HOUSE BILL 401

57TH LEGISLATURE - STATE OF NEW MEXICO - FIRST SESSION, 2025

INTRODUCED BY

Linda Serrato

AN ACT

RELATING TO ARTIFICIAL INTELLIGENCE; ENACTING THE ARTIFICIAL INTELLIGENCE SYNTHETIC CONTENT ACCOUNTABILITY ACT; PROVIDING FOR CIVIL AND CRIMINAL ENFORCEMENT FOR IMPROPER USE OF SYNTHETIC CONTENT CREATED BY ARTIFICIAL INTELLIGENCE; PROVIDING PENALTIES.

BE IT ENACTED BY THE LEGISLATURE OF THE STATE OF NEW MEXICO:

SECTION 1. [NEW MATERIAL] SHORT TITLE.--This act may be cited as the "Artificial Intelligence Synthetic Content Accountability Act".

SECTION 2. [NEW MATERIAL] DEFINITIONS.--As used in the Artificial Intelligence Synthetic Content Accountability Act:

A. "artificial intelligence" means an engineered or machine-based system that has various levels of autonomy and, for explicit or implicit objectives, can infer from input the

underscored material = new
[~~bracketed material~~] = delete

system receives how to generate outputs that can influence physical or virtual environments;

B. "artificial intelligence red-teaming" means structured testing of a generative artificial intelligence system to identify harmful or discriminatory outputs, unforeseen or undesirable system behaviors, limitations, potential risks associated with the misuse of the system or other flaws or vulnerabilities of the system;

C. "biometric system" means a technology system that links the identity of a person to the person's unique physical characteristics, including the person's fingerprints, iris or face;

D. "content" means images, videos or audio materials;

E. "covered synthetic content" means all synthetic content except for text;

F. "depicted person" means a person depicted in covered synthetic content;

G. "digital fingerprint" means a unique set of information that can be used to identify identical or similar digital content;

H. "digital identification" means information stored on a digital network that serves as proof of the identity of an individual person;

I. "digital signature" means a cryptography-based

underscored material = new
[~~bracketed material~~] = delete

method that uses provenance data to verify that an individual or an entity participated in the creation of certain digital content;

J. "generative artificial intelligence system" means an artificial intelligence system that is capable of generating synthetic content or derivative synthetic content;

K. "large online platform":

(1) means a public-facing or semi-public-facing, internet-based service or application that:

(a) has had at least one hundred thousand users in New Mexico during the preceding twelve months;

(b) substantially functions to connect platform users to allow social interactions among users within the platform; and

(c) can facilitate the sharing of content; and

(2) does not mean a system that provides email or direct messaging communication services alone;

L. "minor modification", with respect to nonsynthetic content, means content that has been changed in a way that does not significantly affect the meaning or perception of the content, including changes to the brightness or contrast of visual content or reduction or removal of background noise in audio content;

underscored material = new
[~~bracketed material~~] = delete

M. "nonsynthetic content" means content created by a person that includes no modifications or only minor modifications;

N. "provenance data" means information about the history of the creation and modification of content, including:

(1) the name of the provider of a generative artificial intelligence system or the camera or recording device manufacturer that relates to the production of content;

(2) the name and version number of the artificial intelligence system that generated the content;

(3) the name and version of the operating system or application used to capture, create or record the content;

(4) the time and date the content was created;

(5) information on any modifications made to the content; and

(6) information on which portions of the content have been changed by a generative artificial intelligence system, if applicable;

O. "provider" means an individual who or an entity that creates, codes, substantially modifies or otherwise produces a generative artificial intelligence system;

P. "reasonable identity verification method" means a method of collecting a person's identifying information using:

underscored material = new
[~~bracketed material~~] = delete

(1) a person's digital identification; or

(2) a commercial identity verification system that verifies identity using a:

(a) government-issued identification document;

(b) biometric system; or

(c) commercially reasonable method that relies on public or private transactional data to verify the identity of an individual;

Q. "state-of-the-art techniques" means techniques with similar performance, reliability and cost compared to the most advanced techniques internally or commercially available within the twelve months preceding the date on which the technique is used;

R. "synthetic content" means content that has been produced or significantly modified from its original form by a generative artificial intelligence system;

S. "transactional data" means a sequence of information that documents an exchange, agreement or transfer between two or more parties used for the purpose of completing a request or event. "Transactional data" includes records from mortgage, educational and employment entities;

T. "watermark" means information that is embedded into content for the purpose of communicating the content's provenance, history of modification or history of conveyance;

underscored material = new
[~~bracketed material~~] = delete

and

U. "watermark decoder" means a software tool or online service that can read or interpret a watermark and provide as output the provenance data associated with the watermark.

