People-First Chatbot Bill

Model Legislation

December 2025

Consumer Federation of America

epic.org ELECTRONIC PRIVACY INFORMATION CENTER

fairplay childhood beyond brands

AUTHORS

Ben Winters - Director of AI and Privacy, Consumer Federation of America
bwinters@consumerfed.org

Kara Williams - Counsel, Electronic Privacy Information Center
williams@epic.org

Brendan Bouffard - Staff Attorney, Fairplay
brendan@fairplayforkids.org

FUNCTIONAL NOTES

This model bill, when possible, uses existing legislative language from enacted laws and proposed bills throughout local, state, federal, and international legislatures.

Mindful of recent lawsuits from industry groups against laws designed to make technology safer, we drafted this model bill to withstand constitutional challenges.

Text in [brackets] indicates places where jurisdictions should add content specific to their localities.

If you have questions or are interested in assistance optimizing this bill for your needs, please reach out to Brendan Bouffard at Fairplay (brendan@fairplayforkids.org), Kara Williams at the Electronic Privacy Information Center (williams@epic.org), and Ben Winters at the Consumer Federation of America (bwinters@consumerfed.org).

EXECUTIVE STATEMENT

The People-First Chatbot Bill gives lawmakers a straightforward approach to address the harms caused by chatbot products that have been developed and deployed by tech companies with little oversight or transparency. It does not outlaw chatbots; it provides a workable, clear framework to encourage the development of safer technology.

The bill centers around a simple truth: Chatbots are products, not people. They're created and released by companies that must be held accountable when their products harm people, just like a charger that starts a fire, or faulty airbags that fail to deploy in a car crash.

This bill does not solve all chatbot harms—nor could any single bill—but it is a crucial step forward that would provide meaningful protections for all chatbot users.

While children are particularly susceptible to the harms of AI, this proposal intentionally protects *all* users and is designed to cover *all* chatbots. As recent lawsuits show, chatbots can cause devastating harm for people of all ages, including both children and adults. That is why we are endeavoring to make them safer for everyone.

The bill addresses key data privacy violations present in almost all commercially available chatbots and sets clear limits on the use of users' chat logs for harmful practices like targeted advertising, which Meta has already started doing. It restricts the use of personal data to significantly reduce the ability for chatbots to employ dangerous and manipulative companion-like features. It also requires clear notice that chatbots are not human and prohibits chatbots from representing they can provide qualified medical or legal advice.

This model bill complements existing authorities, including laws on unfair and deceptive acts and practices, product liability, negligence, false advertising, privacy, copyright, and criminal laws and gives enforcers another tool to use to combat ongoing harms from chatbots.

BILL HIGHLIGHTS

- Establishes that chatbots are products and outlines clear liability standards for injuries caused by the use of chatbots
- Prevents companies from using chats to target people with advertisements
- Limits companies' ability to use personal data or chats to profile users
- Requires companies to ensure their chatbots do not falsely represent that they can provide qualified medical, legal, or financial advice
- Bans companies from using minors' input data to train chatbots
- Prohibits companies from training chatbots on people's input data without their knowledge or consent by requiring them to obtain affirmative consent from adults over 18
- Requires companies to make clear explicitly and regularly that their chatbots are not people
- Gives people the right to sue companies for violating privacy protections, data security requirements, and disclosure provisions
- Grants government regulators the authority to enforce the law
- Prevents companies from providing law enforcement access to chats without the proper warrant
- Requires companies to publish key user safety metrics for their chatbots
- Restricts companies from using personal data from outside of chatbot interactions to inform chatbot outputs
- Empowers the Attorney General to adopt rules to keep the law up to date with developing technologies and business practices

DRAFT LEGISLATIVE TEXT

Section 1: Definitions

1) **Advertisement** means any written or oral statement, illustration, or depiction that promotes the sale or use of a good or service or is designed to increase interest in a brand, good, or service where such statement, illustration, or depiction is displayed in exchange for monetary or other valuable consideration, including access to data, between the chatbot provider and the brand, good, or service.

