FUTURE OF PRIVACY FORUM

1350 Eye Street NW, Suite 350, Washington, DC 20005 | 202-768-8950 | fpf.org

July 13, 2026

Via Electronic Submission

Attorney General Phil Weiser
Office of the Attorney General
Colorado Department of Law
1300 Broadway, 10th Floor
Denver, CO 80203

Re: Pre-Rulemaking for the Automated Decision-Making Technology (Senate Bill 26-189) and Chatbot Safety (House Bill 26-1263) Acts

Dear Attorney General Weiser and Staff,

Thank you for your work and the opportunity to provide input on the pre-rulemaking process concerning Senate Bill 26-189 (the "Colorado Automated Decision-making Act") and House Bill 26-1263 (the "Chatbot Safety Act"). FPF is a global non-profit organization dedicated to advancing privacy leadership, scholarship, and principled data practices.[1] We seek to support balanced, informed public policy and equip regulators with the resources and tools needed to craft effective regulation.

In light of the five principles the Department of Law (the Department) outlines regarding the need to balance consumer rights, clarify ambiguities, facilitate compliance, harmonize, and allow for innovation, we write to raise several areas of consideration for the proposed rulemaking.

First, consistent with the Department's goals of clarifying ambiguities and harmonizing regulatory approaches, we highlight several recommendations to ease points of tension between the Colorado Automated Decision-making Act ("ADM Act") and the Colorado Privacy Act ("CPA"), including:

1. Aligning the scope and definition of automated decision-making technology;
2. Clarifying transparency obligations to enable compliance under both laws; and
3. Streamlining consumer rights

[1] The views reflected herein do not necessarily reflect the views of the FPF Advisory Board or FPF members.

For the Chatbot Safety Act, we outline three additional recommendations for the Department's consideration to improve the law's clarity and implementation:

1. Clarifying key exemptions and terms, including the exemption for services limited to a "narrow and discrete topic";
2. Ensuring age estimation rules are both flexible and interoperable with the recently enacted Digital Age Assurance Act; and
3. Specifying rules regarding the creation and implementation of suicide and self-harm crisis intervention protocols

---

I. The Department Should Clarify Points of Tension and Conflict Between the Colorado ADM Act and the Privacy Act

The current text of the ADM Act contains terminology, definitions, and approaches to AI and data governance and consumer rights that are inconsistent with those in the Colorado Privacy Act ("CPA"). Importantly, both laws aim to address two distinct yet intertwined sets of challenges. Enacted in 2021, the CPA aims to promote data protection and safeguard consumer digital privacy. Like many state comprehensive consumer data privacy laws, the CPA includes provisions regarding profiling (both automated and not) in furtherance of decisions that carry a legal or similarly significant effect. These common profiling provisions govern the personal data processing behind automated decision-making (ADM).

The Colorado ADM Act, in contrast, governs the automated decision-making technology itself. It promotes transparency in ADM use and provides consumers enumerated rights when such technology materially influences a consequential decision that impacts them. As a result, many entities will be subject to requirements under both laws, and consumers will exercise rights under both. It is therefore important that the two regimes operate as clearly and cohesively as possible, including by:

- Aligning terminology and definitions related to ADM;
- Ensuring sufficient transparency for controllers and deployers to fulfill their obligations; and
- Aligning and streamlining consumer rights where possible.

Colorado, and the Department in particular, continues to be a leader in both data privacy and AI governance. Given that role, this rulemaking carries weight beyond Colorado's borders. As more states legislate on ADMT and privacy—separately, jointly, or both—how the Department resolves the inconsistencies described above could shape how ADMT and privacy law are reconciled across the country.

A. The Scope and Definition of Automated Decisionmaking Should Align with the CPA

The ADM Act's framework for determining which automated decision-making technologies (ADMT) fall within scope should align with the CPA's approach to profiling. At minimum, the Department should clarify when a covered ADMT under the Act also constitutes regulated profiling under the CPA, and when it does not. Without alignment, covered entities would need to run unnecessary parallel compliance analyses under two differently defined standards, risking inconsistent outcomes for functionally identical processing. Similarly, an automated decision that produces a legal or similarly significant effect on a consumer that qualifies as regulated under one law but not the other could result in that consumer being denied protections — such as the right to explanation, opt-out, or appeal — undermining the efficacy of the two laws' protections as a whole.

