Date of Hearing: June 14, 2022

ASSEMBLY COMMITTEE ON PRIVACY AND CONSUMER PROTECTION
Jesse Gabriel, Chair
SB 1216 (Gonzalez) – As Introduced February 17, 2022

SENATE VOTE: 39-0

SUBJECT: Secretary of the Government Operations Agency: working group: deepfakes

SUMMARY: This bill would, upon appropriation by the Legislature, require the Secretary of the Government Operations Agency (GovOps) to establish the Deepfake Working Group (DWG) to evaluate specified implications of the proliferation of deepfakes and digital content forgery technologies, as defined, for California's government agencies, businesses, and residents; and would require the DWG to report specified findings and recommendations to the Legislature by July 1, 2024. Specifically, **this bill would**:

1) Require the Secretary of GovOps, upon appropriation by the Legislature, to establish the DWG to evaluate all of the following:

   - The impact of proliferation of deepfakes on state government, California-based businesses, and residents of the state.

   - The risks, including privacy risks, associated with the deployment of digital content forgery technologies and deepfakes on state and local government, California based-businesses, and residents of the state.

   - The impact of digital content forgery technologies and deepfakes on civic engagement, including voters.

   - The legal implications associated with the use of digital content forgery technologies and deepfakes.

   - The best practices for preventing digital content forgery and deepfake technology to benefit the state, California-based businesses, and California residents.

2) Require the DWG to develop a coordinated plan to accomplish all of the following:

   - Reduce the proliferation and impact of digital content forgeries and deepfakes, including by exploring how the adoption of a digital content provenance standard could assist with reducing the proliferation of digital content forgeries and deepfakes.

   - Investigate the feasibility of, and obstacles to, developing standards and technologies for state departments for determining digital content provenance.

   - Increase the ability of internet companies, journalists, watchdog organizations, other relevant entities, and members of the public to meaningfully scrutinize and identify digital content forgeries and relay trust and information about digital content provenance to content consumers.

- Develop or identify mechanisms for content creators to cryptographically certify authenticity of original media and non-deceptive manipulations.

- Develop or identify mechanisms for content creators to enable the public to validate the authenticity of original media and non-deceptive manipulations to establish content provenance.

3) Require the DWG to report to the Legislature on the potential uses and risks of deepfake technology to the state government and California-based businesses on or before July 1, 2024; and require the report to include recommendations for modifications to the definition of digital content forgery and deepfakes and recommendations for amendments to other code sections that may be impacted by the deployment of digital content forgery technologies and deepfakes.

4) Specify that the DWG shall consist of participants from all of the following:

   - Three appointees from the technology industry, with technical focus that includes digital content, media manipulation, or related subjects.

   - Three appointees from nontechnology-related industries.

   - Three appointees with a background in law chosen in consultation with the Judicial Council.

   - Two appointees representing privacy organizations.

   - Two appointees representing consumer organizations.

   - The State Chief Information Officer, or the officer’s designee.

   - The Director of Finance or the director’s designee.

   - The chief information officers of three other state agencies, departments, or commissions, or their designees.

   - One member of the Senate, appointed by the Senate Committee on Rules.

   - One member of the Assembly, appointed by the Speaker of the Assembly.

5) Require the Secretary of GovOps to designate the chairperson of the DWG on or before July 1, 2023.

6) Specify that the working group shall take input from a broad range of stakeholders with a diverse range of interests affected by state policies governing emerging technologies, privacy, business, the courts, the legal community, and state government.

7) Specify that the working group shall serve without compensation, but shall be reimbursed for all necessary expenses actually incurred in the performance of their duties.

8) Define "digital content provenance" to mean the verifiable chronology of the original piece of digital content, such as an image, video, audio recording, or electronic document.

9) Define "digital content forgery" to mean the use of technologies, including artificial intelligence (AI) and machine learning (ML) techniques, to fabricate or manipulate audio, visual, or text content with the intent to mislead.

10) Specify that the report submitted pursuant to 3), above, be submitted in compliance with existing law pertaining to the submission of reports to the Legislature.