2) **Affirmative consent** means a clear affirmative act signifying a user's freely given, specific, informed, and unambiguous authorization for an act or practice in response to a specific request from a chatbot provider, provided:

    a) the request is provided to the user in a clear and conspicuous standalone disclosure;

    b) the request includes a description, written in easy-to-understand language, of the act or practice for which the user's consent is sought;

    c) the request is made in a manner reasonably accessible to and usable by users with disabilities;

    d) the request is made available to the user in each language in which the chatbot provider provides a chatbot;

    e) the option to refuse to give consent is at least as prominent as the option to give consent, and the option to refuse to give consent takes the same number of steps or fewer as the option to give consent; and

    f) affirmative consent to an act or practice is not inferred from the inaction of the user or the user's continued use of a chatbot provided by the chatbot provider.

    "Affirmative consent" does not include:

    a) acceptance of a general or broad terms of use or similar document;

    b) hovering over, muting, pausing, or closing a given piece of content;

    c) agreement obtained through the use of a false, fraudulent, or materially misleading statement or representation; or

    d) agreement obtained through the use of other dark patterns.

3) **Chatbot** means any artificial intelligence, algorithmic, or automated system that generates information via text, audio, image, or video in a manner that simulates interpersonal interactions or conversation.

4) **Chat log** means any input data, outputs generated by a chatbot, or record of the input data or outputs from user interactions with a chatbot.

5) **Chatbot provider** means any person creating, distributing, or otherwise making available a chatbot.

6) **Collect or collecting** means creating, buying, renting, gathering, obtaining, receiving, accessing, or otherwise acquiring personal data or input data by any means through individuals' use of chatbots.

7) **Dark pattern** means a user interface designed or manipulated with the substantial effect of subverting or impairing user autonomy, decision-making or choice, and includes, but is not limited to, any practice the Federal Trade Commission refers to as a "dark pattern."

8) **De-identified data** means information that cannot reasonably be used to infer or derive the identity of an individual or does not identify and is not linked or reasonably linkable to an individual or a device that identifies or is linked or reasonably linkable to such individual, regardless of whether the information is aggregated, provided that the chatbot provider:

    a) takes such physical, administrative, and technical measures as are necessary to ensure that the information cannot, at any point, be used to re-identify any individual or device that identifies or is linked or reasonably linkable to one or more individuals;

    b) publicly commits in a clear and conspicuous manner to:

        i) process, retain, or transfer the information solely in a de-identified form without any reasonable means for re-identification; and

        ii) not attempt to re-identify the information with any individual or device that identifies or is linked or reasonably linkable to an individual; and

    c) contractually obligates any entity that receives the information from the chatbot provider to:

        i) comply with all of the provisions of this paragraph with respect to the information; and

        ii) require that such contractual obligations be included in all subsequent instances for which the data may be received.

9) **Input data** means information, including text, photos, audio, video, or files, provided to a chatbot by a user.

10) **Model** means an engineered or machine-based system underlying a chatbot that can, for explicit or implicit objectives, infer from the input it receives how to generate outputs that can influence physical or virtual environments.

11) **Personal data** means any information, including derived data, inferences, or unique identifiers, that is linked or reasonably linkable, alone or in combination with other information, to an identified or identifiable individual or a device that identifies or is linked or reasonably linkable to an individual.

"Personal data" does not include de-identified data or publicly available information.

12) **Publicly available information** means information that has been lawfully made available to the general public from:

a) federal, state or municipal government records, if the person collects, processes, and transfers such information in accordance with any restrictions or terms of use placed on the information by the relevant government entity;

b) widely distributed media; or

c) a disclosure to the general public as required by federal, state, or local law.

"Publicly available information" does not include:

(a) any obscene visual depiction, as defined in section 1460 of title 18, United States Code;

(b) biometric data;

(c) personal data that is created through the combination of personal data with publicly available information;

(d) information that is collated and combined to create user profiles on publicly available or subscription-based websites and inferences generated from such information;

(e) genetic data, unless otherwise made publicly available by the individual to whom the information pertains;

(f) information made available by a user on a website or online service made available to all members of the public, for free or for a fee, where the user has restricted the information to a specific audience; or

(g) intimate images, authentic or computer-generated, known to be nonconsensual.