Four key risk factors determine whether an ADMT falls within the scope of either law, though the laws' treatment of these factors shows both similarities and differences that should be considered:

1. **Personal Data:** Whether personal data is processed;
2. **Judgment:** The degree of independent judgment the ADMT exercises, i.e., the extent to which it operates without human involvement, and, where a human is involved, whether that review is meaningful;
3. **Automation Threshold:** The role of the ADMT to the final decision, i.e., whether its output serves merely as an input to, or is dispositive of, the ultimate decision; and
4. **Consequentiality:** The significance of the decision's impact, i.e., the scope of covered decisions or domains.

Because both laws require personal data processing as a threshold requirement, using the same definition of “personal data,” the analysis below focuses on distinctions between the remaining three factors.[2] A high-level comparison between the laws’ definitions across these factors is provided below in **Table 1**.

First, unlike the CPA, the ADM Act does not account for a covered ADMT’s degree of independent judgment in determining whether it falls within scope. This distinction is increasingly important as AI evolves into generative and agentic systems used in employment and other consequential decision-making contexts. Not all AI systems exercise the same degree of independent judgment: agentic systems, for example, can operate autonomously across multiple steps, taking sequential actions with limited human intervention. In contrast, a predictive system with meaningful human involvement presents less risk: a reviewer who understands, evaluates, and can override the output serves as a checkpoint, catching errors before they reach the

[2] The Colorado Privacy Act defines “personal data” as information that is linked or reasonably linkable to an identified or identifiable individual, and does not include de-identified data or publicly available information. (Sec. 6-1-1303(17)). The Colorado ADM Act states that “personal data” has the same meaning as set forth in the Privacy Act. (Sec. 6-1-1701(16)).

consumer. The CPA accounts for this difference in risk profile through three tiered definitions laid out in the Department's Privacy Act Regulations:[3]

- "**Solely Automated Processing**" means the automated processing of Personal Data with no human review, oversight, involvement, or intervention.
- "**Human Reviewed Automated Processing**" means the automated processing of Personal Data where a human reviews the automated processing, but the level of human engagement does not rise to the level required for Human Involved Automated Processing. Reviewing the output of the automated processing with no meaningful consideration does not rise to the level of Human Involved Automated Processing.
- "**Human Involved Automated Processing**" means the automated processing of Personal Data where a human (1) engages in a meaningful consideration of available data used in the Processing or any output of the Processing and (2) has the authority to change or influence the outcome of the Processing.

The Department should provide guidance on whether, and how, a comparable approach applies under the ADM Act. The California Privacy Protection Agency's (CalPrivacy) California Consumer Privacy Act (CCPA) regulations, for example, require that in-scope ADMT (inclusive of profiling) "replace" or "substantially replace" human decision-making.[4] The regulations, in turn, state that ADMT replaces human decisionmaking if there is not a human involved in the decisionmaking that (A) knows how to interpret and use the technology's output to make the decision; (B) reviews and analyzes that output, along with any other information relevant to the decision; and (C) has the authority to make or change the decision.[5] This three-part test closely parallels the CPA's own approach in tiering automated processing into solely automated, human reviewed, and human involved.

Second, the ADM Act requires that a covered ADMT "materially influence" a consequential decision, in contrast to the CPA's requirement that automated profiling be "in furtherance" of a decision with a legal or similarly significant effect. While the ADM Act defines "materially influence" as a non-de minimis factor that meaningfully alters how a consequential decision is made, this remains an untested standard that will require case-by-case interpretation to apply consistently. Existing thresholds already in use elsewhere in privacy law offer covered entities and regulators alike a more predictable basis for compliance. The Department should consider aligning the ADM Act with the CPA's own standard or with another state's tested threshold, such as California's standard that an ADMT "replace or substantially replace" human decision-making.

Lastly, the ADM Act's scope of covered decisions is both broader and narrower than the CPA's, creating uncertainty for covered entities and consumers alike. On the narrower end, several of the ADM Act's "covered domains" are defined more restrictively than their CPA counterparts. In the employment context, for example, the CPA applies to any employment opportunity, while the

[3] 4 Code of Colorado Regulation 904-3, Rule 2.02
[4] Cal. Code Reg. tit. 11, § 7001(e).
[5] *Id.*

ADM Act limits covered decisions to those "that create[ ] or may create an employer-employee relationship." On the broader end, the ADM Act extends to any decision that "relates to" the provision, price, or terms of a covered domain, whereas the CPA requires that the decision have a legal or similarly significant effect and result in the provision or denial of a listed domain.[6] These two provisions can point in opposite directions for the same use case: an ADMT used to identify employees for termination does not itself create an employer-employee relationship, yet arguably "relates to" the terms of that relationship—leaving covered entities and affected consumers uncertain whether the ADM Act applies. The Department should clarify these boundaries and, as CalPrivacy has done, consider adopting illustrative examples to guide compliance.[7]

