

Date of Hearing: April 10, 2024

ASSEMBLY COMMITTEE ON ELECTIONS
Gail Pellerin, Chair
AB 2655 (Berman) – As Amended April 1, 2024

SUBJECT: Elections: deceptive audio or visual media.

SUMMARY: Requires large online platforms, as defined, to block the posting or sending of materially deceptive and digitally modified or created content related to elections, or to label that content, during specified periods before and after an election. Specifically, **this bill:**

- 1) Requires a large online platform (platform), using state-of-the-art, best available tools to detect digitally modified or created content, to develop procedures for blocking and preventing the posting or sending of materially deceptive and digitally modified or created content, and to block and prevent that content if the platform knows or should know that the content meets the following requirements, during a specified time period, of any of the following:
 - a) A candidate portrayed as doing or saying something that the candidate did not do or say.
 - b) An elections official portrayed as doing or saying something in connection with the performance of their elections-related duties that the official did not do or say.
 - c) An elected official portrayed as doing or saying something that influences the election that the elected official did not do or say.
 - d) A voting machine, ballot, voting site, or other property or equipment related to an election that is portrayed in a materially false way.
- 2) Prohibits a platform from preventing a candidate, notwithstanding the prohibition detailed above, from portraying themselves as doing or saying something that the candidate did not do or say if the digital content includes a disclaimer stating: “This (image/video/audio) has been manipulated.” Requires this disclaimer to comply with the following:
 - a) In the case of visual media, requires the text of the disclaimer to appear in a size that is easily readable, as specified.
 - b) In the case of a video, requires the disclaimer to appear for the duration of the video.
 - c) In the case of media that consists of audio only, requires the disclaimer to be read in a manner that can be easily heard by the average listener at both the beginning and the end of the audio. For audio that is longer than two minutes, the disclaimer must also be included during the audio at intervals of not more than two minutes each.
- 3) Provides that the provisions detailed above apply only during the following time periods:
 - a) Beginning 120 days before any election through the day of the election.

- b) With respect to content pertaining to elections officials, or that depicts or pertains to elections equipment and materials, beginning 120 days before the election and ending on the 60th day after the election.
- 4) Requires a platform, using state-of-the-art, best available tools to detect digitally modified or created content, to develop procedures for labeling materially deceptive and digitally modified or created content that pertains to election processes and that is not subject to the blocking provisions outlined above, and to label such content as inauthentic, fake, or false if the platform knows or should know that the digitally modified or created content meets the requirements of this bill.
- a) Requires the label to permit users to click or tap on it and to inspect all available provenance data about the digitally modified or created content in an easy-to-understand format.
 - b) Provides that the labeling requirement detailed above applies only during the following time periods:
 - i) The period beginning one year before the election and through the day of the election that is specified in or implicated by the content.
 - ii) The period beginning one year before the election process and through the final day of the election process that is specified in or implicated by the content.
 - iii) If the content depicts or pertains to elections officials, the period beginning one year before the election or election process that is specified in or implicated by the content and ending on the 60th day after that election or the 60th day after the final day of that election process, as applicable.
- 5) Requires a platform to provide a way for Californians to report content that was not blocked or labeled as required. Requires the platform to respond within 36 hours and to describe any action taken.
- 6) Permits a Californian who reported content and who does not receive a response in 36 hours or who disagrees with the response, the Attorney General (AG), a district attorney, or a city attorney to seek injunctive or other equitable relief against a platform to compel compliance with this bill. Requires the court to award a prevailing plaintiff reasonable attorney's fees and costs. Provides that such an action is entitled to precedence in court, as specified.
- 7) Requires any label or disclaimer that must appear on content under this bill to appear in English, and in the same language as the content if the content isn't in English.
- 8) Requires a platform to maintain a copy of any content that it blocks or labels under this bill for at least five years from the election or election process specified or implicated in the content. Requires the platform to make that content available to the Secretary of State, the Fair Political Practices Commission, and researchers, if requested.
- 9) Provides that this bill does not apply to any of the following:

- a) A regularly published online newspaper, magazine, or other periodical of general circulation that routinely carries news and commentary of general interest, and that publishes any materially deceptive and digitally altered or digitally created image, audio, or video recording that an online platform is required to block or label by this bill, if the publication contains a clear disclosure that the materially deceptive and digitally altered or digitally created image or audio or video recording does not accurately represent any actual event, occurrence, appearance, speech, or expressive conduct.
- b) Materially deceptive audio or visual media that constitutes satire or parody.

