

**SENATE JUDICIARY COMMITTEE**  
**Senator Thomas Umberg, Chair**  
**2021-2022 Regular Session**

SB 1018 (Pan)  
Version: April 18, 2022  
Hearing Date: April 26, 2022  
Fiscal: Yes  
Urgency: No  
AWM

**SUBJECT**

Platform Accountability and Transparency Act

**DIGEST**

This bill requires online platforms, as defined, to make annual public disclosures relating to the content on the platform, their use of algorithms and metrics, and their content moderation efforts.

**EXECUTIVE SUMMARY**

Given the reach of social media platforms and the role they play in many people's lives, concerns have arisen over what content permeates these sites, entering the lives of the billions of users, and the effects that the content has on them and society as a whole. In particular, the sharpest calls for action focus on the rampant spread of misinformation, hate speech, and sexually explicit content. Social media companies' content moderation of a decade ago involved handfuls of individuals and user policies were minimal. These programs and policies have dramatically evolved over the years, but the proliferation of objectionable content and "fake news" has led to calls for swifter and more aggressive action in response. However, there has also been backlash against perceived censorship in response to filtering of content and alleged "shadow banning."

This bill requires online platforms, as defined and with at least 1,000,000 monthly users in the state, to annually publicly disclose a representative sample of the content on their sites, along with information relating to their uses of algorithms and metrics and content moderation efforts. The bill authorizes the Attorney General enforce the bill through a civil action and to recover a civil penalty of up to \$100,000 per violation. To strengthen the bill and respond to some of the opposition's concerns, the author has agreed to amendments that narrow the disclosure requirements and further clarify what information must be disclosed.

This bill is sponsored by ProtectUS. The bill is opposed by CalChamber, NetChoice, and TechNet.

### PROPOSED CHANGES TO THE LAW

Existing law:

- 1) Prohibits, through the United States Constitution, the enactment of any law respecting an establishment of religion, or prohibiting the free exercise thereof; or abridging the freedom of speech, or of the press; or the right of the people peaceably to assemble, and to petition the Government for a redress of grievances. (U.S. Const., 1st amend.)
- 2) Provides, through the California Constitution, the right of every person to freely speak, write, and publish their sentiments on all subjects, being responsible for the abuse of this right. Existing law further provides that a law may not restrain or abridge liberty of speech or press. (Cal. Const., art. I, § 2(a).)
- 3) Provides, in federal law, that a provider or user of an interactive computer service shall not be treated as the publisher or speaker of any information provided by another information content provider. (47 U.S.C. § 230(c)(2).)
- 4) Provides that a provider or user of an interactive computer service shall not be held liable on account of:
  - a) Any action voluntarily taken in good faith to restrict access to or availability of material that the provider or user considers to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected; or
  - b) Any action taken to enable or make available to information content providers or others the technical means to restrict access to such material. (47 U.S.C. § 230(c)(2).)
- 5) Defines “interactive computer service” as any information service, system, or access software provider that provides or enables computer access by multiple users to a computer server, including specifically a service or system that provides access to the Internet and such systems operated or services offered by libraries or educational institutions. (47 U.S.C. § 230(f)(2).)
- 6) Requires certain businesses to disclose the existence and details of specified policies, including:
  - a) Operators of commercial websites or online services that collect personally identifiable information about individual consumers residing in California who use or visit the website must conspicuously post its privacy policy. (Bus. & Prof. Code, § 22575.)

- b) Retailers and manufacturers doing business in this state and having annual worldwide gross receipts over \$100,000,000 must disclose online whether the business has a policy to combat human trafficking and, if so, certain details about that policy. (Civ. Code, § 1714.43.)
- c) End-users of automated license plate recognition technology must post its usage and privacy policy on its website. (Civ. Code, § 1798.90.53.)
- d) Campus bookstores at public postsecondary educational institutions must post in-store or online a disclosure of its retail pricing policy on new and used textbooks. (Ed. Code, § 66406.7(f).)

