HB0286S01 compared with HB0286

1 **Artificial Intelligence Transparency Amendments**

2026 GENERAL SESSION

STATE OF UTAH

Chief Sponsor: Doug Fiefia

Senate Sponsor:

2

3 **LONG TITLE**

4 **General Description:**

5   This bill enacts the AI Transparency Act relating to transparency and whistleblower

6 protections for frontier artificial intelligence models.

7 **Highlighted Provisions:**

8   This bill:

9   ‣ defines terms;

10   ‣ requires developers of certain artificial intelligence models to create, implement, and publish public safety and child protection plans;

12   ‣ requires developers to publish summaries of risk assessments for certain artificial intelligence models;

14   ‣ prohibits developers from making materially false or misleading statements about covered risks;

16   ‣ requires developers to report certain safety incidents to the Office of Artificial Intelligence Policy (office);

18   ‣ requires the office to provide annual assessments and legislative recommendations regarding regulation of certain artificial intelligence models;

20 ‣ establishes civil penalties for violations;

21 ‣ provides whistleblower protections for employees who report safety concerns of certain artificial
intelligence models;

23 ‣ establishes remedies for employees who suffer adverse action for whistleblower activities; {and}

25 ‣ **creates the AI Transparency Enforcement Restricted Account to fund enforcement
activities; and**

25 ‣ provides a severability clause.

28 **Money Appropriated in this Bill:**

29 None

30 **Other Special Clauses:**

31 None

32 **Utah Code Sections Affected:**

33 ENACTS:

34 **13-72b-101** , Utah Code Annotated 1953

35 **13-72b-102** , Utah Code Annotated 1953

36 **13-72b-103** , Utah Code Annotated 1953

37 **13-72b-104** , Utah Code Annotated 1953

38 **13-72b-105** , Utah Code Annotated 1953

39 **13-72b-106** , Utah Code Annotated 1953

40 **13-72b-107** , Utah Code Annotated 1953

41 **13-72b-108** , Utah Code Annotated 1953

42 **13-72b-109 , Utah Code Annotated 1953**

43 **13-72b-201** , Utah Code Annotated 1953

44 **13-72b-202** , Utah Code Annotated 1953

45 **13-72b-203** , Utah Code Annotated 1953

46 **13-72b-204** , Utah Code Annotated 1953

47

48 *Be it enacted by the Legislature of the state of Utah:*

49 Section 1. Section **1** is enacted to read:

52 **13-72b-101. Definitions.**

72b. AI Transparency Act

1. Artificial Intelligence Transparency and Child Protection

As used in this chapter:

51   (1) "Affiliate" means a person controlling, controlled by, or under common control with a specified person, directly or indirectly, through one or more intermediaries.

53   (2) "Artificial intelligence model" means an engineered or machine-based system that varies in the system's level of autonomy and that can, for explicit or implicit objectives, infer from the input the artificial intelligence model receives how to generate outputs that can influence physical or virtual environments.

57   (3)

   (a) "Catastrophic loss" means:

58   (i) the death or serious bodily injury of more than 50 individuals; or

59   (ii) damage to property, or loss of property, exceeding $1,000,000,000.

60   (b) "Catastrophic loss" does not include the loss of value of equity.

61   (4)

   (a) "Catastrophic risk" means a foreseeable and material risk that a frontier developer's development, storage, use, or deployment of a frontier model will materially contribute to a catastrophic loss in a single incident by:

64   (i) providing assistance in creating or releasing a chemical, biological, radiological, or nuclear weapon;

66   (ii) engaging in a cyberattack, or conduct that, if committed by an individual, would constitute murder, assault, extortion, or theft, including theft by deception, under Utah law, without meaningful human oversight, intervention, or supervision; or

69   (iii) evading control of the frontier developer or user.

70   (b) "Catastrophic risk" does not include a foreseeable and material risk from any of the following:

72   (i) information that a frontier model outputs if the information is otherwise publicly accessible in a substantially similar form from a source other than a foundation model;

75   (ii) lawful activity of the federal government; or

76   (iii) harm caused by a frontier model in combination with other software if the frontier model did not materially contribute to the harm.

