of provenance data if the platform cannot redact personal provenance data without also redacting system provenance data, and requires the platform to attach a label to the content that does all of the following:

   a) Prominently discloses any system provenance data that was redacted by the social media platform.

   b) Not disclose any personal provenance data that was redacted by the social media platform.

   c) Remains attached to the content even if the content is shared, reposted, or otherwise replicated within the social media platform.

7) Requires a platform to embed the stripped system provenance data back into the content or the content's metadata if the content is downloaded, shared to an external internet website, or otherwise distributed such that the original platform can no longer control how it is displayed.

8) Requires platforms to abide by relevant industry standards to the greatest extent possible.

9) Permits a user of a platform to bring a civil action against a platform that fails to comply with this chapter with regard to content uploaded by that user for any of the following relief:

   a) Actual damages or ten thousand dollars per violation, whichever is greater.

   b) Injunctive relief.

c) Reasonable attorney's fees and costs.

EXISTING LAW:

1) Provides, pursuant to the California Constitution, that all people are by nature free and independent and have inalienable rights. Among these are the fundamental right to privacy. (Cal. Const. art. I, § 1.)

2) States that the "right to privacy is a personal and fundamental right protected by Section 1 of Article I of the Constitution of California and by the United States Constitution and that all individuals have a right of privacy in information pertaining to them." Further states these findings of the Legislature:

   a) The right to privacy is being threatened by the indiscriminate collection, maintenance, and dissemination of personal information and the lack of effective laws and legal remedies.

   b) The increasing use of computers and other sophisticated information technology has greatly magnified the potential risk to individual privacy that can occur from the maintenance of personal information.

   c) In order to protect the privacy of individuals, it is necessary that the maintenance and dissemination of personal information be subject to strict limits. (Civ. Code § 1798.1.)

3) Defines "deepfake" to mean audio or visual content that has been generated or manipulated by artificial intelligence (AI) which would falsely appear to be authentic or truthful and which features depictions of people appearing to say or do things they did not say or do without their consent. Requires the Secretary of Government Operations to evaluate the impact of the proliferation of deepfakes on the state. (Gov. Code § 11547.5.)

4) Defines "social media platform" to mean a public or semipublic internet-based service or application that has users in California and meets all of the following criteria: (Bus. & Prof. Code § 22589.)

   a) A substantial function of the service or application is to connect users in order to allow users to interact socially with each other within the service or application.

   b) A service or application that provides email or direct messaging services shall not be considered to meet this criterion on the basis of that function alone.

   c) The service or application allows users to construct a public or semipublic profile for purposes of signing into and using the service or application.

   d) The service or application allows users to populate a list of other users with whom an individual shares a social connection within the system.

   e) The service or application allows users to create or post content viewable by other users, including, but not limited to, on message boards, in chat rooms, or through a landing page or main feed that presents the user with content generated by other users.

5) Requires social media companies to post terms of service for each social media platform owned or operated by the company in a manner reasonably designed to inform all users of the social media platform of the existence and contents of the terms of service. (Bus. & Prof. § Code 22676.)

6) Defines "personal information" to mean information that identifies, relates to, describes, is reasonably capable of being associated with, or could reasonably be linked, directly or indirectly, with a particular consumer or household. States that personal information includes, but is not limited to, the following if it identifies, relates to, describes, is reasonably capable of being associated with, or could be reasonably linked, directly or indirectly, with a particular consumer or household (Civ. Code § 1798.140(v).):

   a) Identifiers such as a real name, alias, postal address, unique personal identifier, online identifier, Internet Protocol address, email address, account name, social security number, driver's license number, passport number, or other similar identifiers.

   b) Any personal information described in Section 1798.80(e).

   c) Characteristics of protected classifications under California or federal law.

   d) Commercial information, including records of personal property, products or services purchased, obtained, or considered, or other purchasing or consuming histories or tendencies.

   e) Biometric information.

   f) Internet or other electronic network activity information, including, but not limited to, browsing history, search history, and information regarding a consumer's interaction with an internet website application, or advertisement.

   g) Geolocation data.

   h) Audio, electronic, visual, thermal, olfactory, or similar information.

   i) Professional or employment-related information.

   j) Education information, defined as information that is not publicly available personally identifiable information as defined in the Family Educational Rights and Privacy Act (20 U.S.C. Sec. 1232g; 34 C.F.R. Part 99).

   k) Inferences drawn from any of the information identified in this subdivision to create a profile about a consumer reflecting the consumer's preferences, characteristics, psychological trends, predispositions, behavior, attitudes, intelligence, abilities, and aptitudes.

   l) Sensitive personal information.