This bill:

- 1) Establishes the Platform Accountability and Transparency Act.
- 2) Provides the following definitions:
  - a) "Personal information" has the same meaning as defined in Civil Code section 1798.140.
  - b) "Platform" means an internet service that meets both of the following criteria:
    - i. The service is an internet website, virtual reality, desktop application, or mobile application that does all of the following:
      - 1. Permits a person to become a registered user, establish an account, or create a profile for the purpose of allowing the user to create, share, and view user-generated content through that account or profile.
      - 2. Enables one or more users to generate content that can be viewed by other users of the platform.
      - 3. Primarily serves as a medium for users to interact with content generated by other users of the platform and for the platform to deliver advertisements to users.
    - ii. The service had at least 1,000,000 unique users in the state in at least 7 of the preceding 12 months.
  - c) "User" means a person that uses a platform, including an advertiser or seller, regardless of whether that person has an account or is otherwise registered with the platform.
- 3) Requires a platform to publicly disclose, on or before July 1, 2023, and on an annual basis after that, a statistically representative sampling of public content that is weighted by the number of impressions the content receives that includes all of the following:
  - a) The underlying content itself, including any public uniform resource locator link to the content.
  - b) The extent of dissemination or engagement with the content.
  - c) The audience reached with the content.
  - d) Whether the content has been determined to violate the platform's policies.

- e) The extent to which the content was recommended by the platform or otherwise amplified by platform algorithms.
  - f) Whether an account responsible for the content posted content deemed to violate the platform's policies in the past.
  - g) Other similar information deemed necessary by the office to accomplish the purposes of the disclosure requirement.
- 4) Requires a platform to publicly disclose, on or before July 1, 2023, and annually after that, all of the following information with respect to its use of algorithms and metrics:
- a) A description of all product features that made use of algorithms during the previous calendar year.
  - b) A summary of signals and features used as inputs to the algorithms referred to in 4)(a), including an explanation of all user data incorporated into these inputs ranked or based on the significance of their impact on the algorithms' outputs.
  - c) A summary of data-driven models, including those based on machine learning or other artificial intelligence techniques, utilized in the described algorithms, including the optimization objective of those models ranked based on the significance of their impact on the algorithms' outputs.
  - d) A summary of metrics used by the platform to score or rank content based on the significance of their impact on the algorithms' outputs.
  - e) A summary of metrics calculated by the platform operator to assess product changes or new features with an assessment of their relative importance in platform decisionmaking.
  - f) Other similar information deemed necessary by the office to accomplish the purposes of this disclosure requirement.
- 5) Requires a platform to publicly disclose, on or before July 1, 2023, and annually after that, all of the following information with respect to content moderation and content violating platform policies:
- a) Statistics regarding the amount of content that the platform determined violated its policies organized by all of the following:
    - i. The violated policy.
    - ii. The action taken in response to the violation.
    - iii. The methods the platform used to identify the violating content.
    - iv. The extent to which the content was recommended or otherwise amplified by platform algorithms.
    - v. The extent to which the user chose to follow the account that originated or spread the violating content, and, if so, whether that account had been recommended to the user by the platform.
  - b) Statistics regarding the number of times violating content was viewed by users and the number of users who viewed it.
  - c) Estimates by the platform about the prevalence of violating content.

- d) Other similar information deemed necessary by the office to accomplish the purposes of this disclosure requirement.
- 6) Provides that the disclosure requirements in 2)-5) do not require the dissemination of confidential business information or trade secrets.
- 7) Provides that a violation of the requirements of this bill is punishable by a civil penalty of up to \$100,000 for each violation, in a civil action that may be brought only by the Attorney General.

### COMMENTS

#### 1. Author's comment

According to the author:

The public should not need to wait for a whistleblower to discover how we are manipulated. Addressing the many public policy concerns involving social media begins with more transparency. SB 1018 would require social media platforms to disclose information about the mechanisms that influence the content a user sees.