78   (5) "Child protection plan" means a documented technical and organizational protocol to manage, assess, and mitigate child safety risks.

80   (6) "Child safety incident" means an occurrence in which a covered chatbot, when interacting with a minor, engages in behavior that, if engaged in by a human, would be considered to intentionally or recklessly:

83   (a) cause death or bodily injury to the minor; or

84   (b) cause severe emotional distress to the minor.

85   (7) "Child safety risk" means a material and foreseeable risk that a frontier developer's foundation model, when used as part of a covered chatbot operated by the frontier developer, will engage in behavior when interacting with a minor that, if the behavior had been engaged in by a human, would be considered to be intentionally or recklessly:

89   (a) cause death or bodily injury to the minor, including as a result of self-harm; or

90   (b) cause severe emotional distress to the minor.

91   (8) "Covered chatbot" means a service that:

92   (a) allows an ordinary person to have conversations in which humanlike responses are generated by a foundation model;

94   (b) is foreseeably likely to be accessed by minors; and

95   (c) has at least 1,000,000 monthly active users.

96   (9) "Covered risk" means a catastrophic risk or a child safety risk.

97   (10) "Critical safety incident" means any of the following:

98   (a) unauthorized access to, modification of, inadvertent release of, or exfiltration of, the model weights of a frontier model;

100   (b) the death of, or serious injury to, more than 50 people or more than $1,000,000,000 in damage to, or loss of, property resulting from the materialization of a catastrophic risk;

103   (c) loss of control of a frontier model that:

104   (i) causes death or bodily injury; or

105   (ii) demonstrates materially increased catastrophic risk; or

106   (d) a frontier model that uses deceptive techniques against the frontier developer to subvert the controls or monitoring of the frontier developer outside of the context of an evaluation designed to elicit this behavior and in a manner that demonstrates materially increased catastrophic risk.

110   (11)

      (a) "Deploy" means to make a frontier model available to a third party for use, modification, copying, or combination with other software.

112   (b) "Deploy" does not include making a frontier model available to a third party for the primary purpose of developing or evaluating the frontier model.

114   (12) "Foundation model" means an artificial intelligence model that is all of the following:

115   (a) trained on a broad data set;

116   (b) designed for generality of output; and

117   (c) adaptable to a wide range of distinctive tasks.

118   (13)

    (a) "Frontier developer" means a person who has used, or initiated the use of, a quantity of computing power of at least 10^{26} integer or floating-point operations to train a frontier model, including computing used for the original training run and for any subsequent fine-tuning, reinforcement learning, or other material modifications.

122   (b) "Frontier developer" does not include an accredited college or university to the extent the college or university is developing or using frontier models exclusively for academic research purposes.

125   (14) "Frontier model" means a foundation model that was trained using a quantity of computing power of at least 10^{26} integer or floating-point operations, including computing for the original training run and for any subsequent fine-tuning, reinforcement learning, or other material modifications the developer applies to a {preceding } foundation model.

130   (15) "Large frontier developer" means a frontier developer who together with the frontier developer's affiliates, had annual revenue of at least $500,000,000 in the {preceding } calendar year.

133   (16) "Minor" means an individual younger than 18 years old.

134   (17) "Model weight" means a numerical parameter in a frontier model that is adjusted through training and that helps determine how inputs are transformed into outputs.

136   (18) "Office" means the Office of Artificial Intelligence Policy created in Section 13-72-201.

137   (19) "Property" means tangible or intangible property.

138   (20) "Public safety plan" means a documented technical and organizational protocol to manage, assess, and mitigate catastrophic risks.

140   (21) "Safety incident" means a child safety incident or a critical safety incident.