13) Process or processing means any operation or set of operations performed, whether by manual or automated means, on personal data or input data or on sets of personal data or input data, such as the use, storage, disclosure, analysis, deletion, or modification of such data.

14) Profiling means any form of processing performed on input data or personal data to detect and classify or designate personality and behavioral characteristics of an individual. "Profiling" does not include processing of chat logs for purposes of user safety or to otherwise comply with this Act.

15) Sell means exchanging personal data or input data for monetary or other valuable consideration, or making available such data or use of such data, by the chatbot provider to a third party.

"Sell" does not include:

(a) the disclosure of personal data or input data to a third party that processes the data on behalf of the chatbot provider;

(b) with the user's affirmative consent, the disclosure of personal data or input data where the user affirmatively directs the chatbot provider to disclose the data or intentionally uses the chatbot provider to interact with a third party; or

(c) the disclosure of personal data that the user:

(i) intentionally made available to the general public via a channel of mass media; and

(ii) did not restrict to a specific audience.

16) Training means the use of input data to adjust or modify a model. "Training" does not include:

a) testing to identify risks of harm to users;

b) adjustments or modifications to address identified risks of harm to users; or

c) any actions necessary to comply with this Act or otherwise required by law.

17) User means any natural person, regardless of age.

18) Widely distributed media means information that is available to the general public, including information from a telephone book or online directory, a television, internet, or radio program, the news media, or an internet site that is

available to the general public on an unrestricted basis; and does not include an obscene visual depiction (as defined in section 1460 of title 18, United States Code).

Section 2: Data Privacy and Security

1) A chatbot provider shall not:
    a) process personal data other than input data to inform chatbot outputs unless the processing of personal data is necessary to fulfill an express request made by a user and that user has provided affirmative consent;
    b) process a user's chat log:
        i) To determine whether to display an advertisement for a product or service to the user;
        ii) To determine a product, service, or category of product or service to advertise to the user; or
        iii) To customize an advertisement or how an advertisement is presented to the user;
    c) process a user's chat log or personal data:
        i) if the chatbot provider knows or should know, based on knowledge fairly implied on the basis of objective circumstances, that the user is under the age of [age based on state/lawmaker preference, 13 or 18], without the affirmative consent of that user's parent or legal guardian;
        ii) for training purposes, if the chatbot provider knows or should have known, based on knowledge fairly implied on the basis of objective circumstances, that a user is under 18 years of age;
        iii) of a user over 18 years of age for training purposes, unless the chatbot provider first obtains affirmative consent; or
        iv) to engage in profiling beyond what is necessary to fulfill an express request;
    d) use any classification or designation of a user's personality or behavioral characteristics created through profiling beyond what is necessary to fulfill an express request made by a user;
    e) sell a user's chat logs;

f) retain a user's chat log for longer than 10 years, unless retention is necessary to comply with this Act or otherwise required by law; or

g) discriminate or retaliate against any user, including by denying products or services, charging different prices or rates for products or services, or providing lower quality products or services to the user, for refusing to consent to the use of chat logs or personal data for training purposes.

2) A user has the right to access, at any time, any of the user's own chat logs that a chatbot provider has retained in a portable and readily usable format.

   a) Chat logs must be made available to users in a downloadable and human- and machine-readable format.

   b) A chatbot provider shall not discriminate or retaliate against any user, including by denying products or services, charging different prices or rates for products or services, or providing lower quality products or services to the user, for accessing their own chat logs.

3) A government entity shall not compel the production of or access to input data or chat logs from a chatbot provider, except as pursuant to a wiretap warrant [insert relevant wiretap warrant standard].

4) A chatbot provider shall develop, implement, and maintain a comprehensive data security program that contains administrative, technical, and physical safeguards that are proportionate to the volume and nature of the personal data and chat logs maintained by the chatbot provider. The program shall be written and made publicly available on the chatbot provider's website.

Section 3: Transparency for Users

1) A chatbot provider shall not use any term, letter, or phrase in the advertising, interface, or outputs of a chatbot that indicates or implies that any output data is being provided by, endorsed by, or equivalent to those provided by:

   a) a licensed healthcare professional;

   b) a licensed legal professional;

   c) a licensed accounting professional;

   d) a certified financial fiduciary or planner; or

   e) another licensed or certified professional in the jurisdiction [insert other

professions that are licensed within the jurisdiction here with references to the jurisdiction's licensing laws].

