Date of Hearing: April 30, 2024

ASSEMBLY COMMITTEE ON PRIVACY AND CONSUMER PROTECTION
Rebecca Bauer-Kahan, Chair
AB 2013 (Irwin) – As Amended April 22, 2024

AS PROPOSED TO BE AMENDED

SUBJECT: Artificial intelligence: training data transparency

SYNOPSIS

There is a common saying in data science: "garbage in, garbage out." When it comes to artificial intelligence (AI), data is everything. Whether an AI's outputs are useful and fair depends entirely on the data used to train it – at present, however, Californians have no insight into which data are used to train which products. As a result, they cannot make informed decisions when purchasing AI products to help maintain their businesses, or when exchanging sensitive personal information for services that will ostensibly improve their quality of life.

This bill would require developers of AI systems and services to publicly disclose specified information related to the datasets used to train their products. In doing so, this bill would allow Californians to make informed decisions about the AI systems they purchase and engage with.

This bill is author-sponsored and supported by Oakland Privacy, Secure Justice, Transparency Coalition.ai, Concept Art Association, and Santa Monica Democratic Club. A coalition of industry associations, including California Chamber of Commerce and Technet, takes an "oppose unless amended" position. The bill is opposed by Chamber of Progress.

SUMMARY: Requires a developer of an AI system or service to publicly disclose specific information related to the system or service's training data. Specifically, **this bill**:

1) Requires a developer of an AI system or service to post documentation related to its training data to the developer's internet website on or before January 1, 2026, and before each time thereafter than an AI system or service is made available to Californians.

2) Requires documentation related to training data to contain a description of each dataset used to develop the AI system or service, including:

    a) The source or owner of the dataset.

    b) A description of how the dataset furthers the intended purpose of the system or service.

    c) The number of data points included in the dataset, with estimated figures for dynamic datasets.

    d) A clear definition of each category associated with data points within the dataset, including the format of data points and sample values.

    e) Whether the dataset includes any data protected by copyright, trademark, or patent, requiring the purchase or licensure of the data, or whether the dataset is entirely in the public domain.

f) Whether the data was purchased or licensed by the developer.

g) Whether the dataset includes personal information.

h) Whether the dataset includes aggregate consumer information.

i) A description of any cleaning, processing, or modification to the dataset by the developer, including the intended purpose of those efforts.

j) The time period during which the data was collected.

k) Whether data collection is ongoing.

l) The dates the dataset was first and last used during development of the AI system or service.

3) Requires a developer of a system or service to disclose whether the system or service used or uses synthetic data generation in its development.

4) Exempts AI systems or services whose sole purpose is to help ensure security and integrity.

EXISTING LAW:

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

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

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

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

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

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

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

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

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

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

e) Biometric information.

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

g) Geolocation data.

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

i) Professional or employment-related information.

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

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

l) Sensitive personal information.

4) Defines biometric information to mean an individual's physiological, biological, or behavioral characteristics, including information pertaining to an individual's deoxyribonucleic acid (DNA), that is used or is intended to be used singly or in combination with each other or with other identifying data, to establish individual identity. (Civ. Code § 1798.140(c).)

5) Further defines "personal information" to include any information that identifies, relates to, describes, or is capable of being associated with, a particular individual, including, but not limited to, his or her name, signature, social security number, physical characteristics or description, address, telephone number, passport number, driver's license or state identification card number, insurance policy number, education, employment, employment history, bank account number, credit card number, debit card number, or any other financial information, medical information, or health insurance information. (Civ. Code § 1798.80(e).)

a) States that personal information does not include publicly available information that is lawfully made available to the general public from federal, state, or local government records.