11) Provide that the provisions of the bill shall remain in effect only until January 1, 2025, and as of that date are repealed, unless a later enacted statute that is enacted before January 1, 2025 deletes or extends that date.

EXISTING LAW:

1) Establishes GovOps under the direction of an executive officer known as the secretary; and specifies that GovOps shall consist of all of the following: the Office of Administrative Law; the Public Employees' Retirement Systems; the State Teachers' Retirement System; the State Personnel Board; the California Victim Compensation Board; the Department of General Services; the Department of Technology; the Franchise Tax Board; the Department of Human Resources; and the California Department of Tax and Fee Administration. (Gov. Code Secs. 12800(a), 12801, and 12803.2(a).)

2) Provides that a depicted individual, as defined, has a cause of action against a person who does either of the following, except as specified, and may recover up to $150,000 in statutory damages, instead of or in addition to other available relief:

   - Creates and intentionally discloses sexually explicit material and the person knows or reasonably should have known the depicted individual in that material did not consent to its creation or disclosure.

   - Intentionally discloses sexually explicit material that the person did not create and the person knows the depicted individual in that material did not consent to the creation of the sexually explicit material. (Civ. Code Sec. 1708.86.)

3) Defines "depicted individual" for the purposes of 2), above, to mean an individual who appears, as a result of digitization, to be giving a performance they did not actually perform or to be performing in an altered depiction. (Civ. Code Sec. 1708.86(a)(4).)

4) Prohibits a person, committee, or other entity, within 60 days of an election at which a candidate for elective office will appear on the ballot, from distributing with actual malice materially deceptive audio or visual media, as defined, of the candidate with the intent to injure the candidate's reputation or to deceive a voter into voting for or against the candidate, unless a specified disclosure is included with the audio or visual media. (Elec. Code Sec. 20010.)

5) Defines "materially deceptive audio or visual media" for purposes of 4), above, to mean an image or an audio or video recording of a candidate's appearance, speech, or conduct that has

been intentionally manipulated in a manner such that both of the following conditions are met:

- The image or audio or video recording would falsely appear to a reasonable person to be authentic.

- The image or audio or video recording 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 that person would have if the person were hearing or seeing the unaltered, original version of the image or audio or video recording. (Elec. Code Sec. 20010(e).)

6) Requires that a report required or requested by law to be submitted to the Members of either house of the Legislature generally, instead be submitted as a printed copy to the Secretary of the Senate, as an electronic copy to the Chief Clerk of the Assembly, and as an electronic or printed copy to the Legislative Counsel, as specified. (Gov. Code Sec. 9795.)

7) Requires that a bill that would require a state agency to submit a report on any subject to either house of the Legislature generally, a committee or office of either house of the Legislature, or the Legislative Counsel Bureau, include a provision that repeals the reporting requirement, or makes the requirement inoperative, no later than a date four years following the date upon which the bill becomes operative, or four years after the due date of any report required every four or more years. (Gov. Code Sec. 10231.5(a).)

FISCAL EFFECT: According to the Senate Appropriations Committee, "[GovOps] anticipates total costs ranging from $514,884 to $600,348 for multiple, temporary, full-time positions to implement the program, support the working group, and for other operating expenses such as facilities and equipment."

COMMENTS:

1) Purpose of this bill: This bill seeks to combat potential detrimental effects of the increasing proliferation of deepfakes and digital content forgery technologies by soliciting the input of subject-matter experts in assessing policy solutions. This bill is sponsored by Adobe, Inc.

2) Author's statement: According to the author:

> Deepfakes are deceptive life-like videos and recordings that can effectively make it appear as though an individual said or did something that never actually took place. This type of manufactured media can have entertaining and innocent uses such as viral TikToks; or nefarious uses like the dissemination of forged sexually explicit material, or videos of influential political leaders that incite political violence.
>
> The potential of these digital forgeries is far reaching and will have implications for national security, influence on elections, and even how journalists and media sources verify the provenance of videos before they report them as factual news.
>
> This new frontier of technology has created a number of ethical, legal, and policy questions that are not easily answered and creates numerous complex implications for privacy rights, governmental communication, media accuracy, copyright infringement,

> and many other legal repercussions that can’t be easily addressed without thoughtful dialogue amongst informed stakeholders.
>
> This bill will allow for the exploration and examination of best practices being used to reduce digital content forgeries, help in identifying mechanisms to certify the authenticity of original content, and evaluate the impact of deepfakes throughout the state.