SECTION 3. [NEW MATERIAL] IMPROPER DISSEMINATION OF COVERED SYNTHETIC CONTENT--CIVIL LIABILITY.--

A. A private cause of action against a person for the nonconsensual dissemination of covered synthetic content exists when:

(1) the person publicly disseminates covered synthetic content with:

(a) knowledge that a depicted person in the covered synthetic content did not consent to the dissemination; and

(b) the intent to harass, entrap, defame, extort or otherwise cause financial or reputational harm to the depicted person;

(2) the covered synthetic content realistically represents a depicted person engaging in conduct that the depicted person did not actually engage in; and

(3) the depicted person is identifiable from:

(a) the covered synthetic content alone; or

(b) other personal information displayed

underscored material = new
[~~bracketed material~~] = delete

or disseminated in connection with the covered synthetic content.

B. The fact that a depicted person consents to the creation of covered synthetic content or to the nonpublic distribution of the covered synthetic content shall not constitute a defense to liability for a person who improperly disseminates the covered synthetic content as provided in Subsection A of this section.

C. A person shall not be liable for improper dissemination of covered synthetic content as provided in Subsection A of this section if:

(1) the dissemination is made:

(a) for the purpose of a criminal investigation or prosecution that is otherwise lawful;

(b) for the purpose of or in connection with a report of unlawful conduct to appropriate authorities; or

(c) in the course of seeking or receiving medical or mental health treatment, and the covered synthetic content is protected from further dissemination by the recipient;

(2) the person who disseminated the covered synthetic content commercially obtained the content for the purpose of the person's lawful sale of goods or services, including artistic creations, and the depicted person knew that

underscored material = new
[~~bracketed material~~] = delete

the covered synthetic content would be created and disseminated commercially;

(3) the covered synthetic content relates to a matter of public interest; the dissemination of the content serves a lawful public purpose; and the person that disseminates the content clearly identifies that the content is covered synthetic content;

(4) the dissemination is for legitimate scientific research or educational purposes, the covered synthetic content is clearly identified as such and the person who disseminates the content acts in good faith to minimize the risk that the covered synthetic content will be further disseminated;

(5) the dissemination is made for use in legal proceedings and:

(a) is consistent with common practice in civil proceedings necessary for the proper functioning of the court system; or

(b) the content is protected by court order that prohibits any further dissemination; or

(6) the dissemination constitutes criticism, comment, satire, parody, news reporting, teaching, scholarship or research and a reasonable consumer receiving the content would not believe it to accurately represent the depicted person's speech or conduct.

underscored material = new
[~~bracketed material~~] = delete

D. In a civil action filed pursuant to this section, the court may issue an order to protect the privacy of the plaintiff, including protection by:

(1) allowing the plaintiff to use a pseudonym in any documents filed in the action that will be publicly available;

(2) requiring the parties to the action to redact all of the plaintiff's personal identifying information from any documents filed in the action that will be publicly available or to file such documents under seal; or

(3) issuing a protective order for purposes of discovery in the action, which may include an order indicating that any intimate visual depiction or digital forgery shall remain in the care, custody and control of the court.

E. In an action filed pursuant to this section, a prevailing plaintiff may recover reasonable attorney fees and costs and:

(1) liquidated damages in the amount of ten thousand dollars ($10,000); or

(2) actual damages sustained by the plaintiff.

SECTION 4. [NEW MATERIAL] IMPROPER DISSEMINATION OF COVERED SYNTHETIC CONTENT--CRIMINAL LIABILITY.--

A. Improper dissemination of covered synthetic content consists of knowingly disseminating or presenting any likeness of an identifiable person in covered synthetic content

underscored material = new
[~~bracketed material~~] = delete

with the purpose of harassing, entrapping, defaming, extorting or otherwise causing financial or reputational harm to the depicted person.

B. A person who commits improper dissemination of covered synthetic content is guilty of a fourth degree felony and shall be sentenced pursuant to the provisions of Section 31-18-15 NMSA 1978.

C. The attorney general and the district attorney in the county with jurisdiction shall have concurrent jurisdiction to enforce the provisions of this section.

SECTION 5. [NEW MATERIAL] IDENTIFICATION OF, LABELING AND CLASSIFYING SYNTHETIC CONTENT.--

A. A provider shall place an imperceptible watermark that is designed to be as difficult to remove as is reasonably possible using state-of-the-art techniques into covered synthetic content that is produced or significantly modified by a generative artificial intelligence system that the provider makes available. A watermark shall:

(1) identify content as synthetic and identify the provider to ensure that if a sample of the content is corrupted, downscaled, cropped or otherwise damaged, the watermark information will remain; and

(2) be compatible with widely used industry standards.