#### 2. Social media content, the lack of transparency into platforms' content moderation efforts, and considerations for how to solve the problem

The concerns underlying this bill – that social media platforms are responsible for spreading dangerous disinformation that has hurt everything from our national health to the integrity of our democracy – are the same ones behind another bill pending before this committee, AB 587 (Gabriel, 2021). The analysis for that bill provided an eloquent summary of the issues posed by social media and the lack of transparency into how they moderate content, which is set forth in part here:

In recent years, the clamor for more robust content moderation on social media has reached a fever pitch. This includes calls to control disinformation or “fake news,” hate speech, political interference, and other online harassment.

The 2016 election was a major breaking point for many. Investigations uncovered attempted interference in the United States Presidential election through a social media “information warfare campaign designed to spread disinformation and societal division in the United States.”<sup>1</sup> [Despite platforms' awareness of the interference with the

---

<sup>1</sup> Select Committee on Intelligence, Russian Active Measures, Campaigns, and Interference in the 2016 U.S. Election, United States Senate, *available at* [https://www.intelligence.senate.gov/sites/default/files/documents/Report\\_Volume2.pdf](https://www.intelligence.senate.gov/sites/default/files/documents/Report_Volume2.pdf) (last visited Apr. 15, 2022).

election in 2016], [t]his again became a threat in the 2020 election, with social media rife with misinformation such as the incorrect election date,<sup>2</sup> and then social media became a hotbed of misinformation about the results of the election.<sup>3</sup> The author points to investigations that have found the violent insurrectionists that stormed the Capitol on January 6, 2021, were abetted and encouraged by posts on social media sites.<sup>4</sup> In response to indications that social media provided a venue for those who overran and assaulted police officers, Facebook deflected blame, asserting that “these events were largely organized on platforms that don’t have our abilities to stop hate, don’t have our standards, and don’t have our transparency.”<sup>5</sup> However, later indictments of those perpetrating the attack “made it clear just how large a part Facebook had played, both in spreading misinformation about election fraud to fuel anger among the Jan. 6 protesters, and in aiding the extremist militia’s communication ahead of the riots.”<sup>6</sup>

Misinformation also poses a danger to public health: One study found that the more people rely on social media as their main news source, the more likely they are to believe misinformation about the COVID-19 pandemic.<sup>7</sup> Another found that a mere 12 people are responsible for 65 percent of the false and misleading claims about COVID-19 vaccines on Facebook, Instagram, and Twitter.<sup>8</sup> Misinformation can also hinder emergency responses to natural disasters, when social media posts contain incorrect or out-of-date information.<sup>9</sup>

---

<sup>2</sup> Fessler, *Robocalls, Rumors And Emails: Last-Minute Election Disinformation Floods Voters*, NPR (Oct. 24, 2020), <https://www.npr.org/2020/10/24/927300432/robocalls-rumors-and-emails-last-minute-election-disinformation-floods-voters> (last visited Apr. 15, 2022).

<sup>3</sup> Frenkel, *How Misinformation ‘Superspreaders’ Seed False Election Theories*, New York Times (Nov. 23, 2020), available at <https://www.nytimes.com/2020/11/23/technology/election-misinformation-facebook-twitter.html> (last visited Apr. 15, 2022); Bump, *The chain between Trump’s misinformation and violent anger remains unbroken*, Washington Post (May 12, 2021), <https://www.washingtonpost.com/politics/2021/05/12/chain-between-trumps-misinformation-violent-anger-remains-unbroken/> (last visited Apr. 15, 2022).

<sup>4</sup> Dilanian & Collins, *There are hundreds of posts about plans to attack the Capitol. Why hasn’t this evidence been used in court?* NBC News (April 20, 2021), <https://www.nbcnews.com/politics/justice-department/we-found-hundreds-posts-about-plans-attack-capitol-why-aren-n1264291> (last visited Apr. 15, 2022).

<sup>5</sup> Frenkel & Kang, *Mark Zuckerberg and Sheryl Sandberg’s Partnership Did Not Survive Trump*, New York Times (July 8, 2021; updated Oct. 20, 2021), <https://www.nytimes.com/2021/07/08/business/mark-zuckerberg-sheryl-sandberg-facebook.html> (last visited Apr. 15, 2022).