3) Deepfakes, generally: As AI and ML have become increasingly sophisticated, the capacity to manipulate audiovisual media to create realistic representations of fabricated events has grown exponentially. These so-called “deepfakes” can be harmless or even entertaining, but can also have serious social costs if used toward nefarious ends. Though the ability to manipulate digital media has been present since at least the 1990s, the realism made possible by these AI/ML-driven techniques, along with the ability to synthesize realistic media to represent nearly any occurrence, make these recent developments particularly insidious. In support of this bill, the Anti-Defamation League (ADL) explains:

> Deepfakes are an emerging technology that [] combines multiple real images/videos/audio with machine learning technology to create a new, synthetic piece of media (e.g. image, audio or video). This technique has been used to create machine-made media of all kinds, including some with the intention of deceiving audiences. Some examples of deceptive deepfakes include videos of politicians depicted in situations that never happened or fabricated pornographic videos targeting specific individuals. Audio deepfakes could lead to serious forms of fraud and identity theft.
>
> The proliferation of deepfakes and misinformation continue to increase at an alarming rate, and the public policy solutions needed to protect California residents, businesses, and government institutions remain unclear. Policy solutions continue to allude [*sic.*] policy makers across the globe.

Echoing the concerns raised by ADL, a recent report published by the United States Department of Homeland Security outlines several scenarios in which the use of deepfakes could be extremely dangerous, including: inciting violence; producing false evidence undermining scientific consensuses such as climate change and vaccine efficacy; falsifying evidence in a criminal case; corporate sabotage; social engineering attacks targeting corporate and financial institutions; stock manipulation; and cyberbullying.[1]

Since 2019, this Legislature has demonstrated recognition of the risks inherent to deepfake technology. Specifically, by passing AB 730 (Berman, Ch. 493, Stats. 2019) and AB 602 (Berman, Ch. 491, Stats. 2019), respectively, California has taken decisive action to address the sensitive circumstances in which deepfakes can be weaponized to influence political outcomes in elections or to manufacture explicit content that could harm the reputation of those falsely depicted.

Still, regulating the use of deepfakes in all but the most egregious of circumstances is extraordinarily complex. Doing so requires technical literacy with respect to the technology

[1] Department of Homeland Security, “Increasing Threats of Deepfake Identities”, https://www.dhs.gov/sites/default/files/publications/increasing_threats_of_deepfake_identities_0.pdf [as of Jun. 12, 2022].

itself, a sociocultural understanding of the media environment in which both problematic and legitimate uses may arise, and an in-depth understanding of the legal constraints surrounding such regulation, including potential impositions on First Amendment rights related to free expression.

Recognizing this complexity, SB 1216 seeks to provide the Legislature with recommendations from a panel of experts spanning various facets of the issues arising from deepfakes. Doing so has the potential to better inform future policy in order to effectively balance the myriad considerations thoughtful policymaking in this space must address.

4) SB 1216 would create the DWG, modelled after the Blockchain Working Group: In 2018, this Legislature passed, and the Governor signed into law, AB 2658 (Calderon, Ch**.** 875, Stats. 2018), which tasked the Secretary of GovOps with appointing a Blockchain Working Group to evaluate several facets of blockchain technology, including: uses of blockchain in state government and business; risks, including privacy risks, associated with the use of blockchain; benefits associated with the use of blockchain; legal implications associated with the use of blockchain; and best practices for enabling blockchain technology to benefit the state of California. AB 2658 also required the Blockchain Working Group to report to the Legislature on its findings, and on July 1, 2020, the Blockchain Working Group submitted its report, recommendations from which have generated several Legislative proposals in the years since. This bill is modelled after AB 2658, mirroring its structure and repurposing several of its provisions nearly verbatim.