6) Defines sensitive personal information to mean any of the following:

   a) Personal information that reveals:

      i) A consumer's social security, driver's license, state identification card, or passport number.

      ii) A consumer's account log-in, financial account, debit card, or credit card number in combination with any required security or access code, password, or credentials allowing access to an account.

      iii) A consumer's precise geolocation.

      iv) A consumer's racial or ethnic origin, citizenship or immigration status, religious or philosophical beliefs, or union membership.

      v) The contents of a consumer's mail, email, and text messages unless the business is the intended recipient of the communication.

      vi) A consumer's genetic data.

   b) The processing of biometric information for the purpose of uniquely identifying a consumer.

   c) Personal information collected and analyzed concerning a consumer's health.

   d) Personal information collected and analyzed concerning a consumer's sex life or sexual orientation. (Civ. Code § 1798.140(ae).)

7) Defines "aggregate personal information" to mean information that relates to a group or category of consumers, from which individual consumer identities have been removed, that is not linked or reasonably linkable to any consumer or household, including via a device. Excludes from this definition one or more individual consumer records that have been deidentified. (Civ. Code § 1798.140(b).)

8) Defines "security and integrity" to mean the ability of:

   a) Networks or information systems to detect security incidents that compromise the availability, authenticity, integrity, and confidentiality of stored or transmitted personal information.

   b) Businesses to detect security incidents, resist malicious, deceptive, fraudulent, or illegal actions and to help prosecute those responsible for those actions.

   c) Businesses to ensure the physical safety of natural persons. (Civ. Code § 1798.140(ac).)

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

COMMENTS:

1) **Artificial intelligence.** The development of AI is creating exciting opportunities to grow California's economy and improve the lives of its residents. AI can generate compelling text and convincing images in an instant. It can automate painstaking tasks, identify subtle patterns in large datasets, and make accurate predictions in the face of incomplete information. But with novel technologies come novel safety concerns. The present bill furthers consumer protection in California by granting the state's residents insight into how the AI systems and services they engage with are trained.

2) **The importance of training.** AI uses algorithms – sets of rules – to transform inputs into outputs. Inputs and outputs can be anything a computer can process: numbers, text, audio, video, or movement. This is because AI is not fundamentally different from other computer functions. Its novelty lies in its application: unlike normal computer functions, AI is able to accomplish tasks that are normally performed by humans.

Training is the secret sauce of machine learning; it is the principle innovation that allows modern AI to be both efficient and versatile. During training, a naïve AI is exposed to data and allowed to automatically explore its structure. As the AI explores, it alters itself in an attempt to better represent the data. Each piece of data affects every part of an AI. In a sense, AI "digest" and integrate the data they train on in order to learn, just as humans digest and integrate the foods we eat in order to grow.

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

3) **Haphazard training data.** There is a common saying in computer science: "garbage in, garbage out." The performance of an AI product is directly impacted by the quality, quantity, and relevance of the data used to train it. Before training, datasets are often categorized to make them easier for AI to work with. Rigorously categorizing the data in a dataset becomes more difficult as the dataset becomes larger, but failing to organize its contents can lead to meaningless, false, or harmful outputs.

The biggest names in AI – OpenAI, Meta, and Google – understand AI's critical need for data better than anyone else. According to a recent New York Times examination, the race to lead in the AI space has become a desperate hunt for digital data. To obtain that data, these tech companies have cut corners, ignored corporate policies and debated bending the law:

> At Meta, which owns Facebook and Instagram, managers, lawyers and engineers last year discussed buying the publishing house Simon & Schuster to procure long works, according to recordings of internal meetings obtained by The Times. They also conferred on gathering copyrighted data from across the internet, even if that meant facing lawsuits. Negotiating licenses with publishers, artists, musicians and the news industry would take too long, they said.