144   Section 2. Section **2** is enacted to read:

145   **13-72b-102. Public safety plan for catastrophic risks -- Requirements.**

143

(1) A large frontier developer shall write, implement, comply with, and clearly and conspicuously publish on the large frontier developer'sinternet website a public safety plan that describes in detail how the large frontier developer:

146 (a) incorporates national standards, international standards, and industry-consensus best practices into the public safety plan;

148 (b) defines and assesses thresholds used by the large frontier developer to identify and assess whether a frontier model has capabilities that could pose a catastrophic risk, which may include multiple-tiered thresholds;

151 (c) applies mitigations to address the potential for catastrophic risks based on the results of assessments undertaken {pursuant to } Subsection (1)(b);

153 (d) reviews assessments of catastrophic risk and adequacy of mitigations of catastrophic risk as part of the decision to deploy a frontier model or use the frontier model extensively internally;

156 (e) uses third parties to assess the potential for catastrophic risks and the effectiveness of mitigations of catastrophic risks;

158 (f) revisits and updates the public safety plan, including any criteria that trigger updates and how the large frontier developer determines when the large frontier developer's frontier models are substantially modified enough to require disclosures {pursuant to } Section 13-72b-104;

162 (g) implements cybersecurity practices to secure unreleased frontier model weights from unauthorized modification or transfer by internal or external parties;

164 (h) identifies and responds to critical safety incidents;

165 (i) institutes internal governance practices to ensure implementation of the processes described in this Subsection (1); and

167 (j) assesses and manages catastrophic risk resulting from the internal use of the large frontier developer's frontier models, including risks resulting from a frontier model circumventing oversight mechanisms.

170 (2) If a large frontier developer makes a material modification to the large frontier developer's public safety plan, the large frontier developer shall clearly and conspicuously publish the modified public safety plan and a justification for that modification within 30 days after the day on which the large frontier developer makes the material modification.

178   Section 3. Section **3** is enacted to read:

179   **13-72b-103. Child protection plan -- Requirements.**

177   (1) A large frontier developer that operates a covered chatbot shall write, implement, comply with, and clearly and conspicuously publish on the large frontier developer's internet website a child protection plan that describes in detail how the large frontier developer:

181   (a) incorporates national standards, international standards, and industry-consensus best practices into the child protection plan;

183   (b) assesses potential for child safety risks;

184   (c) applies mitigations to address the potential for child safety risks based on the results of assessments undertaken {~~pursuant to~~} Subsection (1)(b);

186   (d) uses third parties to assess the potential for child safety risks and the effectiveness of mitigations of child safety risks;

188   (e) revisits and updates the child protection plan, including any criteria that trigger updates and how the large frontier developer determines when the large frontier developer's foundation models are substantially modified enough to require disclosures {~~pursuant to~~} Section 13-72b-104;

192   (f) identifies and responds to child safety incidents; and

193   (g) institutes internal governance practices to ensure implementation of the processes described in this Subsection (1).

195   (2) If a large frontier developer makes a material modification to the large frontier developer's child protection plan, the large frontier developer shall clearly and conspicuously publish the modified child protection plan and a justification for that modification within 30 days after the day on which the large frontier developer makes the material modification.

203   Section 4. Section **4** is enacted to read:

204   **13-72b-104. Publication requirements -- Frontier models and foundation models.**

202   (1) A large frontier developer shall conspicuously publish on the developer's i website summaries of the following before deploying a new or substantially modified foundation model as part of a covered chatbot operated by the developer:

205   (a) assessments of child safety risks conducted {~~pursuant to~~} the developer's child protection plan;

207   (b) the results of the assessments described in Subsection (1)(a);

208   (c) the extent to which third-party evaluators were involved in the assessments described in Subsection (1)(a); and

210   (d) other steps taken by the developer to fulfill the requirements of the child protection plan.

212

(2) A large frontier developer shall conspicuously publish on the large frontier developer's website summaries of the following before deploying a new frontier model or a frontier model that was substantially modified by the large frontier developer:

215 (a) assessments of catastrophic risks from the frontier model conducted {~~pursuant to~~} the developer's public safety plan;

217 (b) the results of the assessments described in Subsection (2)(a);

218 (c) the extent to which third-party evaluators were involved in the assessments described in Subsection (2)(a); and

220 (d) other steps taken to fulfill the requirements of the public safety plan with respect to the frontier model.