<sup>6</sup> *Ibid.*

<sup>7</sup> Su, *It doesn’t take a village to fall for misinformation: Social media use, discussion heterogeneity preference, worry of the virus, faith in scientists, and COVID-19-related misinformation belief* Telematics and Information, Vol. 58 (May 2021), available at <https://www.sciencedirect.com/science/article/abs/pii/S0736585320302069?via%3Dihub> (last visited Apr. 15, 2022)

<sup>8</sup> Bond, *Just 12 People Are Behind Most Vaccine Hoaxes On Social Media, Research Shows*, NPR (May 14, 2021), <https://www.npr.org/2021/05/13/996570855/disinformation-dozen-test-facebooks-twitters-ability-to-curb-vaccine-hoaxes> (last visited Apr. 15, 2021).

<sup>9</sup> United States Department of Homeland Security, *Countering False Information on Social Media in Disasters and Emergencies* (March 2018), available at

A recent Congressional Research Services Report discussed the issue of content moderation and, specifically, the spread of misinformation and the role that social media companies play in worsening the issue:

Two features of social media platforms—the user networks and the algorithmic filtering used to manage content—can contribute to the spread of misinformation. Users can build their own social networks, which affect the content that they see, including the types of misinformation they may be exposed to. Most social media operators use algorithms to sort and prioritize the content placed on their sites. These algorithms are generally built to increase user engagement, such as clicking links or commenting on posts. In particular, social media operators that rely on advertising placed next to user-generated content as their primary source of revenue have incentives to increase user engagement. These operators may be able to increase their revenue by serving more ads to users and potentially charging higher fees to advertisers. Thus, algorithms may amplify certain content, which can include misinformation, if it captures users’ attention.<sup>10</sup>

The role that content moderation, or the lack of it, has in alleviating or exacerbating these issues has been a source of much debate. A policy paper published by the Shorenstein Center on Media, Politics, and Public Policy at the Harvard Kennedy School, *Countering Negative Externalities in Digital Platforms*, focuses on the costs associated with various internet platforms that are not absorbed by the companies themselves:

Today, in addition to the carcinogenic effects of chemical runoffs and first and second hand tobacco smoke, we have to contend with a new problem: the poisoning of our democratic system through foreign influence campaigns, intentional dissemination of misinformation, and incitements to violence inadvertently enabled by Facebook, YouTube and our other major digital platform companies.<sup>11</sup>

The paper asserts that these major platform companies “enable exceptionally malign activities” and “experience shows that the companies have not made sufficient investments to eliminate or reduce these negative externalities.”

---

[https://www.dhs.gov/sites/default/files/publications/SMWG\\_Countering-False-Info-Social-Media-Disasters-Emergencies\\_Mar2018-508.pdf](https://www.dhs.gov/sites/default/files/publications/SMWG_Countering-False-Info-Social-Media-Disasters-Emergencies_Mar2018-508.pdf) (last visited Apr. 15, 2022).

<sup>10</sup> Gallo & Cho, *Social Media: Misinformation and Content Moderation Issues for Congress*, Congressional Research Service (Jan. 27, 2021) available at <https://crsreports.congress.gov/product/pdf/R/R46662> (last visited Apr. 15, 2022).

<sup>11</sup> Verveer, *Countering Negative Externalities in Digital Platforms*, Shorenstein Center on Media, Politics and Public Policy (Oct. 7, 2019) <https://shorensteincenter.org/countering-negative-externalities-in-digital-platforms/> (last visited Apr. 15, 2022).