10) Defines the following terms, for the purposes of this bill:

- a) “Election processes” to mean any government process related to an election, including, but not limited to, elections, candidates, vote counting, redistricting, and proceedings or processes of the electoral college.
- b) “Materially deceptive and digitally modified or created content” to mean an image or an audio or video recording or other digital content, including a chatbot, that has been intentionally manipulated such that all of the following conditions are met:
 - i) The digital content is the product of digital manipulation, artificial intelligence (AI), or machine learning, including deep learning techniques, that merges, combines, replaces, or superimposes content onto an image or an audio or video recording, creating an image or an audio or video recording that appears authentic, or that otherwise generates an inauthentic image or an audio or video recording that appears authentic.
 - ii) The content contains a false portrayal of a candidate for elective office, an elected official, an elections official, or a voting machine, ballot, voting site, other property or equipment related to an election, or elections process. Provides that a “false portrayal” means the content would cause a reasonable person to have a fundamentally different understanding or impression of the content than the person would have if they were hearing or seeing the unaltered, original version of the content.
 - iii) The person or entity who attempted to post or send, or who did post or send, the content did so knowing the portrayal was false, or did so with reckless disregard for whether the portrayal was false. Provides that if the content is intentionally manipulated and contains a false portrayal, there is a rebuttable presumption that the person or entity knew the portrayal was false or that they acted with reckless disregard for whether the portrayal was false.
- c) “Large online platform” to mean a public-facing internet website, web application, or digital application, including a social network, video sharing platform, advertising network, or search engine that had at least 1 million California users during the preceding 12 months.

- 11) Provides that content that contains only minor modifications that do not lead to significant changes to the perceived contents or meaning of the content are not “materially deceptive and digitally modified or created content” for the purposes of this bill, as specified.
- 12) Contains various findings and declarations and contains a severability clause.

EXISTING STATE LAW:

- 1) Prohibits a person, committee, or other entity, until January 1, 2027, from distributing with actual malice, within 60 days of an election at which a candidate for elective office will appear on the ballot, materially deceptive audio or visual media of a candidate with the intent to injure the candidate’s reputation or to deceive a voter into voting for or against the candidate.
 - a) Defines “materially deceptive audio or visual media,” for these purposes, as an image or an audio or visual recording of a candidate’s appearance, speech or conduct that has been intentionally manipulated in a manner that both of the following are true about the image or audio or video recording:
 - i) It would falsely appear to a reasonable person to be authentic; and,
 - ii) It would cause a reasonable person to have a fundamentally different understanding or impression of the expressive content of the image or audio or video recording than the person would have if the person were hearing or seeing the unaltered, original version of the image or audio or video recording.
 - b) Provides that this prohibition does not apply if the audio or visual media includes a disclaimer stating “This (image/video/audio) has been manipulated,” and the disclaimer complies with specified requirements.
 - c) Permits a candidate whose voice or likeness appears in deceptive audio or visual media distributed in violation of this provision to seek the following relief:
 - i) Injunctive or other equitable relief prohibiting the distribution of the materially deceptive audio or visual media in violation of this bill. Provides that such an action is entitled to precedence in court, as specified.
 - ii) General or special damages against the person, committee, or other entity that distributed that audio or visual media. Permits the court to award reasonable attorney’s fees and costs to a prevailing party in such an action.
 - d) Provides that in any civil action brought pursuant to these provisions, the plaintiff bears the burden of establishing the violation through clear and convincing evidence.
 - e) Provides that this prohibition shall not be construed to alter or negate any rights, obligations, or immunities of an interactive service provider under Section 230 of the federal Communications Decency Act.
 - f) Provides that this prohibition does not apply to any of the following:

- i) A radio or television broadcasting station, as specified, in either of the following circumstances:
 - (1) When it broadcasts materially deceptive audio or visual media as part of a bona fide newscast, news interview, news documentary, or on-the-spot coverage of bona fide news events, if the broadcast clearly acknowledges through content or disclosure that there are questions about the authenticity of the audio or visual media, as specified.
 - (2) When it is paid to broadcast materially deceptive audio or visual media.
 - ii) An internet website, or a regularly published newspaper, magazine, or other periodical of general circulation, including an internet or electronic publication, that routinely carries news and commentary of general interest, and that publishes materially deceptive audio or visual media covered by this prohibition, if the publication clearly states that the media does not accurately represent the speech or conduct of the candidate.
 - iii) Materially deceptive audio or visual media that constitute satire or parody. (Elections Code §20010, as amended by Section 3 of Chapter 745 of the Statutes of 2022)
- 2) Prohibits a person, firm, association, corporation, campaign committee, or organization, beginning January 1, 2027, with actual malice, from producing, distributing, publishing, or broadcasting campaign material, as defined, that contains either of the following types of pictures or photographs, as specified, unless the campaign material includes a disclaimer that the picture is not an accurate representation of fact:
- a) A picture or photograph of a person or persons into which the image of a candidate for public office is superimposed.
 - b) A picture or photograph of a candidate for public office into which the image of another person or persons is superimposed. (Elections Code §20010, as amended by Section 4 of Chapter 745 of the Statutes of 2022)

EXISTING FEDERAL LAW provides, pursuant to Section 230 of the federal Communications Decency Act, that no 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. (47 U.S.C. §230)

FISCAL EFFECT: Unknown

COMMENTS:

- 1) **Purpose of the Bill:** According to the author:

AB 2655 will ensure that online platforms restrict the spread of election-related deceptive deepfakes meant to prevent voters from voting or to deceive them based on fraudulent content. Deepfakes are a powerful and dangerous tool in the arsenal of those that want to wage disinformation campaigns, and they have the potential

to wreak havoc on our democracy by attributing speech and conduct to a person that is false or that never happened. Advances in AI make it easy for practically anyone to generate this deceptive content, making it that much more important that we identify and restrict its spread before it has the chance to deceive voters and undermine our democracy.

- 2) **Threats of Manipulated Media in Campaign Communications:** The use of false and deceptive information in campaigns to influence election outcomes is not a new phenomenon. Laws aimed at curbing such practices and preserving the integrity of elections have a long history in California. In 1850, the First Session of the California State Legislature created penalties for election misconduct, including for “deceiving [an elector] and causing him to vote for a different person for any office than such elector desired or intended to vote for” (Chapter 38, Statutes of 1850). California law today includes various provisions criminalizing deceptive tactics that undermine election integrity or interfere with voters’ ability to participate in elections. This includes laws that prohibit distribution of false and misleading information about qualifications to vote or about the days, dates, times, and places where voting may occur (Elections Code §18302); prohibit the misleading use of government seals in campaign literature (Elections Code §18304); and prohibit coercing or deceiving people into voting in a way that was inconsistent with the person’s intent (Elections Code §§18573, 18573.5).

Advancements in technology have made it increasingly simple to produce false and misleading media that closely resembles authentic content. Moreover, platforms like social media have facilitated the rapid dissemination of deceptive media to large audiences at minimal cost. Given these developments, the potential threat posed by manipulated media to future elections’ integrity may be more significant than in the past.

As described in greater detail below, past legislative efforts have addressed concerns about manipulated media’s use to deceive voters during elections. Those laws, however, are limited, and are designed primarily to target the harms to *candidates* that may result from the distribution of manipulated media of those candidates. In contrast, this bill aims to regulate materially deceptive and digitally altered media depicting not only candidates, but also elections officials and elected officials who are not candidates. Additionally, this bill targets media that portrays elections materials and equipment in materially deceptive ways. The author and supporters of this bill believe that these provisions will safeguard voters against deceitful media that could undermine trust in the electoral process.

- 3) **Recent Examples of Materially Deceptive Campaign Communications:** As evidence of the need for this bill, the author points to the following incidents, as reported in the media:
- AI tools were used to create deepfake video ads of British Prime Minister Sunak on Facebook.
 - A Chinese disinformation campaign in Taiwanese elections made use of deepfakes, and AI-generated videos, images and audio clips.
 - AI tools additionally were used to disrupt elections in Argentina, Bangladesh, Pakistan, Slovakia.