225   Section 5. Section **5** is enacted to read:

226   **13-72b-105. Prohibited conduct -- Redactions.**

224 (1)

(a) A frontier developer may not make a materially false or misleading statement or omission about covered risks from the developer's activities or the developer's management of covered risks.

227 (b) A large frontier developer may not make a materially false or misleading statement or omission about the large frontier developer's implementation of, or compliance with, the large frontier developer's public safety plan.

230 (c) A large frontier developer that operates a covered chatbot may not make a materially false or misleading statement or omission about the large frontier developer's implementation of, or compliance with, the large frontier developer's child protection plan.

234 (d) This Subsection (1) does not apply to a statement that was made in good faith and was reasonable under the circumstances.

236 (2)

(a) When a frontier developer publishes documents to comply with this part, the frontier developer may make redactions to those documents that are necessary to protect:

239 (i) the frontier developer's trade secrets;

240 (ii) the frontier developer's cybersecurity;

241 (iii) public safety;

242 (iv) the national security of the United States; or

243 (v) compliance with any federal or state law.

244   (b) If a frontier developer redacts information in a document {~~pursuant to~~} Subsection (2)(a), the large
frontier developer shall:

246   (i) describe the character and justification of the redaction in any published version of the document to
the extent permitted by the concerns that justify the redaction; and

248   (ii) retain the unredacted information for five years after the day on which the developer makes the
redaction.

253   Section 6. Section **6** is enacted to read:

254   **13-72b-106. Safety incident reporting mechanism -- Rulemaking -- Annual report.**

253   (1) The office may make rules in accordance with Title 63G, Chapter 3, Utah Administrative
Rulemaking Act, to:

255   (a) establish a mechanism for a large frontier developer or a member of the public to report a safety
incident; and

257   (b) establish alternate compliance procedures if substantially equivalent or stricter federal reporting
requirements or guidance documents are established.

259   (2) A large frontier developer shall report a safety incident to the office within 15 days after the day on
which the large frontier developer discovers the incident.

261   (3) A large frontier developer that discovers a critical safety incident that poses an imminent risk of
death or serious physical injury, shall disclose that incident within 24 hours to a law enforcement
agency or public safety agency with appropriate jurisdiction based on the nature of the incident.

265   (4) A large frontier developer shall submit to the office a report summarizing assessments of
catastrophic risk resulting from internal use of the large frontier developer's frontier models:

268   (a) at least once every three months; or

269   (b) {~~pursuant to~~} an alternate schedule if:

270   (i) the large frontier developer requests the alternate schedule from the office in writing; and

272   (ii) the office agrees to the alternate schedule.

273   (5) The office may transmit the reports described in Subsections (2) and (4) to the Legislature, the {~~r,~~
} governor, the federal government, or appropriate state agencies, but may consider risks related to
trade secrets, public safety, cybersecurity, or national security when transmitting reports.

277   (6) A report submitted under Subsection (2) or (4) may be classified as a protected record under
Subsections 63G-2-305(1) and (2) if the requirements of Subsection 63G-2-309(1)(a)(i) are met.

280

(7) On or before November 1, 2027, and annually thereafter, the office shall prepare a report for the Business and Labor Interim Committee that includes recommendations for modifying this chapter as well as anonymized, aggregated information about reports received {~~pursuant to~~} this chapter, without including information that would compromise the trade secrets or cybersecurity of a frontier developer, public safety, or the national security of the United States or that would be prohibited by any federal or state law.

289   Section 7. Section **7** is enacted to read:

290   **13-72b-107. Civil penalty.**

288 (1) A large frontier developer that violates this part is subject to a civil penalty that does not exceed:

290 (a) for a first violation, $1,000,000; or

291 (b) for each subsequent violation, $3,000,000.

292 (2) A civil penalty under this section may be recovered in a civil action brought by the attorney general on behalf of the office.