FISCAL EFFECT: As currently in print, this bill is keyed nonfiscal.

COMMENTS: 1) **AI and GenAI.** AI is the mimicking of human intelligence by artificial systems such as computers. AI uses algorithms – sets of rules – to transform inputs into outputs.

Inputs and outputs can be anything a computer can process: numbers, text, audio, video, or movement. AI is not fundamentally different from other computer functions; its novelty lies in its application. Unlike normal computer functions, AI is able to accomplish tasks that are normally performed by humans.

AI that are trained on small, specific datasets in order to make recommendations and predictions are sometimes referred to as "predictive AI." This differentiates them from GenAI, which are trained on massive datasets in order to produce detailed text and images. When Netflix suggests a TV show to a viewer, the recommendation is produced by predictive AI that has been trained on the viewing habits of Netflix users. When ChatGPT generates text in clear, concise paragraphs, it uses GenAI that has been trained on the written contents of the internet.

2) **Deepfakes.** Image manipulation and video doctoring have existed for nearly as long as photography and recording equipment, but they have historically required great effort and talent. In the past few years the rapid development of GenAI has drastically reduced those barriers to entry, allowing a vast quantity of convincing – but ultimately fake – content to be generated in an instant. The creation of text, imagery, video, and audio by GenAI has the potential to change the world by automating repetitive tasks and fostering creativity. When employed by bad actors, however, these capabilities have the potential to destroy lives and destabilize societies.

Nonconsensual pornography. GenAI has been used to create pornography since its inception. This content is inevitably nonconsensual, and as GenAI technology improves, these products will become harder to distinguish from reality. Women have always been the primary victims of these efforts; in the run-up to the 2024 Super Bowl, a series of images involving Taylor Swift began to appear on the platform X (formerly Twitter.) These images were removed over the following days, but the damage had been done:

> "We are too little, too late at this point, but we can still try to mitigate the disaster that's emerging," says Mary Anne Franks, a professor at George Washington University Law School and president of the Cyber Civil Rights Initiative. Women are "canaries in the coal mine" when it comes to the abuse of artificial intelligence, she adds. "It's not just going to be the 14-year-old girl or Taylor Swift. It's going to be politicians. It's going to be world leaders. It's going to be elections."[1]

The harms of nonconsensual AI-powered pornography are already being felt in California:

> A third school in Southern California has been hit with allegations of digitally manipulated images of students circulating around campus . . . "Sixteen eighth-grade students were identified as being victimized, as well as five egregiously involved eighth-grade students," Superintendent Michael Bregy wrote. While Bregy acknowledged that children "are still learning and growing, and mistakes are part of the process," he affirmed disciplinary measures had been taken and noted that the incident was swiftly contained. The district vowed to hold accountable any other students "found to be creating, disseminating, or in possession of AI-generated images of this nature."[2]

[1] Brian Contreras, "Tougher AI Policies Could Protect Taylor Swift—And Everyone Else—From Deepfakes," Feb. 8. 2024, www.scientificamerican.com/article/tougher-ai-policies-could-protect-taylor-swift-and-everyone-else-from-deepfakes/.

