April 24, 2024

The Honorable Julie Gonzales
Colorado State Capitol
200 East Colfax Avenue
Denver CO 80203

Dear Chair Gonzales:

BSA | The Software Alliance appreciates the opportunity to share insights from the enterprise software sector on artificial intelligence (AI) generally and SB 205. BSA is the leading advocate for the global software industry.[1] BSA members are at the forefront of developing cutting edge services, and their products are used by businesses of all sizes across every sector of the economy. AI is much more than robots, self-driving vehicles, or social media; it is used by companies large and small to create and improve the products and services they provide to consumers, to streamline their internal operations, and to enhance their capacity to make data-informed decisions. BSA members are on the leading edge of providing businesses-to-business tools that help companies leverage the remarkable benefits of AI.[2]

As leaders in the development of enterprise AI, BSA members have unique insights into the technology's tremendous potential to further spur digital transformation in the private and public sectors and the policies that can best support the responsible use of AI, especially high-risk uses of AI. BSA's views are informed by our recent experience with members developing BSA Framework to Build Trust in AI,[3] a risk management framework for mitigating the potential for unintended bias throughout an AI system's lifecycle. Built on a vast body of research and informed by the experience of leading AI developers, the BSA Framework outlines a lifecycle-based approach for performing impact assessments to identify risks of AI bias and highlights corresponding risk mitigation best practices. BSA's extensive experience has helped us identify effective policy solutions for addressing AI risks.

We outline several priorities below that we believe policymakers should focus on when examining AI. We also make a number of specific recommendations to SB 205 to help ensure the legislation

[1] BSA's members include: Adobe, Alteryx, Asana, Atlassian, Autodesk, Bentley Systems, Box, Cisco, CNC/Mastercam, Databricks, DocuSign, Dropbox, Elastic, Graphisoft, Hubspot, IBM, Informatica, Kyndryl, MathWorks, Microsoft, Okta, Oracle, PagerDuty, Palo Alto Networks, Prokon, Rubrik, Salesforce, SAP, ServiceNow, Shopify Inc., Siemens Industry Software Inc., Splunk, Trend Micro, Trimble Solutions Corporation, TriNet, Twilio, Workday, Zendesk, and Zoom Video Communications, Inc.
[2] *See* BSA | The Software Alliance, Artificial Intelligence in Every Sector, *available at* https://www.bsa.org/files/policy-filings/06132022bsaaieverysector.pdf.
[3] *See* BSA | The Software Alliance, Confronting Bias: BSA's Framework to Build Trust in AI, *available at* https://www.bsa.org/reports/confronting-bias-bsas-framework-to-build-trust-in-ai.

200 Massachusetts Avenue, NW
Suite 310
Washington, DC 20001

P 202-872-5500
W bsa.org

is workable in practice and generally encourage continued alignment between SB 205 and Connecticut SB 2.

I. Focus on High-Risk Uses of AI

BSA commends you and the committee for focusing on high-risk uses of AI in SB 205. We recommend policymakers focus on AI systems that determine an individual's eligibility for housing, employment, credit, education, access to physical places of public accommodation, healthcare, or insurance. These systems have the potential to affect important life opportunities — and are a key area for policymakers to address. In contrast, many everyday uses of AI present few risks to individuals and create significant benefits, like helping organize digital files, auto-populate common forms for later human review, improve a company's ability to forecast supply chain issues, and detect, prevent, and respond to cybersecurity threats.

The provisions in Sections 6-1-1602 and 6-1-1603 of SB 205 create a strong foundation for addressing the risks posed by high-risk uses of AI. We have several recommendations for improving these provisions so that they work in practice.

a. **The definition of consequential decision should be revised.** While we appreciate that high-risk uses are tied to consequential decisions, to avoid overbroad application, we recommend focusing the definition of this term on eligibility determinations, changing "access to, or the availability, cost, or terms of" to "eligibility for and results in the provision or denial of" in the definition. Focusing consequential decisions on eligibility determinations and the actual extension or denial of public goods and services helps capture the key aspects of these decisions that have the most impact to consumers' lives.