297   Section 8. Section **8** is enacted to read:

298   **13-72b-108. {~~Transfer of frontier developer obligations -- Severability~~} AI Transparency Enforcement Restricted Account -- Creation -- Deposits into account -- Distribution.**

296 {~~(1)~~ {~~If any provision of this chapter or the application of any provision to any person or circumstance is held invalid by a final decision of a court of competent jurisdiction, the remainder of this chapter shall be given effect without the invalid provision or application.~~} }

300 (1) There is created within the General Fund a restricted account known as the "AI Transparency Enforcement Restricted Account."

302 (2) The account consists of:

303 (a) money collected by the attorney general from civil penalties, settlements, judgments, and other relief obtained in civil actions brought under Section 13-72b-107;

305 (b) appropriations made to the account by the Legislature; and

306 (c) interest and earnings on account money.

307 (3) The Division of Finance shall deposit money described in Subsection (2)(a) into the account.

309 (4) Upon appropriation by the Legislature, money in the account shall be distributed to the Office of the Attorney General for:

311 (a) investigations and enforcement of Part 1, Artificial Intelligence Transparency and Child Protection;

313 (b) attorney fees and litigation costs related to enforcement actions under this chapter;

314 (c) expert witnesses, consultants, and technical advisors with expertise in artificial intelligence safety and frontier models;

316 (d) specialized equipment, technology, and facilities necessary for enforcement activities;

317 (e) coordination with the Office of Artificial Intelligence Policy, federal agencies, and other state agencies; and

300 (2){(f)} {The provisions } other expenses related to the administration and enforcement of this chapter {are severable} .

301 {(3) {The duties and obligations imposed by this chapter are cumulative with any other duties or obligations imposed under other law and shall not be construed to relieve any party from any duties or obligations imposed under other law and do not limit any rights or remedies under existing law.}

 }

320 Section 9. Section **9** is enacted to read:

321 **13-72b-109. Transfer of frontier developer obligations -- Severability.**

322 (1) If any provision of this chapter or the application of any provision to any person or circumstance is held invalid by a final decision of a court of competent jurisdiction, the remainder of this chapter shall be given effect without the invalid provision or application.

326 (2) The provisions of this chapter are severable.

327 (3) The duties and obligations imposed by this chapter are cumulative with any other duties or obligations imposed under other law and shall not be construed to relieve any party from any duties or obligations imposed under other law and do not limit any rights or remedies under existing law.

331 Section 10. Section **10** is enacted to read:

333 **13-72b-201. Definitions.**

 2. Public Safety and Child Protection Whistleblower Protections

 As used in this part:

309 (1) "Adverse action" means to discharge, threaten, harass, or otherwise discriminate against an employee in any manner that affects the employee's employment, including:

311 (a) compensation;

312 (b) terms;

313 (c) conditions;

314 (d) location;

315 (e) rights;

316   (f) immunities;

317   (g) promotions; or

318   (h) privileges.

319   (2) "Employee" means an individual who performs a service for wages or other remuneration under a

         contract of hire, written or oral, express or implied, for a frontier developer.

322   (3) "Reporter" means an individual who provides information relating to a violation in accordance with

         Section 13-72b-202.

350   Section 11. Section **11** is enacted to read:

351   **13-72b-202. Procedure for disclosure -- Internal whistleblower process.**

326   (1) To be a reporter for purposes of this part, an individual shall:

327   (a) reasonably believe that an act poses a specific and substantial threat to public health or safety or to

         the health or safety of a minor, or is a violation of Part 1, Artificial Intelligencend Child Protection;

         and

330   (b) provide information to the office:

331   (i) in writing; and

332   (ii) in accordance with procedures established by the office by rule made in accordance with Title 63G,

         Chapter 3, Utah Administrative Rulemaking Act.

334   (2)

       (a) Notwithstanding Title 63G, Chapter 2, Government Records Access and Management Act,

         and except as provided in Subsection (2)(b), the office may not disclose information that could

         reasonably be expected to reveal the identity of a reporter.

338   (b) Subsection (2)(a) does not limit the office's ability to present evidence to a grand jury or share

         evidence with witnesses or defendants in an ongoing criminal investigation.