[2] Mackenzie Tatananni, "'Inappropriate images' circulate at yet another California high school, as officials grapple with how to protect teens from AI porn created by classmates," *Daily Mail*, Apr. 11, 2024,

Scams. GenAI-powered speech and video is driving a new era in scamming. These AI tools are trained on publicly available data – the more data a target has online, the easier it is to develop a passable imitation of them or their loved ones. This is especially true of wealthy clients, whose public appearances, including speeches, are often widely available on the internet.[3] For example, a complicated scam utilizing both deepfake video and false audio was recently performed in Hong Kong. A multinational company lost $25.6 million after employees were fooled by deepfake technology, with one incident involving a digitally recreated version of its chief financial officer ordering money transfers in a video conference call. Everyone present on the video call, except the victim, was a fake representation of real people. The scammers applied deepfake technology to turn publicly available video and other footage into convincing versions of the meeting's participants.[4]

Political propaganda and disinformation. Deepfake technology is being used around the world to spread disinformation and propaganda. 2024 is a major election year in democracies around the globe: at least 64 countries will hold elections, representing close to 49% of the world's population.[5] It is also likely to be the first of many election years in which AI plays a pivotal role, as the technology becomes more widely available and easier to use. This has already been observed in Slovakia, where deepfake audio influenced an election in 2023:

> Days before a pivotal election in Slovakia to determine who would lead the country, a damning audio recording spread online in which one of the top candidates seemingly boasted about how he'd rigged the election. And if that wasn't bad enough, his voice could be heard on another recording talking about raising the cost of beer. The recordings immediately went viral on social media, and the candidate, who is pro-NATO and aligned with Western interests, was defeated in September by an opponent who supported closer ties to Moscow and Russian President Vladimir Putin.[6]

Deepfakes are not only being deployed by third parties; they can be used by the candidates themselves, either to improve their own self-images or to detract from their opponents. In mid-2023, former Republican presidential candidate Governor Ron DeSantis used AI to add fighter jets to one of his campaign videos.[7] Around the same time, Governor DeSantis' super PAC

https://www.dailymail.co.uk/news/article-13295475/Inappropriate-images-California-Fairfax-High-School-AI-deepfake.html

[3] Emily Flitter and Stacy Cowley, "Voice Deepfakes Are Coming for Your Bank Balance," *New York Times*, Aug. 30, 2023, www.nytimes.com/2023/08/30/business/voice-deepfakes-bank-scams.html.

[4] Harvey Kong, "'Everyone looked real': multinational firm's Hong Kong office loses HK$200 million after scammers stage deepfake video meeting," *South China Morning Post*, Feb. 4, 2024, www.scmp.com/news/hong-kong/law-and-crime/article/3250851/everyone-looked-real-multinational-firms-hong-kong-office-loses-hk200-million-after-scammers-stage.

[5] Koh Ewe, "The Ultimate Election Year: All the Elections Around the World in 2024," *Time*, Dec. 28, 2023, https://time.com/6550920/world-elections-2024/.

[6] Curt Devine, Donie O'Sullivan, Sean Lyngass, "A fake recording of a candidate saying he'd rigged the election went viral. Experts say it's only the beginning," *CNN*, Feb. 1, 2024, www.cnn.com/2024/02/01/politics/election-deepfake-threats-invs/index.html.

[7] Ana Faguy, "New DeSantis Ad Superimposes Fighter Jets In AI-Altered Video Of Speech," *Forbes*, May 25, 2023, www.forbes.com/sites/anafaguy/2023/05/25/new-desantis-ad-superimposes-fighter-jets-in-ai-altered-video-of-speech/.

released an ad containing an AI-generated speech by former president Donald Trump.[8] The Republican National Committee also released a 30-second ad that displayed images of disorder and destruction, with a voiceover that described the "consequences" of re-electing President Biden.[9] None of the images in this ad were real.

3) **Content provenance.** Many of the issues associated with deepfakes could be resolved, if only there were a reliable way to identify GenAI content. While at present no single solution exists, there are ongoing efforts to embed information related to "content provenance" – the verifiable history of a piece of content – into both GenAI products and the products of real-life recorders, such as digital cameras. Under this framework, the users of social media platforms would be able to rely on provenance data to identify trustworthy content.