> Like OpenAI, Google transcribed YouTube videos to harvest text for its A.I. models, five people with knowledge of the company's practices said. That potentially violated the copyrights to the videos, which belong to their creators.
>
> Last year, Google also broadened its terms of service. One motivation for the change, according to members of the company's privacy team and an internal message viewed by The Times, was to allow Google to be able to tap publicly available Google Docs, restaurant reviews on Google Maps and other online material for more of its A.I. products.[1]

In their race to obtain vast quantities of training data, major AI developers have not hesitated to move fast and break things. The Stanford Internet Observatory recently discovered that a common image training dataset known as LAION-5B contains many instances of child sexual abuse materials. Their study identified 3226 dataset entries of suspected child pornography, much of which was later confirmed as such by third parties.[2] This dataset was built by automatically scraping the internet, and images containing child pornography were found to have originated from large, well-known websites such as Reddit, Twitter, Blogspot, and Wordpress, as well as mainstream adult sites such as XHamster and XVideos.

4) **An AI never forgets.** Just as humans cannot intentionally forget information they have learned, it is not currently possible to remove data from a trained AI.[3] Unlike an Excel spreadsheet, which stores data in neat columns, AI stores data in the connections between "neurons" in a "neural network." Every one of these connections is influenced by every piece of training data, and a large model like ChatGPT-4 is reported to have more than 1.7 trillion connections.[4] It is not possible to specifically alter these connections in order to remove data without fundamentally changing the model; as a result, for data to be removed, the model must be retrained from scratch. ChatGPT-4 is estimated to have taken 4-7 months to train.[5]

5) **Synthetic data.** AB 2013 requires AI developers disclose "whether [their] system or service used or continuously uses synthetic data generation in its development." Synthetic data is artificially generated data that is created, rather than collected from real-world events. It is designed to mimic the statistical properties of authentic data, and can be useful for training AI when actual data may be limited, sensitive, or biased. Having already gobbled up most of the high-quality data that humanity has produced, major generative AI developers have begun looking to synthetic data:

> OpenAI's Mr. Altman had a plan to deal with the looming data shortage.

---

[1] Cade Metz, Cecilia Kang, Sheera Frenkel, Stuart A. Thompson and Nico Grant, "How Tech Giants Cut Corners to Harvest Data for A.I.," *New York Times*, Apr. 6, 2024, https://www.nytimes.com/2024/04/06/technology/tech-giants-harvest-data-artificial-intelligence.html.
[2] David Thiel, "Identifying and Eliminating CSAM in Generative ML Training Data and Models," *Stanford Internet Observatory*, Dec. 23, 2023.
[3] Stephen Pastis, "A.I.'s un-learning problem: Researchers say it's virtually impossible to make an A.I. model 'forget' the things it learns from private user data," *Yahoo! Finance*, Aug. 30, 2023, finance.yahoo.com/news/un-learning-problem-researchers-virtually-164342971.html.
[4] Reed Albergotti, "Microsoft pushes the boundaries of small AI models with big breakthrough," *SEMAFOR*, Nov. 1, 2023, www.semafor.com/article/11/01/2023/microsoft-pushes-the-boundaries-of-small-ai-models.
[5] Stephen McAleese, "Retrospective on 'GPT-4 Predictions' After the Release of GPT-4," *LESSWRONG*, Mar. 17, 2023, https://www.lesswrong.com/posts/iQx2eeHKLwgBYdWPZ/retrospective-on-gpt-4-predictions-after-the-release-of-gpt.

> Companies like his, he said at the May conference, would eventually train their A.I. on text generated by A.I. — otherwise known as synthetic data.
>
> Since an A.I. model can produce humanlike text, Mr. Altman and others have argued, the systems can create additional data to develop better versions of themselves. This would help developers build increasingly powerful technology and reduce their dependence on copyrighted data.
>
> "As long as you can get over the synthetic data event horizon, where the model is smart enough to make good synthetic data, everything will be fine," Mr. Altman said.
>
> A.I. researchers have explored synthetic data for years. But building an A.I system that can train itself is easier said than done. A.I. models that learn from their own outputs can get caught in a loop where they reinforce their own quirks, mistakes and limitations.
>
> "The data these systems need is like a path through the jungle," said Jeff Clune, a former OpenAI researcher who now teaches computer science at the University of British Columbia. "If they only train on synthetic data, they can get lost in the jungle."[6]