340   (3) A large frontier developer shall provide a reasonable internal process through which an employee

         may anonymously report information if the employee believes in good faith that:

343   (a) the large frontier developer's activities pose a specific and substantial threat to public health or

         safety or to the health or safety of a minor; or

345   (b) the large frontier developer has violated Part 1, Artificial Intelligence and Child Protection.

347   (4) The process required by Subsection (3) shall include monthly updates to the reporting employee

         regarding the status of the investigation and actions taken in response to an anonymous report

         described in Subsection (3).

350   (5)

        (a) Except as provided in Subsection (5)(b), disclosures and responses under this section shall be shared
            with officers and directors of the large frontier developer at least once each quarter.

353   (b) If an employee alleges wrongdoing by an officer or director, Subsection (5)(a) does not apply with
            respect to that officer or director.

381   Section 12. Section **12** is enacted to read:

382   **13-72b-203. Reporter protected from adverse action -- Exceptions.**

357   (1) A frontier developer may not take adverse action against an employee because of a lawful act of the
            employee, or a person authorized to act on behalf of the employee, to:

359   (a) provide information to the office in accordance with Section 13-72b-202, if the employee is a
            reporter;

361   (b) initiate, testify in, or assist in any investigation, judicial action, or administrative action based on or
            related to information provided to the office, if the employee is a reporter; or

364   (c) provide information through an internal reporting process established by the frontier developer.

366   (2) A frontier developer may not make, adopt, enforce, or enter into a rule, regulation, policy, or
            contract that would prevent an employee, or a person authorized on behalf of the employee, from
            taking any of the actions described in Subsection (1).

369   (3) An employee is not protected under this section if the employee:

370   (a) knowingly or recklessly makes a false, fictitious, or fraudulent statement or misrepresentation;

372   (b) uses a false writing or document knowing that, or with reckless disregard as to whether, the writing
            or document contains false, fictitious, or fraudulent information; or

375   (c) knows that, or has a reckless disregard as to whether, the disclosure is of information that is false or
            frivolous.

377   (4) Information provided {~~pursuant to~~} this section may be classified as a protected record under
            Subsections 63G-2-305(1) and (2) if the requirements of Subsection 63G-2-309(1)(a)(i) are met.

406   Section 13. Section **13** is enacted to read:

407   **13-72b-204. Remedies for employee bringing action.**

382   (1)

        (a) An employee who alleges a violation of Section 13-72b-203 may bring an action for injunctive
            relief, actual damages, or both, in a court with jurisdiction under Title 78A, Judiciary and Judicial
            Administration.

385 (b) An employee may not bring an action under this section more than:

386 (i) four years after the day on which the violation of Section 13-72b-203 occurs; or

387 (ii) two years after the day on which facts material to the right of action are known or reasonably should be known by the employee.

389 (2) To prevail in an action under this section, an employee shall establish, by a preponderance of the evidence, that the employee suffered an adverse action because the employee, or a person acting on the employee's behalf, engaged or intended to engage in an activity protected under Section 13-72b-203.

393 (3) A court may award relief for an employee prevailing in an action under this section:

394 (a) reinstatement with the same fringe benefits and seniority status that the individual would have had, but for the adverse action;

396 (b) two times the amount of back pay otherwise owed to the individual, with interest;

397 (c) compensation for litigation costs, expert witness fees, and reasonable attorney fees;

398 (d) actual damages; or

399 (e) any combination of the remedies listed in this Subsection (3).

400 (4)

 (a) An employer may file a counterclaim against an employee who files a civil action under this section seeking attorney fees and costs incurred by the employer related to the action and the counterclaim.

403 (b) The court may award an employer who files a counterclaim under Subsection (4)(a) attorney fees and costs if the court finds that:

405 (i) there is no reasonable basis for the civil action filed by the employee; or

406 (ii) the employee is not protected under Section 13-72b-203 because the employee engaged in an act described in Subsection 13-72b-203(3).

434 Section 14. **Effective date.**

 Effective Date.

 This bill takes effect on May 6, 2026.

1-27-26 10:28 AM