- Governor Ron DeSantis created a deepfake of former President Trump hugging Anthony Fauci.
 - AI was used by a democratic operative and a magician to generate deepfake audio of President Biden's voice that was used in a robocalls to dissuade voters from voting in the primary election.
- 4) **Previous Legislation Related to Materially Deceptive Media in Campaigns:** In 2019, in response to concerns that deepfake technology could be used to spread misinformation in political campaigns, the Legislature approved and Governor Newsom signed AB 730 (Berman), Chapter 493, Statutes of 2019. Deepfake technology refers to software capable of producing a realistic looking video of someone saying or doing something that they did not, in fact, say or do. This technology has advanced rapidly in recent years thanks to the use of AI to help train the software.

AB 730 prohibits the distribution of materially deceptive audio or visual media with actual malice with the intent to injure a candidate's reputation or to deceive a voter into voting for or against a candidate, unless the materially deceptive audio or visual media includes a disclaimer that it has been manipulated. AB 730 does not apply exclusively to deepfakes, but rather applies to any intentional manipulation of audio or visual images that results in a version that a reasonable observer would believe to be authentic. Nonetheless, the increasing availability and advancing capability of deepfake technology was the immediate impetus for that bill.

AB 730 was designed as an update to California's "Truth in Political Advertising Act," a law enacted in 1998 (through the passage of AB 1233 (Leach), Chapter 718, Statutes of 1998) that prohibited campaign material that contains a picture of a person into which a candidate's image is superimposed, or contains a picture of a candidate into which another person's image is superimposed, except if a specified disclaimer was included. The Truth in Political Advertising Act was introduced in response to the use of photoshopped pictures in campaign materials, and accordingly was designed to target the manipulation of photographs in campaign materials. In the 20 years following its passage, however, it was never amended to update the law to address more modern techniques of manipulating campaign materials in a manner that can mislead voters. AB 730 replaced the Truth in Political Advertising Act with a law that regulates not only altered photographs in campaign materials, but also audio and video media that have been altered in a materially deceptive manner.

AB 730 included a January 1, 2023 sunset date. In 2022, however, the Legislature approved AB 972 (Berman), Chapter 745, Statutes of 2022, which extended the sunset date to January 1, 2027. AB 972 did not otherwise change the provisions of AB 730. If the current January 1, 2027 sunset date is not repealed or extended, the original Truth in Political Advertising Act as enacted by AB 1233 of 1998 would go back into effect.

Because the impetus for AB 730 was concern about the potential that people might create deepfake media appearing to be accurate representations of the conduct of candidates for office, its provisions apply exclusively to images or audio or video recordings of a candidate's appearance, speech, or conduct. Relatedly, candidates for elective office who are the target of materially deceptive media are the only entities that can seek injunctive relief or damages under AB 730. Materially deceptive images, audio, or video that appear in

campaign communications are not covered by AB 730 if that media is not of a candidate. For instance, if a candidate digitally manipulated video or a photo of a campaign rally to make the crowd look significantly larger than it actually was, such manipulation would not be covered by AB 730 as long as the manipulated image or video was not materially deceptive about a candidate's appearance, speech, or conduct. Similarly, manipulated and materially deceptive content in advertisements related to ballot measures, or in communications that seek to undermine confidence in the electoral process but that do not mention candidates directly, generally would not be covered by AB 730.

- 5) **Online Platforms and Materially Deceptive Content:** This is one of two bills being considered by the committee today that seek to address materially deceptive and digitally altered elections-related content in an effort to protect the integrity of elections in California. While the related bill that is being considered today (AB 2839 (Pellerin)) applies broadly to the distribution of such content through various mediums, this bill specifically targets the distribution of deceptive content through online platforms, including social media. Recognizing that those online platforms can facilitate the rapid spread of deceptive content, this bill seeks to minimize that potential by obligating large online platforms to block or label offending content. In order to do that, the platforms necessarily will need to be able to identify the content that must be blocked or labeled.

In recognition that the regulation of the distribution of content can create free speech concerns, this bill contains various provisions that tailor the content to which it applies, such that it targets content that has the highest likelihood of deceiving voters and undermining electoral integrity. While that tailoring does limit the content that online platforms would be required to block or label, it also adds additional factors that platforms must consider in order to identify content that is required to be blocked or labeled under this bill.