Table 1. Comparison of Covered ADMT Between the Colorado ADM Act and Privacy Act
Key: **Bold** = Defined Term; underline = key element or test

	Colorado ADM Act	Colorado Privacy Act
Covered ADMT	**"Automated decision-making technology" or "ADMT"** means a technology that processes personal data and uses computation to generate output, including predictions, recommendations, classifications, rankings, scores, or other information that is used to make, guide, or assist a decision, judgment, or determination concerning an individual. (Sec. 6-1-1701(2)(a)). "**Covered ADMT**" means automated decision-making technology that is used to materially influence a consequential decision. (Sec. 6-1-1701(5)).	**"Profiling"** means any form of automated processing of personal data to evaluate, analyze, or predict personal aspects concerning an identified or identifiable individual's economic situation, health, personal preferences, interests, reliability, behavior, location, or movements. (Sec. 6-1-1303(20)). Relevant CPA requirements and rights are triggered when profiling is "in furtherance of a decision that produces a "legal or other similarly significant effect concerning a consumer."
Degree of Judgment	N/A	**"Human Involved Automated Processing"** means the automated processing of Personal Data where a human (1) engages in a meaningful consideration of available data used in the Processing or any output of the Processing and (2) has the authority to change or influence the outcome of the Processing. (CPA Rule 2.02)

[6] The ADM Act's predecessor, the Colorado AI Act (SB 205), more closely paralleled the CPA's text by defining "consequential decision" as "any decision that has a material, legal or similarly significant effect on the provision or denial to any consumer of, or the cost or terms of: (A) Education enrollment or education opportunity; (B) Employment or employment opportunity; (C) Financial or lending services; (D) Essential government services; (E) Healthcare service; (F) Housing, (G) Insurance, or (H) Legal services."
[7] *See* "Fact Sheet: Draft Automated Decisionmaking Technology Regulations," CalPrivacy.

		"Human Reviewed Automated Processing" means the automated processing of Personal Data where a human reviews the automated processing, but the level of human engagement does not rise to the level required for Human Involved Automated Processing. Reviewing the output of the automated processing with no meaningful consideration does not rise to the level of Human Involved Automated Processing. (CPA Rule 2.02) **"Solely Automated Processing"** means the automated processing of Personal Data with no human review, oversight, involvement, or intervention. (CPA Rule 2.02)
Automation Threshold	"**Materially influence**" means: (I) An ADMT output is a non-de minimis factor that is used in making a consequential decision; and (II) An ADMT output affects the outcome of a consequential decision, including by constraining, ranking, scoring, recommending, classifying, or otherwise meaningfully altering how a consequential decision is made. "Materially influence" does not include incidental, trivial, or clerical uses.	**Profiling** must be "in furtherance of decisions that produce legal or other similarly significant effects concerning a consumer" for certain automated profiling requirements to be triggered.
Covered Decisions	**"Consequential decision"** means: (I) A decision, determination, or action made about a consumer that relates to the provision of or a consumer's access to, eligibility for, selection for, or compensation for a covered domain; or (II) A decision, determination, or action about a consumer that relates to a differentiated price, cost sharing, compensation, or other material terms in a manner that is reasonably likely to materially limit, delay, effectively deny, or otherwise fundamentally alter the consumer's access, eligibility, or opportunity for a covered domain. (Sec. 6-1-1701(3)(a) **"Covered domain"** means: an education enrollment or an education opportunity;	**"Decisions that produce legal or similarly significant effects concerning the consumer"** means decisions that result in the provision or denial of financial or lending services, housing, insurance, education enrollment or opportunity, criminal justice, employment opportunities, health care services, or access to essential goods or services. (Sec. 6-1-1303(10)).

	employment or an employment opportunity that creates or may create an employer-employee relationship; the lease or purchase of residential real estate in Colorado; a financial or lending service; insurance, including underwriting, pricing, coverage, claims adjudication, or other determinations that materially affect access to benefits; health-care services; or essential government services and public benefits, including eligibility and renewal determinations. (Sec. 6-1-1701(6)).	