SB 1216 would task GovOps with creating the DWG, made up of appointees from the technology industry, nontechnology-related industries, legal backgrounds, privacy organizations, consumer organizations, state agencies, and the Legislature. The bill would require the DWG to explore certain impacts and implications of the proliferation of digital content forgery technologies, as defined, and deepfakes, including: impacts on state government, California-based businesses, and residents of the state; risks, including privacy risks, associated with the deployment of digital content forgery technologies and deepfakes; impacts of digital content forgery technologies and deepfakes on civic engagement, including voters; legal implications associated with the use of digital content forgery technologies and deepfakes; and best practices for preventing digital content forgery and deepfake technology. The bill in print would also define the term "digital content forgery" to mean the use of technologies, including AI and ML techniques, to fabricate or manipulate audio, visual, or text content with the intent to mislead; and would define the term "digital content provenance" to mean the verifiable chronology of the original piece of digital content, such as an image, video, audio recording, or electronic document.

SB 1216 would further require the DWG to develop a coordinated plan to accomplish several goals, including: reducing the proliferation and impact of digital content forgeries and deepfakes; investigating the feasibility of, and obstacles to, developing standards and technologies for state departments for determining digital content provenance; increasing the ability of various entities, including members of the public, to meaningfully scrutinize and identify digital content forgeries, and to relay trust and information about digital content provenance to content consumers; developing or identifying mechanisms for content creators to cryptographically certify authenticity of original media and non-deceptive manipulations; and developing or identifying mechanisms for content creators to enable the public to validate the authenticity of original media and non-deceptive manipulations to establish content provenance.

Finally, the bill in print would require the DWG to report to the Legislature on or before July 1, 2024 regarding potential uses and risks of deepfake technology to the state government and California-based businesses, including recommendations for modifications to the definition of digital content forgery and deepfakes, and "recommendations for amendments to other code sections that may be impacted by the deployment of digital content forgery technologies and deepfakes."

The Silicon Valley Leadership Group argues in support of the bill:

> The recent proliferation of deepfakes has harmed the public's ability to discern fact from disinformation. This legislation would wisely bring together a diverse group of experts that represent privacy organizations, consumer advocacy associations, the tech industry, non-tech industry stakeholders, state agency representatives and legal experts to work with the Judicial Council.
>
> The creation of a Deepfake Working Group is a needed step for California to bring together subject matter experts that will combine their perspectives to address the challenges of deepfakes. We support the duties assigned to the Working Group that would include a report to the Legislature and a coordinated plan to reduce the proliferation and impact of deepfakes.

Adobe, Inc., who sponsor the bill, add:

> SB 1216 [] represents an important step toward increasing public and private sector collaboration in combatting the unique threat that digital content forgeries and misinformation campaigns pose to our state and our democracy. As studies have shown, we will continue to consume content digitally and we must find ways to protect against the dangers of falsely manipulated digital content. There could be 100 times more visual content by 2027, according to one study. One expert [] estimates that synthetic video may account for as much as 90% of online video in just three to five years. We applaud your effort to create the Deepfake Working Group which will bring together leading minds from across industries and sectors to help find solutions to these challenges.

5) **Author's amendments**: Though the bill in print takes a seemingly thoughtful approach to policymaking in the complex arena of digital content forgery and deepfakes by soliciting expert advice, the author has prudently offered several amendments to the bill to improve clarity and ensure that critical privacy considerations related to verifying digital content provenance do not go overlooked.

Amendment #1: While the bill in print repeatedly refers to "deepfakes" and requires that the DWG's report to the Legislature include recommendations for modifying the definition of deepfakes, the bill in print does not provide an initial definition of "deepfakes." Because there is no existing definition for deepfakes within California statute, this could result in ambiguity with respect to the mandate of the DWG, and provides no initial definition upon which the DWG can recommend modifications. To better define the boundaries of the topic the DWG would be tasked with exploring and to provide a starting point for the development of a precise definition of deepfakes, the author has offered the following amendment, providing a placeholder definition for "deepfake" to be honed by the DWG.