There are a number of considerations when addressing how to approach the proliferation of these undesirable social media posts and the companies' practices that fuel the flames. A number of methods of content moderation are being deployed and have evolved from simply blocking content or banning accounts to quarantining topics, removing posts from search results, barring recommendations, and down ranking posts in priority. However, there is a lack of transparency and understanding of exactly what companies are doing and why it does not seem to be enough. A recent article in the MIT Technology Review articulates the issues with content moderation behind the curtain:

As social media companies suspended accounts and labeled and deleted posts, many researchers, civil society organizations, and journalists scrambled to understand their decisions. The lack of transparency about those decisions and processes means that—for many—the election results end up with an asterisk this year, just as they did in 2016.

What actions did these companies take? How do their moderation teams work? What is the process for making decisions? Over the last few years, platform companies put together large task forces dedicated to removing election misinformation and labeling early declarations of victory. Sarah Roberts, a professor at UCLA, has written about the invisible labor of platform content moderators as a shadow industry, a labyrinth of contractors and complex rules which the public knows little about. Why don't we know more?

In the post-election fog, social media has become the terrain for a low-grade war on our cognitive security, with misinformation campaigns and conspiracy theories proliferating. When the broadcast news business served the role of information gatekeeper, it was saddled with public interest obligations such as sharing timely, local, and relevant information. Social media companies have inherited a similar position in society, but they have not taken on those same responsibilities. This situation has loaded the cannons for claims of bias and censorship in how they moderated election-related content.<sup>12</sup>

### 3. This bill requires certain online platforms to annually disclose information relating to their content and content moderation efforts

Like AB 587, this bill is intended to give the public greater information about how social media platforms moderate their content, including insight into what content is moderated and by whom. This bill is adapted from a bipartisan federal bill that was

---

<sup>12</sup> Donovan, *Why social media can't keep moderating content in the shadows*, MIT Technology Review (Nov. 6, 2020), <https://www.technologyreview.com/2020/11/06/1011769/social-media-moderation-transparency-censorship/> (last visited Apr. 16, 022).



announced in late 2021 by co-authors Senator Amy Klobuchar, Senator Chris Coons, and Senator Rob Portman.<sup>13</sup> The bill has yet to be formally introduced in Congress. The bill – also named the Platform Accountability and Transparency Act (PATA) – is broader than this bill: the federal PATA establishes in the Federal Trade Commission a Platform Accountability and Transparency Office that would serve as a bridge between researchers and social media platforms; imposes on platforms certain obligations to turn over information to researchers; narrows the federal law that gives online platforms wide discretion in how they moderate content; and establishes certain safe harbors and immunities for researchers and platforms.<sup>14</sup>

This bill borrows the federal bill’s requirement that online platforms disclose certain information relating to content moderation on an annual basis. Specifically, this bill requires an online platform to annually disclose a statistically representative sample of public content on the platform, along with specified information about the content; information relating to an online platform’s use of algorithms and the metrics the platform uses to analyze content, along with related information; and information relating to the platform’s content moderation efforts, along with related information. The federal PATA also requires online platforms to disclose certain information relating to advertisements, but the author has not included that requirement here.

The bill’s definition of “online platform” appears to be intended to be limited to “traditional” social media platforms (e.g., Facebook, Instagram, YouTube). The bill’s definition defines a platform as a website, virtual reality, or application that permits a person to become a user, share and create content on the platform, and view other users’ content, and that *primarily* serves as a medium for users to act with each others’ content and for the platform to deliver advertisements. The inclusion of this primary purpose likely excludes other platforms that incidentally involve features such as usernames or user interactions, such as online commerce sites that allow for comments or platforms that provide professionally created media (e.g., Netflix or Hulu).

Additionally, the definition of “online platform” is limited to platforms with 1,000,000 or more discrete users in the state of California in seven of the prior 12 months. While the largest social media platforms will likely clear this hurdle easily, it is not clear how many platforms will be covered.

The bill provides that an online platform need not disclose confidential business or trade secrets as part of its required annual disclosures. While Committee staff is not

---

<sup>13</sup> See Wright, *The Platform and Accountability Act: New legislation addresses platform data secrecy*, Stanford Cyber Policy Center (Dec. 9, 2021), <https://cyber.fsi.stanford.edu/news/platform-transparency-and-accountability-act-new-legislation-addresses-platform-data-secrecy> (last visited Apr. 15, 2022).