B. Developer's Transparency Obligations Should Enable Controllers and Deployers To Comply with the ADM Act and CPA

Because many entities are both a controller under the CPA and a deployer under the ADM Act for their use of ADMT, developer disclosures must be specific and comprehensive enough to encompass the information required for deployers under the ADM Act and controllers using profiling under the CPA. These controller-deployers must also ensure that their ADMT does not violate existing anti-discrimination law, as addressed in Section 1707 of the ADM Act. For many entities, compliance across all three regimes for a single ADMT would ideally occur through one unified process, which depends, in part, on the information ADMT developers provide.[8] However, only the ADM Act imposes developer transparency requirements; and the scope of information developers must provide does not match what controllers and deployers need for compliance. This mismatch leaves deployers—particularly small businesses with fewer resources to independently fill these gaps—unable to demonstrate compliance.

To comply with the CPA, for example, Rule 9.03 requires a controller to know the logic used in the profiling process and whether, and to what extent, the system has been evaluated for accuracy, fairness, or bias.[9] Similarly, to allocate responsibility appropriately between developers and deployers under anti-discrimination law, as detailed in Section 1707 of the ADM Act, developers must supply information regarding testing and evaluation for demographic differentials. Consider a developer of an ADMT used to evaluate college or university applicants for admission: the developer would need to disclose not only the general criteria or logic the system applies to score applicants, but also the results of any testing for disparate impact across protected characteristics such as race, sex, or age—information an admissions office deploying the tool needs both to satisfy the CPA's disclosure requirements and to defend its own compliance with anti-discrimination law.

---

[8] *See "Best Practices for AI and Workplace Assessment Technologies," at 7-8, Future of Privacy Forum (Sept. 2023) (explaining role-specific transparency between developers and deployers).*

[9] 4 Code of Colorado Regulation 904-3, Rule 9.03.

The current text of the ADM Act requires only that a developer provide deployers a general statement addressing intended use, known risks, a summary of training data, known limitations, and instructions on monitoring and human review. As the Department considers implementing regulations, one interpretation worth exploring is whether "known risks" should encompass ADMT risks and measures required under the CPA, as well as any information needed to satisfy Section 1707's anti-discrimination allocation of responsibility.

C. Consumer Rights Under Both Laws Should Be Streamlined

For consumers to be able to effectively exercise their rights under both the ADM Act and CPA, and for controllers/deployers to be able to efficiently create a single data subject request (DSR) process, consumer rights under both laws should be aligned and streamlined. Though the two laws provide similar rights, they differ in three key ways: (1) the amount and timing of transparency required for ADM use; (2) when rights are triggered; and (3) whether consumers are afforded a right to opt out of, or to appeal, a decision.

First, the CPA and ADM Act diverge in what transparency they require, and when. The CPA requires controllers to disclose, before profiling occurs, what decisions are subject to profiling, what personal data is used, how the profiling factors into the decision-making process (including the role of human involvement), and the logic used in the profiling.[10] The ADM Act, by contrast, requires only that developers and deployers provide advance notice that a covered ADMT is in use, along with instructions on how to request additional information. Only after a covered ADMT materially influences an adverse consequential decision does a consumer become entitled to more information, such as the role of the ADMT in the decision. As a result, the same use of ADMT can afford consumers markedly different amounts of transparency. The Department should consider clarifying how these disclosure requirements interact to help ensure consumers receive a consistent baseline of transparency across both laws and that controllers/deployers can rely on a consistent disclosure template.

Second, the CPA and ADM Act diverge in who can exercise a correction right and when it becomes available. Under the CPA, consumers have the standard privacy rights to confirm whether a controller is processing their personal data, access that data, correct inaccuracies in it, and delete it. The ADM Act, by contrast, incorporates the CPA's right to correct inaccurate personal data used by a covered ADMT, even where the deployer is not itself a controller under the CPA, but *only* once the ADMT has materially influenced an adverse consequential decision. As a result, the same right for a singular ADM process is available immediately under the CPA once personal data is processed for profiling, but only after an adverse decision under the ADM Act. The Department should therefore consider clarifying how this shared right applies consistently across both laws, to help ensure consumers can exercise a single right of correction

[10] 4 Code of Colorado Regulation 904-3, Rule 9.03

in these instances and that controllers/deployers can implement a singular correction process for the ADM activity.

