
A BILL FOR AN ACT

RELATING TO ARTIFICIAL INTELLIGENCE.

BE IT ENACTED BY THE LEGISLATURE OF THE STATE OF HAWAII:

1 SECTION 1. This Act shall be known and may be cited as the
2 "Artificial Intelligence Disclosure and Safety Act".

3 SECTION 2. Chapter 481B, Hawaii Revised Statutes, is
4 amended by adding a new section to part I to be appropriately
5 designated and to read as follows:

6 "§481B- Artificial intelligence; conversational
7 artificial intelligence services; disclosures; reports;

8 penalties. (a) If a reasonable person interacting with a
9 conversational artificial intelligence service would be led to
10 believe that the person is interacting with a human, an operator
11 shall issue a clear and conspicuous notification indicating that
12 the service is artificial intelligence and not human.

13 (b) If an operator has actual knowledge or reasonable
14 certainty that an account holder or a user is a minor, the
15 operator shall clearly and conspicuously disclose to the minor
16 account holder or user that the minor is interacting with
17 artificial intelligence:



- 1 (1) As a persistent visible disclaimer; or
- 2 (2) Both:
- 3 (A) At the beginning of each session; and
- 4 (B) Appearing at least every three hours in a
- 5 continuous conversational artificial intelligence
- 6 service interaction that reminds the user to take
- 7 a break from the chat and that the conversation
- 8 is artificially generated and not with a human.
- 9 (c) An operator shall:
- 10 (1) Adopt a protocol for the conversational artificial
- 11 intelligence service to respond to user prompts
- 12 regarding suicidal ideation or self-harm that includes
- 13 but is not limited to making reasonable efforts to
- 14 provide a response that refers the user to crisis
- 15 intervention service providers, such as a suicide
- 16 hotline, crisis text line, or other appropriate crisis
- 17 services;
- 18 (2) Use evidence-based methods for measuring suicidal
- 19 ideation;
- 20 (3) Not cause or program a conversational artificial
- 21 intelligence service to make any representation or



1 statement that explicitly indicates that the service
2 is designed to provide professional mental or
3 behavioral health care; and

4 (4) Institute reasonable measures to prevent the
5 conversational artificial intelligence service from
6 making any representation or statement that would lead
7 a reasonable person to believe that the person is
8 interacting with a human where the person is seeking
9 or receiving crisis intervention services for self-
10 harm or suicide.

11 (d) Where an operator knows or has reasonable certainty
12 that an account holder or user is a minor, an operator shall:

13 (1) Not provide the account holder or user with points or
14 similar rewards at unpredictable intervals with the
15 intent to encourage increased engagement with the
16 conversational artificial intelligence service;

17 (2) Institute reasonable measures to prevent the
18 conversational artificial intelligence service from:

19 (A) Producing visual material of sexually explicit
20 conduct;

21 (B) Generating or deploying:



- 1 (i) Direct statements that the account holder or
2 user should engage in sexually explicit
3 conduct;
- 4 (ii) Representations that would reasonably lead
5 the minor to believe that the conversational
6 artificial intelligence service is a human
7 or sentient being, unless the
8 representations are simultaneously
9 accompanied by a clear and conspicuous
10 disclosure that the account holder or user
11 is not engaging with a human or sentient
12 being;
- 13 (iii) Representations that would reasonably lead
14 the minor to think that the conversational
15 artificial intelligence service is capable
16 of engaging in a romantic relationship or
17 sexual relationship with the minor;
- 18 (iv) Manipulative engagement techniques designed
19 to foster prolonged interaction or emotional
20 dependency by minors, including techniques
21 that create artificial rewards, simulated



1 exclusivity, or pressure to continue
2 interacting with the service; or
3 (v) Representations that would reasonably lead
4 the minor to believe that the conversational
5 artificial intelligence service is providing
6 professional mental health, medical, or
7 therapeutic services; or
8 (C) Generating statements that sexually objectify the
9 account holder or user; and
10 (3) Offer tools to manage the minor account holder's or
11 user's privacy and account settings; provided that the
12 operator shall also make the tools required by this
13 paragraph available to the parent or guardian of a
14 minor account holder or user, as appropriate based on
15 relevant risks;
16 provided that this subsection shall not prohibit conversational
17 artificial intelligence services that express generic
18 expressions of politeness or encouragement; neutral assistance
19 or information provided without fostering emotional attachment;
20 or safety-oriented responses intended to de-escalate harm or
21 redirect an account holder or a user to appropriate resources.



1 (e) Beginning January 1, 2028, an operator shall submit to
2 the department of commerce and consumer affairs an annual report
3 that includes:

4 (1) The number of times an operator has issued a crisis
5 intervention services provider referral in the
6 preceding calendar year;

7 (2) Protocols put in place to detect, remove, and respond
8 to instances of suicidal ideation by users; and

9 (3) Protocols put in place to prohibit a service response
10 about suicidal ideation or actions with the user;

11 provided that the report shall include only the information
12 listed in this subsection and shall not include any identifiers
13 or personal information about users.