There are risks associated with relying on synthetic data. First, synthetic datasets may not perfectly replicate the complexity and variability of real-world data. This discrepancy can lead to models that perform well when tested in isolation, but falter in real-world applications. Second, training on synthetic data can introduce biases into a system's output if the dataset is not carefully designed. Third, training an AI system on its own outputs can lead to a phenomenon known as "model collapse," where errors and biases become continuously amplified until the AI's outputs are no longer correct or useful. A recent Scientific American article likens this problem to the scramble to obtain low-radioactivity metal in the 20th-century:

> The possibility of AI models tainting themselves may be a bit analogous to a certain 20th-century dilemma. After the first atomic bombs were detonated at World War II's end, decades of nuclear testing spiced Earth's atmosphere with a dash of radioactive fallout. When that air entered newly-made steel, it brought elevated radiation with it. For particularly radiation-sensitive steel applications, such as Geiger counter consoles, that fallout poses an obvious problem: it won't do for a Geiger counter to flag itself. Thus, a rush began for a dwindling supply of low-radiation metal. Scavengers scoured old shipwrecks to extract scraps of prewar steel. Now some insiders believe a similar cycle is set to repeat in generative AI—with training data instead of steel.
>
> Researchers can watch AI's poisoning in action. For instance, start with a language model trained on human-produced data. Use the model to generate some AI output. Then use that output to train a new instance of the model and use the resulting output to train a third version, and so forth. With each iteration, errors build atop one another. The 10th model,

---

[6] Cade Metz, Cecilia Kang, Sheera Frenkel, Stuart A. Thompson and Nico Grant, "How Tech Giants Cut Corners to Harvest Data for A.I.," *New York Times,* Apr. 6, 2024, https://www.nytimes.com/2024/04/06/technology/tech-giants-harvest-data-artificial-intelligence.html.

prompted to write about historical English architecture, spews out gibberish about jackrabbits.[7]

Model collapse will become of a more pressing issue as more of the internet's content is AI-generated. The intentional use of synthetic data to train AI may expedite this process.

6) **What this bill would do.** This bill would require developers of AI systems and services to publicly disclose specified information about the datasets used to train, test, and validate their models.

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

> Artificial Intelligence has become nearly unavoidable in Californians' daily lives, with new exciting generative AI tools being introduced daily, and the companies who make up the cornerstones of our digital lives either adopting AI or identifying their existing tools as falling under the AI umbrella. However consumer confidence in AI systems has not grown at the same rapid pace as industry adoption. Many consumers have valid questions about how these AI systems and services are created, and if they truly are better than what they seek to replace.
>
> To build consumer confidence we need to start with the foundations, and for AI that is the selection of training data. AB 2013 provides transparency to consumers of AI systems and services by providing important documentation about the data used to train the services and systems they are being offered, including if synthetic data has or is being used to fill gaps in data sources.
>
> Consumers may use this knowledge to better evaluate if they have confidence in the AI system or service, compare competing systems and services, or put into place mitigation measures to address any shortcomings of the particular system or service.

8) **Analysis**. Overall, this bill would impose modest requirements on developers of AI systems and services in exchange for providing Californians with a fuller understanding of the state's AI information ecosystem. This is consistent with the policy aims of a related bill this Committee recently passed, AB 3204 (Bauer-Kahan), which would require organizations that train AI using personal information to register with the California Privacy Protection Agency. The two bills are broadly compatible: AB 3204 would create a registry to identify the universe of entities that train AI with personal information, while this bill would require more specific disclosures by a subset of those entities to better understand the types of data being used to train AI.

An industry coalition takes an "oppose unless amended" position on AB 2013; their opposition letter, penned by the California Chamber of Commerce, describes four perceived weaknesses of the bill:

> *Opposition claim #1:* AB 2013 should clearly delineate what is and is not considered "training" and narrow the scope of AI systems or services subject to these transparency measures to high-risk AI systems.