Lastly, the CPA and ADM Act provide different consumer rights to opt out of, or appeal, automated decision-making. Under the CPA, a consumer may opt out of profiling in furtherance of a decision that carries a legal or similarly significant effect, provided the profiling is solely automated or human-reviewed (human-involved profiling carries no such right).[11] The ADM Act, by contrast, requires consumers to be given the opportunity for meaningful human review and reconsideration when an ADMT materially influences an adverse consequential decision. These rights serve different purposes and typically are not treated as cumulative: opt-out is a preventive right that lets a consumer avoid automated profiling before any decision is made, while appeal is a remedial right that lets a consumer challenge an automated decision's outcome after the fact. Because both rights address the same underlying risk—that a decision was made without adequate human judgment—requiring both simultaneously for the same activity would be duplicative rather than protective. To reconcile these rights, California's approach offers a potentially useful model, where covered entities may offer either an opt-out or an appeal for automated decision-making, but must offer an appeal if they do not offer consumers an opt-out right.[12]

II. For the Chatbot Safety Act, the Department Should Clarify Key Ambiguities Around Scope and Key Terms, Age Estimation, Safety Protocols, and Transparency Reporting

Because chatbot safety laws are relatively new, and the Department is the first U.S. regulator to issue implementing regulations on the topic, the Department is poised to provide critical clarifications on the scope and implementation of these laws. The first laws in this area were enacted only last year, first in California and New York in 2025, and other state laws in this space have not yet taken effect and/or been interpreted through rulemaking or enforcement.[13] As a result, Colorado has limited precedent to draw from in determining how undefined terms and novel obligations should operate in practice in this Act. At the same time, rulemaking offers Colorado the opportunity to provide influential guidance for other jurisdictions, specifically around common areas of ambiguity in scope and key terms, age assurance, and the implementation of protocols related to suicidal ideation and self-harm.

A. Key Exemptions and Terms Should Be Clarified

The Department's pre-rulemaking considerations for the Chatbot Safety Act appropriately identify "emotional dependence" as a key area where guidance may be useful, but other undefined phrases could also substantially affect the Chatbot Safety Act's scope. In particular, the Chatbot

[11] 4 Code of Colorado Regulation 904-3, Rule 9.03(B).
[12] Cal. Code Reg. tit. 11, art. 11.
[13] *See* Cal. Bus. & Prof. Code §§ 22601–22608 (2025); N.Y. Gen. Bus. Law §§ 1700–1704 (2025); *see also* Justine Gluck & Rafal Fryc, *2026 Chatbot Legislation Tracker*, Future of Privacy Forum, https://fpf.org/2026-chatbot-legislation-tracker/ (last visited July 13, 2026).

Safety Act exempts certain conversational AI services designed to provide outputs relating to a "narrow and discrete topic."[14] Depending on how that phrase is interpreted, the exemption could exclude only routine transactional tools, *or* it could more broadly sweep tools that are task-specific but still enable open-ended, personalized, or emotionally sensitive conversation.

The Department should clarify this exemption through illustrative examples and factors. A chatbot that answers billing questions, tracks orders, or schedules appointments may clearly be limited to a narrow and discrete topic. A general-purpose chatbot or open-ended wellness assistant likely not. Harder edge cases may include tutoring tools or workplace assistants that begin with a narrow purpose but can adapt to the user, remember prior interactions, or sustain broader conversation.

Other states have begun to use similar narrow-purpose carveouts, but the terms remain similarly undefined. Connecticut's chatbot law, SB 5, for example, excludes certain narrow, task-specific tools only where the tool's primary function is not to discuss mental health.[15] Colorado's Chatbot Safety Act takes a different approach: it exempts services designed to provide outputs relating to a narrow and discrete topic, but limits that exemption where the service produces sexually explicit outputs or dialogue concerning suicidal ideation, self-harm, or harm to others. Without further guidance, operators may face uncertainty about whether health-adjacent, wellness, education, or other support tools remain *within* the carveout when user interactions move into emotionally sensitive territory.

Clear examples would help preserve the Chatbot Safety Act's focus on open-ended companions while avoiding scoping in low-risk tools that do not implicate the same minor-safety concerns. Because similar carveouts are appearing in other state chatbot laws, Colorado's rulemaking could provide influential guidance on how to distinguish genuinely narrow tools from services that function more like open-ended conversational systems.

State Law	Narrow-purpose carveout language
Iowa SF 2417 / Idaho SB 1297 / Nebraska LB 525	Services designed to provide outputs relating to a narrow and discrete purpose
Connecticut SB 5	Narrow, task-specific tools, with limits where the primary function involves mental health discussion
Colorado HB 1263	Services designed to provide outputs relating to a narrow and discrete topic, subject to limits involving sexual content and suicide/self-harm dialogue

[14] Conversational Artificial Intelligence Service Operator Requirements, ch. 208, 2026 Colo. Sess. Laws 1197, § 6-1-1701(3.5)(b)(III).
[15] *See* Public Act No. 26-15, 2026 Conn. Acts (Reg. Sess.); *see also* Justine Gluck & Daniel Hales, *SB 5 in Five: What to Know About Connecticut's New AI Law*, Future of Privacy Forum (May 27, 2026), https://fpf.org/blog/sb-5-in-five-what-to-know-about-connecticuts-new-ai-law/.