14 (f) Any violation of this section shall constitute an
15 unfair or deceptive act or practice in the conduct of trade or
16 commerce within the meaning of section 480-2.

17 (g) The attorney general and director of the office of
18 consumer protection may bring an action based on a violation of
19 this section to recover all of the following relief:

20 (1) Injunctive relief;

21 (2) Damages in an amount equal to the greater of:



1 (A) Actual damages; or

2 (B) A civil penalty of \$1,000 per violation; provided
3 that damages awarded under this subparagraph
4 shall not exceed \$1,000,000 per operator; and

5 (3) Reasonable attorneys' fees and costs.

6 (h) Notwithstanding section 480-2 or any other law to the
7 contrary and notwithstanding subsection (j), no person,
8 including a consumer, may bring a private action to enforce this
9 section or support a private action under any other law.

10 (i) This section shall not create liability for the
11 developer of an artificial intelligence model for any violation
12 of this section by an artificial intelligence system developed
13 by a third party to provide a conversational artificial
14 intelligence service.

15 (j) The duties, remedies, and obligations imposed by this
16 section are cumulative to the duties, remedies, or obligations
17 imposed under other law and shall not be construed to relieve an
18 operator from any duties, remedies, or obligations imposed under
19 any other law.

20 (k) For the purposes of this section:



1 "Account holder" or "user" means a person who has or
2 generates an account or profile to use a conversational
3 artificial intelligence service.

4 "Artificial intelligence" means an engineered or
5 machine-based system that varies in its level of autonomy and
6 that can, for explicit or implicit objectives, infer from the
7 input it receives how to generate outputs that can influence
8 physical or virtual environments.

9 "Conversational artificial intelligence service" or
10 "service" means an artificial intelligence software application,
11 web interface, or computer program that is accessible to the
12 general public that primarily simulates human conversation and
13 interaction through textual, visual, or aural communications.

14 "Crisis intervention" means communication intended to
15 provide immediate support or assistance in response to a user
16 seeking help for, referencing, or expressing self-harm, suicidal
17 ideation, or suicide.

18 "Emotional attachment" means a sustained bond that develops
19 through repeated interaction.



1 "Emotional dependency" means a pattern of interaction in
2 which a minor develops an excessive emotional reliance on a
3 conversational artificial intelligence service.

4 "Minor" means any person under eighteen years of age.

5 "Operator" means a person who develops or makes available a
6 conversational artificial intelligence service to the public. A
7 mobile application store or search engine that merely provides
8 access to a conversational artificial intelligence service shall
9 not by itself constitute an "operator".

10 "Personal information" has the same meaning as defined in
11 section 487N-1.

12 "Romantic relationship" means interactions in which a
13 conversational artificial intelligence service represents or
14 simulates a relationship based on emotional and physical
15 attraction that could potentially lead to a long-term intimate
16 relationship.

17 "Sexual relationship" means interactions in which a
18 conversational artificial intelligence service represents or
19 simulates physical attraction, desire, and sexual contact that
20 may or may not include an emotional connection.



1 "Sexually explicit conduct" has the same meaning as defined
2 in title 18 United States Code section 2256.

3 "Sexually objectify" means to make sexual comments directed
4 at the account holder's or user's body or appearance."

5 SECTION 3. If any provision of this Act, or the
6 application thereof to any person or circumstance, is held
7 invalid, the invalidity does not affect other provisions or
8 applications of the Act that can be given effect without the
9 invalid provision or application, and to this end the provisions
10 of this Act are severable.

11 SECTION 4. This Act does not affect rights and duties that
12 matured, penalties that were incurred, and proceedings that were
13 begun before its effective date.

14 SECTION 5. New statutory material is underscored.

15 SECTION 6. This Act shall take effect on July 1, 3000;
16 provided that section 2 shall take effect on July 1, 2027.



Report Title:

DCCA; AG; OCP; Artificial Intelligence; Conversational Artificial Intelligence Services; Disclosures; Protections; Minors; UDAP; Reports

Description:

Requires operators of conversational artificial intelligence services in the State to issue certain disclosures to account holders and users. Requires operators to develop protocols to prevent the production of suicidal ideations in account holders and users. Establishes protections for account holders and users of conversational artificial intelligence services. Establishes protections for minor account holders and users of conversational artificial intelligence services. Beginning 1/1/2028, requires operators to submit annual reports to the Department of Commerce and Consumer Affairs containing certain information. Allows the Department of the Attorney General and Office of Consumer Protection to bring a civil action against operators who violate certain requirements. Establishes that violations are to be considered unfair or deceptive acts or practices. Effective 7/1/3000. (HD3)

The summary description of legislation appearing on this page is for informational purposes only and is not legislation or evidence of legislative intent.