<sup>14</sup> See S. \_\_\_\_, proposed Platform Accountability and Transparency Act, *available at* [https://www.coons.senate.gov/imo/media/doc/text\\_pata\\_117.pdf](https://www.coons.senate.gov/imo/media/doc/text_pata_117.pdf) (last visited Apr. 15, 2022). SB 1018 had significantly more overlap with the federal PATA prior to the author’s April 18, 2022, amendments.

aware of what information required by the disclosures would fall into these categories, this appears to be a reasonable measure to protect online platforms.

Finally, violations of this bill are enforceable only through a civil action brought by the Attorney General. The bill authorizes a court to award a civil penalty of up to \$100,000 per violation.

The author has agreed to amend the bill to address certain concerns of the opposition and to clarify the bill's disclosure requirements. The amendments are discussed in Part 4 below and set forth in the appendix. As noted above, this bill broadly shares the goals and outcomes of AB 587 (Gabriel, 2022). Going forward the authors may wish to work together to develop a unified scheme for ensuring transparency from social media platforms.

#### 4. Amendments

The author has agreed to amend the bill in several ways. Broadly speaking, the amendments are intended to (1) harmonize certain definitions with the definitions in other social media bills passed by this Committee; (2) remove the disclosure requirements relating to the "statistically representative sampling" of content, (3) clarify some of the remaining disclosure requirements to, among other things, ensure that platforms are not required to disclose proprietary information or trade secrets and provide greater specificity as to what is required, and (4) remove references to an office that would have been established in a prior version of the bill. These amendments address many of the opposition's concerns. A full mock-up of the amendments is set forth at the end of this analysis, subject to any technical and/or nonsubstantive changes Legislative Counsel may make.

#### 5. Arguments in support

According to bill sponsor ProtectUS:

With social media algorithms notoriously prioritizing content based on popularity and similarity to previously seen content, users exposed to misinformation once may end up seeing more and more of it over time, further reinforcing the user's misunderstanding. Given the deadly impacts of public health disinformation, it is imperative that we hold social media platforms accountable for prioritizing profits over people.

From our support for evidence-based public health policy to our creation of Masked Up, a show dedicated to debunking health disinformation, combating online disinformation has been a top priority for our organization. SB 1018 would assist in our goal of protecting the health of our most vulnerable

communities by making it easier for the public to hold social media companies accountable for their spread of disinformation.

6. Arguments in opposition

According to CalChamber, NetChoice, and TechNet, writing together in opposition:

While our companies and platforms believe in the value of transparency regarding content moderation and have taken many steps in recent years to release more information and data about community standards and efforts to enforce them, this bill's disclosure requirements go too far and would not further its intent.

The disclosure requirements in SB 1018 would require companies to disclose granular details about content on their platform, as well as require the disclosure of sensitive company information and intellectual property. First, the bill requires companies to disclose a "statistically representative sampling of public content," which could mean millions of entries... This bill fails to take into account that hundreds of millions of new pieces of content are posted or shared on social media platforms every single day. Not only would this information be impossible to compile, it would be useless to anyone seeking to use it to better understand harmful content on our platforms. This requirement would completely overshadow the bill's intent of providing greater transparency...

[O]ur companies use algorithms to protect users' data and identify and remove threats. The disclosures required by SB 1018 would give hackers a leg up in trying to breach our systems by providing descriptions of each one as well as identifying where and how we deploy automated systems. The same holds true for our efforts to moderate content on our platforms because due to the sheer volume of content we use automated systems to identify and remove harmful content. This bill would require our companies to provide detailed descriptions of how our automated systems identify and remove harmful content. These descriptions will provide roadmaps to get around our protections to host of bad actors including extremists, terrorist organizations, child predators, scammers, and serial abusers.