Additional terms may also benefit from future rule-making guidance. For example:

- **"Emotional dependence:"** The Chatbot Safety Act requires operators to take reasonable measures to prevent responses that simulate emotional dependence, a term now appearing in several enacted and proposed chatbot laws, including in Connecticut, Iowa, Idaho, Washington, Nebraska, New York(pending), but without definition in each of these laws.[16] Guidance could help clarify whether the term is meant to capture clinical or addiction-like symptoms, such as difficulty disengaging or fear of abandonment, or a broader pattern in which a chatbot encourages substitution for real-world relationships and support.
- **"Romantic companionship:"** The Chatbot Safety Act refers to statements that simulate "romantic companionship," but does not define the term.[17] Guidance could help distinguish prohibited romantic or relationship-like outputs directed to minors from other interactions that may involve narrative content or general expressions of warmth or encouragement.
- **"Self-harm:"** The Chatbot Safety Act does define self-harm, albeit minimally, as intentional self-injury, with or without intent to cause death.[18] Guidance could help clarify how operators should treat adjacent or ambiguous user statements, including references to eating disorders, substance misuse, or non-imminent distress, particularly for purposes of crisis protocols and annual reporting.
- **False representation of professional services:** The Chatbot Safety Act prohibits operators from using terms or outputs suggesting that chatbot-generated output is provided by, endorsed by, or equivalent to services provided by licensed health-care professionals, legal professionals, mental health professionals, or dietitians.[19] Guidance could help clarify how this applies to informational tools that *summarize* legal, health, nutrition, or mental health information while avoiding explicit statements that they are a substitute for licensed professional advice.

B. Age Estimation Rules Should Be Flexible on Methodology and Interoperable with the Recently Enacted Digital Age Assurance Act

The Department's pre-rulemaking questions probe what factors should determine the commercial reasonableness of an operator's age-estimation approach, what evidence should establish a user's age or age range (and whether an operator has sufficiently established 'knowledge'), and, importantly, how age estimation rules can remain adaptable as new technologies surface novel

---

[16] 2026 Colo. Sess. Laws 1197, § 6-1-1708(2)(a)(III)(d). *See also* 2026 Conn. Pub. Acts No. 15 (Reg. Sess.); 2026 Iowa Acts, Senate File 2417; 2026 Idaho Sess. Laws 249 (establishing SB 1297); 2026 Wash. Sess. Laws 168 (establishing HB 2225); Neb. Rev. Stat. § 84-712.05 (as amended by LB525); and S9051-B, 2025-2026 Reg. Sess. (N.Y. 2026) (passed Legislature but pending Governor signature).
[17] 2026 Colo. Sess. Laws 1197, § 6-1-1708(2)(a)(III)(d)(II).
[18] *Id.* at § 6-1-1708(16.5).
[19] *Id.* at § 6-1-1708(5).

approaches and methodologies. In crafting rules responsive to these questions, the Department should prioritize: (a) flexibility in methodologies available to operators, while establishing clear standards for acceptable methods; and (b) interoperability with other state laws to best protect user privacy while promoting achievable compliance outcomes.

a. The Rules Should be Flexible

A flexible regulatory framework could give operators the agency to select the technologies best suited to their specific services and users while maintaining technical resilience as innovation continues to evolve. The rules mandate operators to age-estimate users, and that the estimated age or age range of that user is "knowledge" for purposes of implementing the Chatbot Safety Act's minor-specific protections. According to FPF's *Unpacking Age Assurance* infographic, the term "age estimation" typically refers to mechanisms that leverage AI-powered deductions of a user's likely age based on biological or behavioral traits (e.g., face, voice, or typing patterns).[20] Although categorized as technically separate for the infographic, a conceptually similar method for estimating a user's likely age is through "age inferences"—reasonable conclusions about a user's age range derived based on a user's behavior over time or other account history, including a user's in-service activity, financial transactions, and email address lifespan.[21] Both of these broader methodologies encapsulate several technologies and mechanisms for determining a user's likely age or age range and are subject to further change and innovation as this budding world of technologies advances. A regulatory framework that merely seeks to identify specific approved methods and technologies through bright-line rules may struggle to adapt to technological change, different services or data contexts, and consumer expectations across use cases.