---

[7] Rahul Rao, "AI-Generated Data Can Poison Future AI Models," *Scientific American,* Jul. 28, 2023, https://www.scientificamerican.com/article/ai-generated-data-can-poison-future-ai-models/

The bill defines training as "testing, validating, or fine tuning the artificial intelligence system or service." The author may wish to instead tie the definition of training to the provided definition of artificial intelligence; for example, AB 3204 defines training to mean "exposing artificial intelligence to data in order to alter the relationship between inputs and outputs." The requirement to describe datasets used to test, validate, or fine-tune an AI system or service could be placed elsewhere in the bill, as they are not fundamental to "training." The author may also wish to provide a definition for "fine-tune" somewhere in the bill.

> *Opposition claim #2:* AB 2013's definition of "developer" is both overbroad and vague and should include guardrails to address compliance challenges.

The bill defines "developer" to mean "a person, partnership, state or local government agency, or corporation that designs, codes, or produces an artificial intelligence system or service, or substantially modifies an artificial intelligence system or service for use by a third party for free or for a fee." This definition is broad, but not vague: any entity who produces or substantially modifies an AI system or service through training is required to provide information about the datasets they used.

> *Opposition claim #3:* AB 2013 should not apply to AI systems and services that were in use prior to the bill's effective date.

AB 2013 is oriented towards consumer protection, and it is not clear why products that are already available to Californian consumers should be broadly exempted from the bill's requirements. The coalition letter specifies one particular descriptor that may be genuinely hard for developers to comply with: the "dates the dataset was first and last used during the development of the system or service." Developers of existing systems, unaware of this bill's provisions, may simply have no record of when training began and ended for particular datasets. The author may wish to include a narrow exemption for existing systems on this point.

> *Opposition claim #4:* AB 2013 should expressly preclude any private right of action.

This bill does not outline specific enforcement mechanisms for its provisions. In their absence, enforcement will likely occur on the basis of California's Unfair Competition Law.[8] It is not clear why the bill would need to specifically preclude a Private Right of Action. Bills are generally written to permit or require specific actions, rather than to laboriously outline actions that are not meant to be taken.

Writing in support of the bill, Oakland Privacy describes the importance of training data transparency:

> Visibility of data sources is one transparency factor in a number of multi-factor AI transparency models being designed. While it is far from the only one (and many focus on AI explainability and AI auditability as crucial factors as well), the ubiquitous presence of data set transparency on all of these models indicates the consensus that responsible AI always includes disclosure of the data sets used to train the system.

---

[8] Bus. & Prof. Code § 17200 *et seq*.

> One reason data set transparency is important is to root out illegally or unethically sourced data sets. A notorious example is Clearview AI which scraped social media to establish a database now described as comprising almost 20 billion images – larger than the population of the earth. This data was publicly available, but protected by community standards policies which forbade large-scale scraping for profit. Similarly, industries like journalism (NY Times v OpenAI) and playwrights and screenwriters are exploring the legal limits of AI data set use. While the laws that will come to govern the use of data sets for artificial intelligence are wildly unsettled at the moment, a robust transparency mandate can be informative for ongoing efforts to set limits and determine the rules of the road going forward, as well as identifying rogue players in the system crossing lines that we decide should not be crossed. Whatever our society comes to decide constitutes acceptable use, it is unlikely to be "everything and anything".

Oakland Privacy also points out perceived challenges for AB 2013:

> We think the challenges in AB 2013 will be defining the formats for the disclosures. How this is done is likely dependent on the intentions of the bill. If the intent is for the disclosed information to be accessible to developers, engineers and scientists, then highly technical disclosures may meet the intention of the bill, and allow for some checks and balances. However, if the intent is to provide reasonably tech-savvy members of the public with actionable information, then we expect there will have to be some prescriptions regarding disclosure formats to avoid such highly technical and abstruse disclosure documents that they are virtually useless to anyone but a highly trained engineer.