Metadata. The Merriam-Webster Dictionary defines metadata as "data that provides information about other data." In practice, metadata is a structured set of descriptors attached to digital content. A photograph's metadata might include information about the camera used to take the photo, the time and date the photo was taken, and the photo's precise geolocation. A written document's metadata may include details about its author, its creation date, and the number of times the document has been edited. Metadata can be used to verify content authenticity, helping to combat fake news by providing a traceable history of content creation and modification. But metadata can also contain personal information about the individual who creates or modifies a piece of content.

Watermarking. Watermarking is the process of embedding an identifiable marker into digital content. This "watermark" can be visible, such as a logo or text overlay, or it can be invisible, data embedded into a file in a manner that does not noticeably alter the content. As a technology, watermarking is in its infancy. Watermarks can be stripped from content relatively easily by common screenshotting tools, file compression software, and image editing programs like Photoshop. They can also be faked by treating a GenAI system like a copy machine: a real image or video can be fed into a GenAI system and spit back out, unaltered except for the addition of a watermark. The system's user receives a piece of authentic content that has been incorrectly marked as inauthentic, ready to be posted online in order to create confusion and sow discord. Furthermore, while progress has been made towards developing standards for watermarking of images and video, how audio and text should be watermarked is far less clear.

Coalition for Content Provenance and Authenticity. The Coalition for Content Provenance and Authenticity (C2PA) is a tech initiative supported by several major companies, including Adobe, Microsoft, Intel, and Google, that seeks to establish a universal standard for the authentication of digital content:

> The goal of the C2PA specifications is to tackle the extraordinary challenge of trusting media in a context of rapidly evolving technology and the democratization of powerful creation and editing techniques. To this end, the specifications are designed to enable global, opt-in, adoption of digital provenance techniques through the creation of a rich ecosystem of digital provenance enabled applications for a wide range of individuals and organizations while

[8] Alex Isenstadt, "DeSantis PAC uses AI-generated Trump voice in ad attacking ex-president," *Politico*, Jul. 17, 2023, www.politico.com/news/2023/07/17/desantis-pac-ai-generated-trump-in-ad-00106695.
[9] GOP, "Beat Biden," Apr. 25, 2023, https://www.youtube.com/watch?v=kLMMxgtxQ1Y.

> meeting appropriate security and privacy requirements, as well as human rights considerations.[10]

4) **What this bill would do.** How should a social media platform contend with provenance data embedded into user-uploaded content? If the priority of the platform is to guarantee user privacy, the platform should strip provenance data before displaying content to users. However, if the platform's priority is to help users identify deepfakes and other GenAI products, then the platform should preserve provenance data in uploaded content. In reality both goals are important, and it is not immediately obvious which a platform should prioritize.

AB 1791 would give platforms clear instructions for dealing with provenance data in user-uploaded content. This bill defines "personal provenance data" to mean provenance data related to a user's identity, and "system provenance data" to mean provenance data related to the type of system used to generate content. The bill requires social media platforms to strip personal provenance data from content uploaded to a platform, while retaining system provenance data. If it is not possible to do both, the bill requires social media platforms to strip both types of provenance data from content, and to attach a label containing the stripped system provenance data to the content. The bill describes several characteristics the labels are required to have in order for them to persist within and between social media platforms.

5) **Author's statement.** According to the author:

> AI generated images and video are becoming more easily accessible and convincing every day. There are serious consequences to deepfakes from our political dialogue, to rattling the stock market, and fraud. It is why being able to authenticate digital content is incredibly important. AB 1791 is allowing users who decide to opt-in to add transparency to their content, will not be removed by the platform.

6) **Analysis.** A prior version of this bill simply stated that "a social media platform shall not remove digital content provenance verification from content uploaded to the social media platform by a user." Several organizations submitted letters in response to that version of the bill. It is worth assessing whether the current version of AB 1791 addresses the concerns raised in those letters.

Two of these letters point out that the prior version of this bill does not sufficiently consider user privacy. Taking an "oppose unless amended" position, Electric Frontier Foundation writes:

> In codifying this technology under A.B. 1791, we respectfully suggest that the definition of "digital content provenance" is more finely tuned to avoid affecting measures that protect the privacy of consumers uploading content. EFF is particularly concerned about A.B. 1791 explicitly prohibiting a social media platform from stripping metadata for privacy reasons when content is uploaded by a user.