The FTC's 'approved verifiable parental consent methods' list under the Children's Online Privacy Protection Act (COPPA) illustrates the stark tradeoffs of stagnant approved-methods lists. Under COPPA, when an operator is required to obtain VPC, they must implement a method that is "reasonably designed in light of available technology" to verify that a child's parent gives consent.[22] The FTC has determined that several methods meet the rule's standard and provided a non-exhaustive list of approved VPC methods.[23] Although operators do not need to *only* use approved methods, most do so out of an abundance of caution against liability if a novel chosen method is later found to be noncompliant.

One of the greatest criticisms of this approved-methods structure is that the methods of VPC currently provided within this framework are restrictive and quickly outdated when agency

[20] Future of Privacy Forum, *Unpacking Age Assurance: Technologies and Tradeoffs* (v2.0) (Jan. 27, 2026), https://fpf.org/wp-content/uploads/2026/01/FPF-Age-Assurance-v2.0.pdf.
[21] *Id.*
[22] Federal Trade Commission, *Verifiable Parental Consent and the Children's Online Privacy Rule*, (Accessed July 13, 2026), https://www.ftc.gov/tips-advice/business-center/privacy-and-security/verifiable-parental-consent-childrens-online-privacy-rule.
[23] Future of Privacy Forum, *The State Of Play: Is Verifiable Parental Consent Fit For Purpose?*, at 12 (Jun. 2023), https://fpf.org/wp-content/uploads/2023/06/FPF-VPC-White-Paper-06-02-23-final2.pdf.

rulemaking cannot keep up with technological change.[24] Even though the agency later added an 'approval mechanism' in hopes of expediting new methods for approval, authorized methods remain limited due to the time-consuming approval process, costs of participation, and a high rate of method rejection.[25] And the lapse in technological innovation allowed within those rules generates real implementation challenges—outdated methods relying upon excess data collection or misaligning with consumer expectations within modern services can create privacy incongruencies, confusion, and user drop-off.[26] These consequences demonstrate the limited utility of an 'approved methods' approach.

Comparatively, California and New York are in more advanced stages of rulemaking on compliant age assurance methodologies and may serve as useful reference points for the Department as it develops rules under the Chatbot Safety Act. Both states have taken slightly different approaches to permitting flexible age determinations by operators, including:

- **California:** provides a nonexhaustive list of example categories and methodologies of commercially reasonable age determinations, alongside criteria for determinations (i.e., reasonably effective at determining users under 18, consistently measurable, and quantifiably testable), and provisions allowing for age determinations under other state and federal laws to serve the requirements of the proposed regulations.[27]
- **New York:** pursuing an "accuracy benchmarking" framework that favors standards of accuracy over solely defining specific methods or technologies for compliance. The Rules leverage illustrative examples to demonstrate accuracy standards and identify noncompliant methods to define a regulatory floor—affording a flexible space for industry innovation within regulatory requirements.[28]

b. The Rules Should be Interoperable

The Department should ensure that age estimation rules promulgated under the Chatbot Safety Act are interoperable with the Digital Age Assurance Act to promote pragmatic compliance. In addition to enacting the Chatbot Safety Act this year, the legislature also enacted SB 51, or the "Digital Age Assurance Act" (DAAA).[29] The DAAA establishes a process by which an operating system (OS) provider generates a signal embedded with a user's age range data and that signal is shared with developers through the app store environment. Developers are then responsible for using that signal to implement minor-specific protections within services across all points of access. That is, an age signal received by a developer through the app store would then be associated with that particular user's account, whether accessed through an app or internet

[24] *Id* at 16-17.
[25] See *Id* at 11-15 as well as the accompanying infographic included therein.
[26] See e.g., *Ibid.*
[27] See *Notice of Proposed Rulemaking, Protecting Our Kids from Social Media Addiction Act Regulations*, California Department of Justice (published May 15, 2026), https://oag.ca.gov/system/files/media/sb976-proposed-rulemaking.pdf.
[28] https://ag.ny.gov/sites/default/files/regulatory-documents/safe-for-kids-act-nprm.pdf
[29] 2026 Colo. Sess. Laws 2086.

browser. As a result, developers subject to both laws may already be receiving and implementing user age signals that could serve as evidence for age determinations or otherwise satisfy the Chatbot Safety Act's age-estimation requirements.