Along with other limitations, this bill applies only to content that (1) is distributed during specified time periods around elections and election processes, (2) includes media relating to elections or the electoral process in specified ways, (3) that was intentionally manipulated digitally to be materially deceptive, and (4) that is not satire or parody. Each of these limitations adds additional factors that online platforms would need to consider when determining whether a specific communication must be blocked or labeled by this bill.

For example, in order to determine whether it must block content that *portrays a candidate for election as doing or saying something that the candidate did not do or say*, the platform would need to know not only that the person portrayed in the content was a candidate for office, but also the date (or dates) of the election when the candidate will appear on the ballot. Similarly, it would need to determine whether the candidate had actually said or done the thing that the candidate is portrayed as doing. While some of that information will be widely available and well known in some cases (e.g., the identity of major party candidates for President of the United States in presidential general elections and the dates of federal elections), it will be more arcane in other situations. Given the number of elections (including standalone local and special elections) and candidates (including write-in candidates and candidates for local elections in smaller jurisdictions) in California at any given time, making the determinations at scale about which content must be blocked or labeled likely will be considerably more challenging than making those determinations on a case-by-case basis in a court of law.

This bill includes numerous provisions that recognize that platforms will face challenges in making some of these determinations, and in limiting those platforms' obligations and liabilities accordingly. For instance, recent amendments to this bill provide for the large online platforms to use "state-of-the-art, best available tools" for detecting digitally modified or created content, recognizing that the identification of such content with perfect accuracy is impossible. Other recent amendments provide that an online platform is obligated to block or label content only if the platform "knows or should know" that the content meets the requirements of the bill, and provide that the bill's requirements relating to content portraying an elections official applies only if the platform knows or should know that the person is an elections official. Furthermore, unlike the related legislation that targets materially deceptive election content more broadly, this bill does not provide for damages to be awarded against platforms that fail to comply with their obligations. Instead, the only legal relief available under this bill is injunctive relief, and a person (other than the AG, a district attorney, or a city attorney) would be required to report the content to the platform and give the platform an opportunity to block or label the content before the person could seek that relief. Notwithstanding these provisions, the extent to which large online platforms will be able to accurately block and label materially deceptive elections-related content at scale, as contemplated by this bill, is unclear.

- 6) **Free Speech Considerations:** The First Amendment to the United States (US) Constitution, which also applies to states under the Fourteenth Amendment, provides in relevant part "Congress shall make no law...abridging the freedom of speech..." Similarly, Section 2 of Article I of the California Constitution provides in relevant part "Every person may freely speak, write, and publish his or her sentiments on all subjects, being responsible for the abuse of this right. A law may not restrain or abridge liberty of speech or press."

This bill seeks to regulate the distribution by online platforms of media containing intentionally manipulated images, audio, or video related to candidates, elections officials, elected officials, and election materials and equipment under certain circumstances. A question could be raised about whether this bill is consistent with the right to freedom of speech that is guaranteed by the US and California constitutions. The US Supreme Court has ruled that even false statements are protected by the First Amendment (*United States v. Alvarez* (2012), 567 U.S. 709). When a law burdens core political speech, the restrictions on speech generally must be "narrowly tailored to serve an overriding state interest," *McIntyre v. Ohio Elections Commission* (1995), 514 US 334.

This bill targets deceptive content that could undermine trust in elections, prevent voters from voting, and distort the electoral process. The US Supreme Court generally has found that the protection of the integrity of elections is an overriding (or compelling) government interest (*Id.* at 349; *Burson v. Freeman* (1992) 504 U.S. 191, 199). A challenge of this bill on First Amendment grounds, then, likely would hinge on whether the court found this bill's provisions to be narrowly tailored.

As discussed in more detail above, this bill includes provisions to limit its scope to communications posing the greatest threat to election integrity. Whether these limitations adequately protect this bill from a potential constitutional challenge is a question that falls more squarely within the jurisdiction of the Assembly Judiciary Committee, where this bill will be heard next if it is approved by this committee. However, while these limitations may help protect the bill against a constitutional challenge, they may also make it harder for the

bill to achieve its aims of limiting the spread of materially deceptive communications that have the potential to undermine election integrity.