SUPPORT

ProtectUS (sponsor)

OPPOSITION

CalChamber  
NetChoice  
TechNet

## RELATED LEGISLATION

### Pending Legislation:

AB 2826 (Muratsuchi, 2022) is similar to this bill and the announced federal PATA legislation, and would require the Department of Technology to establish a program to identify qualified research projects and require online platforms to turn over research material for those projects, and to submit annual reports to specified members of the Legislature relating to the research projects and research conducted. AB 2826 is pending before the Assembly Privacy and Consumer Protection Committee.

AB 587 (Gabriel, 2021) requires social media companies, as defined, to post their terms of service and report certain information to the Attorney General on a quarterly basis. AB 587 is pending before the Senate Judiciary Committee.

### Prior Legislation:

SB 388 (Stern, 2021) would have required a social media platform company, as defined, that, in combination with each subsidiary and affiliate of the service, has 25,000,000 or more unique monthly visitors or users for a majority of the preceding 12 months, to report to the Department of Justice by April 1, 2022, and annually thereafter, certain information relating to its efforts to prevent, mitigate the effects of, and remove potentially harmful content. SB 388 died in the Senate Judiciary Committee.

AB 1114 (Gallagher, 2021) would have required a social media company located in California to develop a policy or mechanism to address content or communications that constitute unprotected speech, including obscenity, incitement of imminent lawless action, and true threats, or that purport to state factual information that is demonstrably false. AB 1114 died in the Assembly Arts, Entertainment, Sports, Tourism, and Internet Media Committee.

AB 35 (Chau, 2021) would have required social media platforms, as defined, to disclose in an easy-to-find location whether they have a policy to combat misinformation, and imposed civil penalties for the failure to comply. AB 35 was pending before the Senate Judiciary Committee when the author left the Assembly.

AB 2442 (Chau, 2020) was substantially similar to AB 35 (Chau, 2020) and would have required social media companies to disclose the existence, or lack thereof, of a misinformation policy, and imposed civil penalties for failing to do so. AB 2442 died in the Senate Judiciary Committee.

AB 1316 (Gallagher, 2019) would have prohibited social media sites from removing user-posted content on the basis of the political affiliation or viewpoint of that content, except where the social media site is, by its terms and conditions, limited to the

promotion of only certain viewpoints and values and the removed content conflicts with those viewpoints or values. AB 1316 was held on the floor of the Assembly and was re-introduced as AB 2931 (2020).

SB 1424 (Pan, 2018) would have established a privately funded advisory group to study the problem of the spread of false information through Internet-based social media platforms, and draft a model strategic plan for Internet-based social media platforms to use to mitigate this problem. SB 1424 was vetoed by Governor Edmund J. Brown, whose veto message stated that, as evidenced by the numerous studies by academic and policy groups on the spread of false information, the creation of a statutory advisory group to examine this issue is not necessary.

\*\*\*\*\*

**Amended Mock-up for SB-1018 (Pan)**

**Mock-up based on Version Number 97 - Amended Senate 4/18/22**

THE PEOPLE OF THE STATE OF CALIFORNIA DO ENACT AS FOLLOWS:

SECTION 1. Chapter 5.9 (commencing with Section 11549.65) is added to Part 1 of Division 3 of Title 2 of the Government Code, to read:

**CHAPTER 5.9. Platform Accountability and Transparency Act**

**11549.65.** This chapter shall be known as the "Platform Accountability and Transparency Act."

**11549.66.** As used in this chapter:

(a) "Personal information" has the same meaning as defined in Section 1798.140 of the Civil Code.

(b) "Social media platform" or "Platformplatform" means an internet service or application that has users in California and that meets both of the following criteria:

(1) The service is an internet website, virtual reality, desktop application, or mobile application that does all of the following: primary purpose of the service or application is to connect users and allow users to interact with each other in the service or application.

(2) The service or application allows users to do all of the following:

(A) Permits a person to become a registered user, establish an account, or create a profile for the purpose of allowing the user to create, share, and view user-generated content through that account or profileConstruct a public or semipublic profile within a bounded system created by the service or application.