As a comparative example, California's Department of Justice faces a similar dynamic in its ongoing SB 976 rulemaking, which must define how age assurance practices under the state's social media law operate alongside the signals received under the DAAA. To reconcile these two processes, the Department's proposed rules provide that using the age range identified by age signals shared with developers by app stores is a compliant age determination under those rules.[30] By allowing operators to rely upon DAAA signals, the Department protects user privacy by disincentivizing companies from collecting more user data than what may otherwise be needed to comply with SB 976's age assurance requirements since sufficient information may be available to developers via the app store age signals. To best guide company compliance practices under the Chatbot Safety Act, the Department should consider clarifying the relationship of age determinations under these two state laws to promote interoperability and user privacy.

C. Expectations Should be Clarified For Sufficient Safety Protocols That Balance Harm Detection and Privacy

The Chatbot Safety Act's requirements for implementing safety protocols raise important implementation questions about what a sufficient protocol should include, when the obligation is triggered, and what escalation entails when crisis indicators are repeated and/or severe. The Chatbot Safety Act requires covered operators to implement a safety protocol for responding to user prompts regarding suicidal ideation or self-harm, including referral to a crisis service provider and escalation procedures for repeated or severe crisis indicators.[31] Separately, the Act's annual reporting provision refers to protocols to detect and respond to instances of suicidal ideation or self-harm, as well as protocols to prevent chatbot responses about suicidal ideation or self-harm. However, there are technical limits to suicide and self-harm detection in chatbots that should be considered by the Department.

While systems may be more likely to identify explicit statements of intent, ambiguous or indirect expressions of suicidal ideation can be harder to reliably detect.[32] Guidance should therefore avoid implying that operators can determine *with certainty* whether a user is suicidal or at imminent risk. Instead, rules should focus on whether operators have reasonable, documented processes for identifying potential risk, testing and updating those processes, and minimizing unnecessary processing of sensitive mental health-related information.

---

[30] See California Department of Justice, *supra* note 21, at 2 (§ 560 permits operators to rely upon DAAA provided age signals for actual knowledge under SB 976).
[31] 2026 Colo. Sess. Laws 1197, § 6-1-1708(4).
[32] *See* Tong Li et al., *Can Large Language Models Identify Implicit Suicidal Ideation? An Empirical Evaluation*, in Findings of the Association for Computational Linguistics: EMNLP 2025, at 18392 (Ass'n for Computational Linguistics 2025), https://aclanthology.org/2025.findings-emnlp.998/.

That final challenge is especially difficult because crisis response requirements can push against privacy and data-minimization principles. To identify signs of crisis, operators may need to analyze highly sensitive user statements about mental health, relationships, violence, distress, or self-harm. Privacy principles, including provisions in the CPA, by contrast, generally favor collecting less data and limiting secondary uses of sensitive information.[33] As crisis response requirements become more detailed, lawmakers may need to clarify what data may be used to detect risk, how long it may be retained, and what limits apply after a crisis signal is identified.

Further, to assist covered operators in complying with the law, the Department could consider clarifying the minimum elements of a protocol, without prescribing a single approach. Relevant elements could include how operators identify prompts involving suicidal ideation or self-harm; what categories of crisis resources may be provided; how operators treat ambiguous or non-imminent harm statements; how protocols are tested, reviewed, and updated; and how operators handle sensitive user information.

Finally, the Department should clarify the meaning of "repeated or severe crisis indicators." Escalation procedures may differ where a user makes a single ambiguous statement of distress versus repeatedly expresses suicidal ideation or asks for instructions for self-harm.[34] Without clear drafting, operators may face uncertainty about both their legal obligations and the consequences of acting on sensitive user statements.

* * *

Thank you for this opportunity to provide comments on these proposed regulations. FPF welcomes any further opportunities to provide resources or information to assist in this important effort. If you have any questions regarding these comments and recommendations, please contact Tatiana Rice at trice@fpf.org.

Sincerely,

Tatiana Rice, Senior Director of Legislation and Regulation
Justine Gluck, AI Policy Analyst
Daniel Hales, Policy Counsel, U.S. Legislation and Regulation

[33] *See* 2026 Colo. Sess. Laws 1197, § 6-1-1308(3); and Jordan Francis, *Data Minimization's Substantive Turn: Key Questions & Operational Challenges Posed by New State Privacy Legislation*, Future of Privacy Forum (June 2025), https://fpf.org/wp-content/uploads/2025/06/FPF_Data-Minimization.pdf.
[34] *See* Anna Vinals Musquera & Scott Babwah Brennen, Between 988 and 911: A New Policy Framework for Violent Planning on AI Chatbots, Ctr. on Tech. Pol'y, N.Y. Univ. (June 22, 2026), https://techpolicynyu.org/between-988-and-911-a-new-policy-framework-for-violent-planning-on-ai-chatbots/.