B. If covered synthetic content is too small to

underscored material = new
[~~bracketed material~~] = delete

directly contain the required provenance data that is part of a watermark, the provider shall, at a minimum, attempt to embed provenance data into the content that identifies the content as partially or entirely synthetic and communicates the following provenance data in order of priority:

(1) the name of the provider;

(2) the name and version number of the artificial intelligence system that generated the content;

(3) the time and date the content was created; and

(4) if applicable, the specific portions of the content that are synthetic.

C. A provider shall:

(1) make available, at no cost to the public, a watermark decoder that:

(a) provides an easy and quick method for a user of the decoder to assess the provenance of a single piece of content; and

(b) to the greatest extent possible, adheres to relevant national or international standards;

(2) before the release of any new generative artificial intelligence system, and annually thereafter, conduct artificial intelligence red-teaming involving third-party experts to test whether watermarks in the system can be easily removed from covered synthetic content produced by a

underscored material = new
[~~bracketed material~~] = delete

provider's generative artificial intelligence system, as well as whether the provider's generative artificial intelligence systems can be used to falsely add watermarks to otherwise nonsynthetic content;

(3) if the provider allows its generative artificial intelligence system to be downloaded and modified, conduct additional artificial intelligence red-teaming to assess whether the provider's system's watermark functionalities can be disabled without authorization;

(4) make summaries of its artificial intelligence red-teaming exercises publicly available in electronic form and provide a clearly labeled link to the summaries on the provider's internet website home page. The link shall be similar in appearance and size relative to other links on the same web page. The provider shall remove from the summaries any details that pose an immediate risk to public safety or provide information that could be used to disable or circumvent the functionality of watermarks specified in the Artificial Intelligence Synthetic Content Accountability Act;

(5) submit a full report of each artificial intelligence red-teaming exercise it conducts to the attorney general within six months of completion of the exercise;

(6) within ninety-six hours of discovering a material vulnerability or failure in a generative artificial intelligence system related to the erroneous or malicious

underscored material = new
[~~bracketed material~~] = delete

inclusion or removal of provenance data or watermarks, report the vulnerability or failure to the attorney general; and:

(a) notify other providers that may be affected by similar vulnerabilities or failures in a manner that allows the other providers to protect their own artificial intelligence systems but does not compromise the reporting provider's systems or disclose the reporting provider's confidential or proprietary information; and

(b) use commercially reasonable efforts to notify parties affected by the vulnerability or failure identified, including notification to online platforms, researchers or users who received incorrect results from a watermark decoder or users who produced covered synthetic content that contained incorrect or insufficient provenance data; provided, however, that a provider shall not be required to notify an affected party whose contact information the provider has not previously collected or retained; and

(7) make any report to the attorney general required pursuant to this section publicly available by providing a clearly labeled link to the report on the provider's internet website home page. The link shall be similar in appearance and size relative to other links on the same web page; provided, however, that if public disclosure of the report could or does present public safety risks, a provider may instead:

underscored material = new
[~~bracketed material~~] = delete

(a) post a summary disclosure of the reported material vulnerability or failure; or

(b) for no longer than thirty days, delay the public disclosure of the report until the public safety risks have been mitigated. If a provider delays public disclosure, the provider shall also document all efforts to resolve the material vulnerability or failure as quickly as possible.

D. A provider and any distributor of software or online services shall not make available to any person a system, application, tool or service that is designed to remove watermarks from covered synthetic content.

E. A large online platform shall use the provenance data of content and state-of-the-art techniques to classify content that is uploaded by users. If the large online platform is able to detect and interpret the provenance data of content, using that provenance data, the large online platform shall classify the content as "fully synthetic", "partially synthetic", "nonsynthetic" or "nonsynthetic with minor modifications". If content uploaded to or distributed on a large online platform by a user does not contain provenance data, or if the provenance data cannot be interpreted or detected by the platform, the platform shall classify the content as "content of unknown provenance".

F. For content classified as content of unknown

underscored material = new
[~~bracketed material~~] = delete

provenance according to Subsection E of this section, a large online platform shall further use state-of-the-art techniques to classify content as "possibly covered synthetic content with unknown provenance" or "possibly nonsynthetic content of unknown provenance".

G. A large online platform shall use labels to disclose the classification assigned to content in accordance with this section. The labels shall prominently display whether content was classified using provenance data or state-of-the-art techniques and which classification the content was assigned, as provided in Subsections E and F of this section. If content was classified according to provenance data, the platform shall ensure that a user is able to click or tap on a label to inspect provenance data, which shall be presented in a clear and simple format. If content was classified according to state-of-the-art techniques, the platform shall include an additional label that warns users that the technique may have incorrectly classified the content and discloses the approximate error rate of the technique used to classify the content.