(B) Enables one or more users to generate content that can be viewed by other users of the platformPopulate a list of other users with whom an individual shares a connection within the system.

(C) Primarily serves as a medium for users to interact with content generated by other users of the platform and for the platform to deliver advertisements to usersView and navigate a list of connections made by other individuals within the system.

**(D) Create or post content viewable by other users.**

~~(2) The service had at least 1,000,000 unique users in the state in at least 7 of the preceding 12 months.~~

(c) "User" means a person that uses a platform, including an advertiser or a seller, regardless of whether that person has an account or is otherwise registered with the platform.

~~11549.67. (a) A platform operator shall disclose, on or before July 1, 2023, and annually thereafter, to the public a statistically representative sampling of public content that is weighted by the number of impressions the content receives that includes all of the following:~~

~~(1) The underlying content itself, including any public uniform resource locator link to the content.~~

~~(2) The extent of dissemination of or engagement with the content.~~

~~(3) The audience reached with the content.~~

~~(4) Whether the content has been determined to violate the platform's policies.~~

~~(5) The extent to which the content was recommended by the platform or otherwise amplified by platform algorithms.~~

~~(6) Whether an account responsible for the content posted content deemed to violate the platform's policies in the past.~~

~~(7) Other similar information deemed necessary by the office to accomplish the purposes of this subdivision.~~

~~(ba) A social media platform operator shall disclose, on or before July 1, 2023, and annually thereafter, to the public all of the following information with respect to its use of algorithms and metrics in the preceding 12-month period:~~

(1) A description of all product features that made use of algorithms during the previous calendar year.

(2) A summary of signals and features used as inputs to the algorithms referred to in paragraph (1), including an explanation of all user data incorporated into these inputs ranked or based on the significance of their impact on the algorithms' outputs. **This paragraph does not require the disclosure of specific algorithms.**

(3) A summary of data-driven models, including those based on machine learning or other artificial intelligence techniques, utilized in the described algorithms, including the optimization objective of those models ranked based on the significance of their impact on the algorithms' outputs. **This paragraph does not require the disclosure of specific models.**

(4) A summary of metrics used by the platform to score or rank content ranked based on the significance of their impact on the algorithms' outputs. **This paragraph does not require the disclosure of specific metrics.**

(5) A summary of metrics calculated by the platform operator to assess product changes or new features with an assessment of their relative importance in platform decisionmaking. **This paragraph does not require the disclosure of specific metrics.**

~~(6) Other similar information deemed necessary by the office to accomplish the purposes of this subdivision.~~

(eb) A **social media** platform ~~operator~~ shall disclose, on or before July 1, 2023, and annually thereafter, to the public all of the following information with respect to content moderation and content violating platform policies **in the preceding 12-month period:**

(1) Statistics regarding the ~~amount-items~~ of content that the platform determined violated its policies, **including organized by** all of the following:

(A) The **breakdown of the total number of items of content by category of** violated policy.

(B) The **breakdown of the total number of items of content that the platform took action against by category of** action taken in response to the violation.

(C) The **breakdown of the total number of items of content that were identified as violating a policy by category of** methods the platform used to identify the violating content.

(D) The extent to which the content was recommended or otherwise amplified by platform algorithms **before and after it was identified as in violation of the platform's policies.**

~~(E) The extent to which the user chose to follow the account that originated or spread the violating content, and, if so, whether that account had been recommended to the user by the platform.~~



(2) Statistics regarding the number of times violating content was viewed by users and the number of users who viewed it as a percentage of total views.

(3) Estimates by the platform about the prevalence of violating content as a percentage of total content.

~~(4) Other similar information deemed necessary by the office to accomplish the purposes of this subdivision.~~

(d) This section does not require the dissemination of confidential business information or trade secrets.

**11549.68.** A violation of this chapter shall subject the violator to a civil penalty of up to one hundred thousand dollars (\$100,000) for each violation that may be recovered only in a civil action brought by the Attorney General.

**11549.69. This chapter does not apply to a social media platform with fewer than 1,000,000 discrete monthly users.**