Similarly, a coalition letter penned by TechNet writes:

[10] C2PA, "C2PA Explainer," https://c2pa.org/specifications/specifications/1.2/explainer/Explainer.html#_what_is_redaction_and_how_does_it_work

> Finally, we want to highlight that the bill could be interpreted to prevent social media platforms from taking privacy protective actions like removing location and personal information metadata from content, which are types of watermarks. We believe this type of action should be encouraged and could be resolved with an amendment to the effect of "except where necessary to protect the safety or privacy of the user."

The current version of AB 1791 specifically requires social media companies to remove content provenance data related to a user's personal information, thus addressing this privacy concern.

The coalition letter also raised concerns related to the bill's definition of "content provenance." Previously, this bill had defined "content provenance" to mean "the verifiable chronology of the original piece of digital content." The coalition letter states:

> Our members have some concerns with the definition of "digital content provenance" and have a suggested amendment to provide some clarity. We believe the current definition should be adjusted to apply more accurately to content credentials, watermarks, and other content provenance methods. The current definition seems to be limited to content credentials because it references a chronology. Watermarking for example doesn't preserve a chronology, it embeds information directly into the content itself. We think broadening the definition will provide greater incentive to companies to use the most appropriate technology for their particular use case. *"Digital content provenance" means information embedded into the outputs or their metadata created by artificial intelligence for the purposes of verifying its authenticity or origination."*

The current version of this bill aligns its definition for "provenance data" closely with the definition provided in the coalition letter: *"Provenance data" means data that is embedded into digital content, or that is included in the digital content's metadata, for the purpose of verifying the digital content's authenticity, origin, or history of modification.* However, the bill's definition is agnostic as to whether the digital content in question is generated by AI. As a result, the definition provided by the bill is more technology-neutral – for example, it applies to provenance information that might be embedded in a digital camera's outputs.

Finally, the coalition letter recommends exempting situations where provenance information is accidently stripped from content:

> We would also suggest adding "intentionally" or "knowingly" to section 21761. There may be some situations where despite a provider's normal practices a bug or glitch could result in a loss of provenance information. Providers should design their systems so that they don't strip this information and try to preserve it, but it should not be a strict liability offense if it happens.

If the author chooses to adopt this language, they should consider making this exemption specific to situations where system provenance data is accidently stripped, and excluding it from situations where personal provenance data is accidently retained. In doing so, they can avoid weakening the bill's privacy protections.

Enforcement. As recently amended, the bill grants a private right of action to users of platforms for violations of the bill's provisions. While a private right of action is a powerful remedy, it is often the case that the Legislature prefers to entrust enforcement authority of certain rights to civil public prosecutors, including the Attorney General, city attorneys, and county counsels. In discussions with the author's office, it is understood that there is room to revisit this issue to determine how best to enforce the bill's provisions.

Relationship to AB 3211. AB 3211 (Wicks, 2024) requires generative AI and digital camera developers to include provenance information in their systems' outputs. AB 3211 also requires "large online platforms," of which social media platforms are a subset, to use labels to prominently disclose provenance data. AB 3211 does not describe how provenance data related to a user's personal information should be treated. As a result, AB 3211 and AB 1791 appear to be compatible.

The nexus between provenance data and speech. Section 230 of the Communications Decency Act of 1996 states that "[n]o provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider."[11] That section also provides that "[n]o cause of action may be brought and no liability may be imposed under any State or local law that is inconsistent with this section."[12] Because this bill's requirements pertain to data surrounding the speech act but not the message itself, it appears that the bill does not treat a platform "as the publisher or speaker," nor hold it liable, for such content. Section 230 is thus not plainly implicated by this bill.

Nor does the bill appear to infringe on freedom of speech, as guaranteed by the United States and California Constitutions.[13] Free speech is implicated when a law has the effect of burdening expression—that is the essential message that a person is attempting to convey. Regulations on information incidental to the expressive value of content—including metadata about the time, place, and means by which the message was conveyed—do not appear to directly implicate First Amendment concerns.