Author's amendment:

> On page 2, after line 4, insert the following, and renumber accordingly: "***(1) "Deepfake" means audio or visual content that has been generated or manipulated by artificial intelligence 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.***"

Amendment #2: The directives provided to the DWG by this bill are fairly comprehensive in contemplating a wide variety of considerations pertaining to digital content forgeries, digital content provenance, and deepfakes. One critical issue that the bill in print does not contemplate, however, is the privacy implications of mechanisms and technologies that permit verification of digital content provenance. By definition, verifying digital content provenance involves documenting and publishing the chronology of a piece of digital content. While doing so may provide benefits for verifying the authenticity of digital content, it also provides potentially invasive metadata regarding where and from whom the content originated, and may include information about its chain of custody depending on how it is chronicled. Considering the value placed on personal privacy and free expression in this state and in the nation as a whole, it is arguably imperative that the DWG ensure their plans and recommendations avoid infringing on privacy and chilling speech. Toward this end, the author has offered the following amendments, which would explicitly require the DWG to explore these topics relating to the privacy and civil liberties implications of technologies allowing public verification of digital content provenance.

Author's amendment:

> On page 2, after line 22, insert the following, and renumber accordingly: "***(3) Potential privacy impacts of technologies allowing public verification of digital content provenance.***"
>
> On page 2, line 26, strike "and deepfakes" and insert: "***, deepfakes, and technologies allowing public verification of digital content provenance.***"
>
> On page 3, line 33, after the word "exploring" insert: "***whether and***"
>
> On page 4, line 9, strike "content provenance" and insert: "***digital content provenance without materially compromising personal privacy or civil liberties***"

Amendment #3: In specifying the required content for the DWG's report to the Legislature, this bill repurposes language from AB 2658 that prescribed the reporting requirements of the Blockchain Working Group. These requirements specify that the report shall include "recommendations for amendments to other code sections that may be impacted by the deployment of digital content forgery technologies and deepfakes." Because this language refers to *amendments to code sections* rather than to legislation generally, it is unclear whether it would be within the purview of the DWG to recommend the development of new laws, rather than the modification of existing ones, should their findings indicate the need.

The author has offered the following technical amendments to correct drafting errors and to resolve the aforementioned ambiguity.

<u>Author's amendment</u>:

On page 4, line 15, strike "definition" and insert: "***definitions***"

On page 4, line 16, after the word "deepfakes" insert: "**,**"

On page 4, lines 16-17, strike "amendments to other code sections that may be impacted by the deployment of" and insert: "***appropriate or necessary legislation related to***"

6) Double referral: This bill has been double-referred to the Assembly Committee on Accountability & Administrative Review, where it will be heard should it pass out of this Committee.

7) Related legislation: AB 972 (Berman) would extend the existing sunset on the provisions of AB 730 (Berman, Ch. 493, Stats. 2019) pertaining to the use of materially deceptive audio and visual media in the context of elections until January 1, 2027.

8) Prior legislation: AB 613 (C. Garcia, 2021) would have required a social media platform, and a user or advertiser on a social media platform, to place a tag identifying that an image of a person has been retouched, and the manner in which that image has been altered from the original depiction of a real person. This bill did not receive a hearing in the Assembly Committee on Privacy & Consumer Protection.

AB 602 (Berman, Ch. 491, Stats. 2019) *See* Comment #3.

AB 730 (Berman, Ch. 493, Stats. 2019) *See* Comment #3.

AB 1280 (Grayson, 2019) would have codified and defined the term "deepfake", criminalized the nonconsensual production with the intent to distribute of a deepfake that depicts an individual engaging in sexual conduct, and would have criminalized the nonconsensual production within 60 days of an election of a deepfake depicting a person with the intent to coerce or deceive any voter into voting for or against a candidate or measure in that election. This bill died in the Assembly Committee on Public Safety.

REGISTERED SUPPORT / OPPOSITION:

Support

Adobe Systems (sponsor)
Anti-Defamation League
BSA The Software Alliance
California Medical Association
Silicon Valley Leadership Group

Opposition

None on file

Analysis Prepared by: Landon Klein / P. & C.P. / (916) 319-2200