9) **Committee amendments.** Two proposed committee amendments would clarify and adjust the scope of this bill. First, the definition of "artificial intelligence system or service" would be replaced with the definition for "artificial intelligence" in other AI-related bills that have passed through this committee:

> ~~(a) "Artificial intelligence system or service" or "system or service" means a machine-based system or service that can, for a given set of human-defined objectives, generate content and make predictions, recommendations, or decisions influencing a real or virtual environment.~~
>
> ***(a) "Artificial intelligence" means an engineered or machine-based system that varies in its level of autonomy and that can, for explicit or implicit objectives, infer from the input it receives how to generate outputs that can influence physical or virtual environments.***

Second, the bill in print requires developers of AI systems to provide a wealth of specific information related to each dataset used to train a model. Each individual description required by this bill may not constitute proprietary information, but their combination may unintentionally allow competitors to infer the precise data that goes into the training of a compliant developer's AI system or service. To avoid this issue, a proposed amendment would require a "high level summary of the datasets" used to develop a given system or service:

> (a) A ~~description~~ ***high-level summary*** of ~~each dataset~~ ***the datasets*** used in the development of the system or service, including, but not limited to…

10) **Related legislation.** AB 3204 (Bauer-Kahan, 2024) would require organizations that train AI using personal information to register with the California Privacy Protection Agency. In registering, organizations would be required to pay an annual fee and disclose the categories of

personal information they use to train AI. This bill is currently pending in the Appropriations Committee.

ARGUMENTS IN SUPPORT:

Oakland Privacy writes:

> From a public transparency view, this is basic and minimal information that the public is entitled to in order to be able to understand the uses to which their personal information may be applied, the potential efficacy of the AI system, and what its output is predicated upon.

Secure Justice writes:

> We believe AB 2013 is a pragmatic proposal that will greatly increase public awareness into the data sets being used to train artificial intelligence models, and further does not impose an undue burden on the developers subject to such a requirement.

Concept Art Association writes:

> AB 2013 lays an imperative first stone on the path to protection for all Californians, and it is a solid concrete action we can take right now to begin bringing some transparency to what so far has mostly been an unscrupulous, opaque and predatory model for data acquisition.

ARGUMENTS IN OPPOSITION:

California Chamber of Commerce, taking an "oppose unless amended" position, writes on behalf of an industry coalition:

> Unfortunately, as currently drafted, we have significant concerns with the approach taken in AB 2013, and specifically around overburdensome mandates, the technical feasibility of the bill's transparency measures (including its assignment of responsibilities and the unique challenges presented for different types of developers in meeting the standards set in this bill), insufficient clarity around key terms, and potential exposure to liability. Moreover, we are heavily concerned about AB 2013's failure to provide protections for trade secrets and intellectual property, though we do not believe that is the intended outcome of this bill. While it may not be obvious on its face, the expertise and judgment as well as selection of data and datasets is part of what differentiates providers, thereby causing significant concern among developers as to the potential of this bill to undermine their intellectual property and harm competition. And lastly, we question whether the disclosure of training data will result in any substantial benefit when it comes to determining an AI model's performance for a particular use case. Stated another way, simply because a model has been trained on certain data does not mean it will perform as needed in a specific use case.

Chamber of Progress writes:

> Requiring online platforms to disclose data used to train their artificial intelligence (AI) systems and services on their website stifles competition in the digital marketplace. A healthy competition marketplace is essential to ensure better quality of services for consumers and encourages platforms to innovate. The disclosure requirement risks revealing important

business information and strategies. Additionally, the inclusion of "but not be limited to" in such requirements makes the expectations placed on online platforms unclear.

REGISTERED SUPPORT / OPPOSITION:

Support

Concept Art Association
Oakland Privacy
Santa Monica Democratic Club
Secure Justice
Transparency Coalition.ai

Opposition

Chamber of Progress

Oppose Unless Amended

California Bankers Association
California Chamber of Commerce
California Land Title Association
Insights Association
National Association of Mutual Insurance Companies
Personal Insurance Federation of California
Software & Information Industry Association
Technet

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