This includes any representation that a user's input data or chat log is confidential. Violations of this subsection are violations of [relevant professional licensing laws, unfair and deceptive trade practice laws, and false advertising laws.]

2) Chatbot providers shall provide clear, conspicuous, and explicit notice to users that they are interacting with a chatbot rather than a human prior to the chatbot generating any outputs, every hour thereafter, and each time a user prompts the chatbot about whether they are a real person.

    a) The text of this notice must appear in the same language as the one in which the user is interacting with the chatbot, in a font size easily readable by an average user, and no smaller than the largest font size of other text appearing on the interface on which the chatbot is provided.

    b) This notice must be accessible to users with disabilities.

    c) This notice must comply with regulations promulgated by the [Attorney General or other appropriate entity] as described in Section 6 of this Act.

Section 4: Assessments and Transparency Requirements

1) A chatbot provider shall assess its chatbot for risks of harm to users on a monthly basis, according to metrics as set forth in rules promulgated by the [Attorney General or other appropriate entity] and shall mitigate any risks of harm as set forth in rules promulgated by the [Attorney General or other appropriate entity] may require;

2) A chatbot provider shall make information about its chatbot publicly available on its website on a monthly basis as set forth in rules promulgated by the [Attorney General or other appropriate entity].

Section 5: Rulemaking

1) The [Attorney General or other appropriate entity] shall promulgate rules or regulations:

    a) describing the form and content of the disclosures required under Section 3 of this Act;

b) providing an example template for the disclosures required under Section 3 of this Act;

c) describing risks of harm to users and the metrics that each chatbot provider shall use to assess its chatbots for these risks of harm to users under Section 4 of this Act;

d) identifying and describing categories of information that each chatbot provider must make publicly available about its chatbots under Section 4 of this Act; and

e) updating annually the inflation-adjusted damages amount, consistent with the Consumer Price Index, as set forth in Section 8(3)(a).

2) The [Attorney General or other appropriate entity] may promulgate any other rules or regulations necessary to implement this Act.

Section 6: Severability and Construction

1) If any provision of this title, or the application thereof to any person or circumstance, is held invalid, the remainder of this title, and the application of such provision to other persons not similarly situated or to other circumstances, may not be affected by the invalidation.

2) Nothing in this Act preempts or otherwise affects any right, claim, remedy, presumption, or defense available at law or in equity, including but not limited to anti-discrimination, consumer protection, labor, and civil rights laws.

Section 7: Liability

1) Chatbots are products for the purposes of product liability actions.

2) A chatbot provider has a duty to ensure that the use of its chatbot does not cause injury to a user.

3) A chatbot provider is liable for any injury it caused a user through the use of its chatbot, even if:

a) The chatbot provider exercised all reasonable care in the design and distribution of the chatbot; or

b) The chatbot provider did not directly distribute the chatbot to the user or otherwise enter into a contractual relationship with the user.

Section 8: Enforcement

1) The Attorney General, a district attorney, or a municipality may bring a civil action against a chatbot provider that violates this Act to:
    a) enjoin the act or practice that is in violation of this Act;
    b) enforce compliance with this Act or a rule adopted under this Act;
    c) obtain damages, civil penalties, restitution, or other remedies on behalf of the residents of [insert appropriate jurisdiction]; or
    d) obtain reasonable attorney's fees and other litigation costs reasonably incurred.
2) A violation of Sections 2 or 3 of this Act constitutes an injury in fact to a user.
3) A user injured by a violation of Sections 2 or 3 of this Act may bring a civil action against the chatbot provider, in which the court may award a prevailing plaintiff:
    a) statutory damages, as updated annually by the [Attorney General or other appropriate entity] pursuant to Section 5(1)(e), of:
        i) $5,000 per violation, for any violation of Section 2, or actual damages, whichever is greater; and
        ii) $5,000 in total for all violations of Section 3, or actual damages, whichever is greater;
    b) punitive damages, for reckless and knowing violations;
    c) injunctive relief;
    d) declaratory relief; and
    e) reasonable attorney's fees and litigation costs.