H. Disclosures required pursuant to this section shall be readily legible to an average viewer or, if the disclosure is made in audio format, shall be clearly audible. A disclosure in audio form shall occur at the beginning and end of a piece of audio content and shall be presented in a

underscored material = new
[~~bracketed material~~] = delete

prominent manner and at a comparable volume and speaking cadence as other spoken words in the content.

I. A user of a large online platform shall have reasonable opportunity to appeal the classification of content by a large online platform.

SECTION 6. [NEW MATERIAL] ENFORCEMENT.--

A. The attorney general may enforce the provisions of the Artificial Intelligence Synthetic Content Accountability Act and may promulgate any rules necessary to implement and enforce the provisions of that act.

B. Prior to filing a civil action to enforce the Artificial Intelligence Synthetic Content Accountability Act, the attorney general may issue a civil investigative demand based on a reasonable belief that a person may be in possession, custody or control of an original or copy of any book, record, report, memorandum, paper, communication, tabulation, map, chart, photograph, mechanical transcription or other document or recording relevant to the subject matter of an investigation of a probable violation of that act. A person issued an investigative demand shall produce the material sought and shall permit it to be copied and inspected. The demand of the attorney general and any material produced in response to it shall not be a matter of public record and shall not be published by the attorney general except by order of the court.

underscored material = new
[~~bracketed material~~] = delete

C. Upon reasonable belief that there has been a violation of the Artificial Intelligence Synthetic Content Accountability Act, the attorney general:

(1) may bring an action in the name of the state to enforce that act;

(2) may petition the district court for injunctive relief;

(3) shall not be required to post bond when seeking a temporary or permanent injunction; and

(4) may recover on behalf of the state a penalty of not less than five thousand dollars ($5,000) and not more than ten thousand dollars ($10,000) for each violation of that act.

SECTION 7. [NEW MATERIAL] POSTING SYNTHETIC CONTENT--IDENTITY VERIFICATION REQUIRED.--

A. A large online platform shall use a reasonable identity verification method to verify a platform user's identity before allowing the user to post content on the platform if the content:

(1) was classified by the platform as fully synthetic, partially synthetic or possibly covered synthetic content of unknown provenance as provided in the Artificial Intelligence Synthetic Content Accountability Act; and

(2) purports to depict reality.

B. The verification process provided for in

underscored material = new
[~~bracketed material~~] = delete

Subsection A of this section shall be performed each time a platform user attempts to post content that meets the descriptions in Paragraph (1) or (2) of that subsection or if more than sixty minutes has elapsed since the previous verification was performed in connection with that user's content; provided, however, that verification shall not be required to be performed more frequently than every sixty minutes.

C. A large online platform shall protect any information obtained while performing actions required pursuant to this section using, at a minimum, the industry standard to protect users' most sensitive information, including medical or financial data.

D. A large online platform shall not use identification information provided by a user or obtained by the platform in the process of actions required pursuant to this section for any purpose other than compliance with this section.

E. A large online platform shall disclose information obtained pursuant to actions performed pursuant to this section only as required by a court order. A court shall only issue an order for disclosure of such information in a civil case if:

(1) the plaintiff in the case undertakes efforts to notify the person that posted the content and whose

underscored material = new
[~~bracketed material~~] = delete

information was obtained through a verification process pursuant to this section that they are the subject of a subpoena or application for an order of disclosure, and the person has had a reasonable opportunity to oppose the subpoena or application;

(2) the plaintiff identifies and sets forth the exact statements or provides information sufficient to identify the content about which the case was filed;

(3) the plaintiff sets forth a prima facie case against the person that posted the content and whose information was obtained through a verification process pursuant to this section by producing evidence for each element of the cause of action; and

(4) the court determines that the rights of the person that posted the content and whose information was obtained through a verification process pursuant to this section under the first amendment to the United States constitution are outweighed by the strength of the prima facie case presented by the plaintiff and the necessity for the disclosure of the person's identity.

F. A large online platform shall disclose information obtained through actions performed pursuant to this section only as required by a court order. A court shall only issue an order for disclosure of such information in a criminal case if the attorney general:

underscored material = new
[~~bracketed material~~] = delete

(1) has a warrant covering the information sought; or

(2) offers specific and articulable facts showing reasonable grounds to believe that the information sought is relevant and material to an ongoing criminal investigation.

G. A large online platform shall clearly and prominently disclose to platform users that information obtained by the platform through the platform's compliance with the Artificial Intelligence Synthetic Content Accountability Act will be released only pursuant to a court order.

SECTION 8. [NEW MATERIAL] SEVERABILITY.--If any part or application of the Artificial Intelligence Synthetic Content Accountability Act is held invalid, the remainder of its application to other situations or persons shall not be affected.

underscored material = new
[~~bracketed material~~] = delete