- 7) **Arguments in Support:** The sponsor of this bill, the California Initiative for Technology & Democracy, a Project of California Common Cause, writes in support:

Those trying to influence campaigns – conspiracy theorists, foreign states, online trolls, and candidates themselves – are already creating and distributing election-threatening deepfake images, audio, and video content in the US and around the world. This threat is not imaginary: generative AI has been used in various ways – most of them deeply deceptive – to influence the national elections in Slovakia, Bangladesh, Argentina, Pakistan, and elsewhere, including in our own country. Examples of this occurring in U.S. elections include Ron Desantis using AI-generated images to attack his opponent in his presidential run, foreign states caught attempting to influence American politics through social media, and just this month, a supporter of former President Trump creating a deepfake image of Trump with Black Americans designed to persuade Black voters to support Trump...

[AB] 2655 strikes the right balance by seeking to ban, for a strictly limited time before and after elections, the online spread of the worst deepfakes and disinformation maliciously intended to prevent voters from voting or getting them to vote erroneously based on fraudulent content. The bill also requires that other fake online content related to elections and elections processes (such as redistricting), which is also designed to undermine election procedures and democratic institutions, must be labeled as fake, again just for a limited time. The bill only applies to the largest online platforms with the greatest reach of potential election disinformation, and we believe it is fully implementable today based on tools these companies already possess. The companies covered by the bill's requirements are all already subject to similar requirements under the European Union's Digital Services Act, which is designed to, among other things, crack down on election interference.

AB 2655's approach is narrowly tailored and does not extend the law to hot button controversies or inflammatory claims – it does not ask social media platforms to adjudicate controversial opinions post by post. It simply stops the use of obviously, demonstrably untrue and provably false content meant to impermissibly influence our elections at peak election times.

- 8) **Arguments in Opposition:** A joint letter of opposition submitted by Chamber of Progress, Computer and Communications Industry Association, NetChoice, Software & Information Industry Association, and TechNet states:

Responsible digital services providers take aggressive steps to moderate dangerous and illegal content, consistent with their terms of service. The companies deliver on the commitments made to their user communities with a mix of automated tools and human review. In 2021, a number of online businesses announced that they had been voluntarily participating in the Digital Trust & Safety Partnership (DTSP) to develop and implement best practices to

ensure a safer and more trustworthy internet, and have recently reported on the efforts to implement these commitments.

AB 2655 appears to be based on the false assumption that online platforms definitively know whether any particular piece of content has been manipulated in such a way that is defined under the bill. While digital services may employ tools to identify and detect these materials with some degree of certainty, it is an evolving and imperfect science in its current form. AB 2655 also presumes that online platforms are an appropriate arbiter of deciding what constitutes accurate election information. However, most digital services are not equipped with the tools or expertise to make such judgments.

AB 2655 would impose significant operationally and practically challenging requirements on large online platforms who may not be best suited to achieve the bill's laudable and important goal of ensuring California's elections remain free and fair.

- 9) **Related Legislation:** AB 2355 (Wendy Carrillo), which is also being heard in this committee today, requires a political advertisement that is generated in whole or in part using AI to include a disclaimer stating that fact.

AB 2839 (Pellerin), which is also being heard in this committee today, prohibits the distribution of campaign advertisements and other election communications that are materially deceptive and digitally altered or created, except as specified.

- 10) **Double-Referral:** This bill has been double-referred to the Assembly Judiciary Committee.

REGISTERED SUPPORT / OPPOSITION:

Support

California Initiative for Technology & Democracy, a Project of California Common Cause
(Sponsor)

Asian Americans Advancing Justice - Asian Law Caucus

Asian Americans and Pacific Islanders for Civic Empowerment

Asian Law Alliance

California Clean Money Campaign

California State Sheriffs' Association

California Voter Foundation

Courage California

Disability Rights California

Indivisible CA Statestrong

Inland Empire United

The Partnership for the Advancement of New Americans

Verified Voting

Opposition

Chamber of Progress

Computer and Communications Industry Association

Electronic Frontier Foundation
NetChoice
Oakland Privacy (unless amended)
Software & Information Industry Association
TechNet

Analysis Prepared by: Ethan Jones / ELECTIONS / (916) 319-2094