Nevertheless, these are emerging issues that may warrant further exploration going forward.

7) **Committee amendments.** Several proposed committee amendments would clarify the provisions of this bill without substantively changing its content. The updated text is reproduced below in full:

21760. As used in this chapter:

(a) "Personal provenance data" means provenance data that contains ~~personal information, as defined in Section 1798.140 of the Civil Code.~~ ***any of the following:***

(1) Personal information, as defined by Section 1798.140 of the Civil Code.

(2) Unique device, system, or service information that is reasonably capable of being associated with a particular user.

(3) Time-of-day information.

(b) "Provenance data" means data that is embedded into digital content, or that is included in the digital content's metadata, for the purpose of verifying the digital content's authenticity,

[11] 47 U.S.C. § 230(c)(1).

[12] *Id.* at (e)(3).

[13] U.S. Const., 1st and 14th Amends; Cal. Const. art. I, § 2.

origin, or history of modification~~, including, but not limited to, personal provenance data and system provenance data~~.

(c) “Social media platform” has the same meaning as defined in Section 22589.

(d) “System provenance data” means provenance data ***that is not reasonably capable of being associated with a particular user and*** that ~~does~~ ***contains*** either of the following:

(1) ~~Identifies the~~ ***Information regarding the type of*** device, system, or service that was used to generate a piece of digital content.

(2) ~~Authenticates the digital content.~~ ***Information that provides proof of content authenticity.***

21761. (a) A social media platform shall redact personal provenance data from content uploaded to the social media platform by a user.

(b) Except as provided in subdivision (c), a social media platform shall not redact system provenance data from content uploaded to the social media platform by a user.

(c) ~~(1)~~ If a social media platform is unable to redact personal provenance data from content without also redacting system provenance data from the content, a social media platform shall redact the personal provenance data and the system provenance data from the content and append a label to the content that meets ~~both~~ ***all*** of the following criteria:

~~(A)~~ ***(1)*** The label prominently discloses any system provenance data that was redacted by the social media platform ~~and does not contain personal provenance data~~.

~~(B)~~ ***(2)*** The label remains appended to the content even if the content is shared, reposted, or otherwise replicated within the social media platform.

(3) The label does not disclose any personal provenance data.

(d) When content to which a social media platform has appended a label pursuant to subdivision (c) is downloaded, shared to an external internet website, or otherwise distributed in a manner that does not permit the social media platform to control how the content is displayed, the social media platform shall embed the ~~system provenance data~~ ***information contained in the label*** into the distributed content or add it to the content’s metadata.

(e) A social media platform shall abide by relevant industry standards to the greatest extent possible when redacting provenance data, labeling content, embedding ~~system provenance data~~ ***information*** into content, or adding ~~system provenance data~~ ***information*** to content metadata pursuant to this section.

21762. ~~(a) A user of a social media platform may bring a civil action against a social media platform that fails to comply with this chapter with regard to content uploaded by that user for either of the following relief:~~

(a) A social media platform that violates this chapter shall be liable in a civil action brought by a user of the social media platform for all of the following:

(1) Actual damages or ***statutory damages of not more than*** ten thousand dollars ($10,000) per violation, whichever is greater.

(2) Injunctive relief.

(3) Reasonable attorney's fees and costs.

~~(b) In an action pursuant to this section, the court shall award reasonable attorney's fees and costs to a prevailing plaintiff.~~

8) Related legislation.

AB 3050 (Low, 2024) would require CDT to issue regulations to establish standards for watermarks to be included in covered AI-generated material. The bill is pending in this Committee.

AB 3211 (Wicks, 2024) would require generative AI and digital camera developers to embed content provenance data into their products' outputs, and would require social media platforms to prominently display labels containing this provenance data. The bill is pending in Assembly Appropriations Committee.

REGISTERED SUPPORT / OPPOSITION:

Oppose Unless Amended

Electronic Frontier Foundation

Opposition

None on file.

Analysis Prepared by: Slater Sharp / P. & C.P. / (916) 319-2200