b. **The list of information the developer provides to a deployer about a high-risk AI system should be revised to reflect the developer's role.** Subsection 2 of Section 6-1-1602 outlines the information a developer must share with a deployer. However, it includes disclosure of an item that would not be within the purview of developers. Specifically, the bill requires developers to explain how an individual can monitor a high-risk AI system when it is used to make, or is a substantial factor in making, a consequential decision. Because developers design, code, or produce AI systems, and deployers use AI systems, they have access to different types of information. In this instance, deployers are best positioned to provide information about how consequential decisions are made and how an individual can monitor the system once deployed.

c. **The bill's requirements for developers and deployers to report when a high-risk AI system has caused algorithmic discrimination should be eliminated.** Subsection 5 of Section 6-1-1602 requires developers to inform all known deployers and the Attorney General when they discover or are informed by a deployer that a deployed high-risk AI system has caused algorithmic discrimination. Additionally, Subsection 6 of Section 6-1-1603 requires deployers to inform the Attorney General when a high-risk AI system has caused algorithmic discrimination. As an initial matter, such requirements envision an ongoing post-deployment relationship with the deployer, which may not be the case. Further, one deployer's use of the high-risk system in a discriminatory manner does not render all other uses discriminatory, and such notice would often be irrelevant to another deployer's use of the system. We suggest aligning with the version of Connecticut SB 2 released on April 23 and striking these requirements.

200 Massachusetts Avenue, NW
Suite 310
Washington, DC 20001

P 202-872-5500
W bsa.org

II. General-Purpose AI Models

The bill's approach to regulating developers of general-purpose AI models raises concerns. As an initial matter, the bill's approach is not risk-based and instead singles out a specific kind of technology to regulate, rather than focusing regulation on particular uses of the technology. Such an approach is overbroad and does not prioritize AI-related uses that pose the most significant risks to consumers. We recommend aligning with the version of Connecticut SB 2 released on April 23 and striking this section.

III. Risk Management Programs

BSA appreciates SB 205's recognition of the importance of risk management programs. Companies should create and maintain risk management programs that help them identify and mitigate risks. Risk management programs establish repeatable processes for companies to identify and mitigate potential risks that can arise throughout the lifecycle of an AI system.

Risk management is particularly important in contexts like AI, privacy, and cybersecurity, where the combination of quickly evolving technologies and highly dynamic threat landscapes can render traditional approaches to compliance ineffective. Risk management programs have two key components: (1) a governance framework of policies, procedures, and personnel that support the company's risk management function, and (2) a scalable process for performing impact assessments that identify and mitigate risks of an AI system.

One way for companies to establish risk management programs is by using the AI Risk Management Framework (AI RMF), which was released earlier this year by the National Institute of Standards and Technology (NIST).[4] The AI RMF builds on NIST's work creating frameworks for managing cybersecurity and privacy risks.[5] The AI RMF helps companies incorporate trustworthiness considerations into the design, development, use, and evaluation of AI products. Ultimately, effective AI risk management programs should support coordination across the company, to promote the identification and mitigation of risks throughout the lifecycle of an AI system.

IV. Impact Assessments

BSA commends the recognition of impact assessments in SB 205 as an important tool for fostering accountability and building trust in AI. BSA recognizes that performing impact assessments is a key part of creating a meaningful risk management program. Impact assessments have three purposes: (1) identifying potential risks that an AI system may pose, (2) quantifying the degree of potential harms the system could generate, and (3) documenting steps taken to mitigate those risks.[6] Impact assessments are already widely used in a range of other fields, including privacy, as an accountability mechanism that demonstrates a product or system has been designed in a manner that accounts for the potential risks it may pose to the public.

Because impact assessments already exist today, they can be readily adapted to help companies

[4] NIST AI Risk Management Framework, available at https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf.
[5] See NIST, Cybersecurity Framework, Questions and Answers, (discussing federal agency use of the NIST CSF), available at https://www.nist.gov/cyberframework/faqs.
[6] See BSA, Impact Assessments: A Key Part of AI Accountability, available at https://www.bsa.org/files/policyfilings/08012023impactassess.pdf.

identify and mitigate AI-related risks.[7] In our view, when AI is used in ways that could adversely impact civil rights or access to important life opportunities, the public should be assured that such systems have been thoroughly vetted and will be continuously monitored to account for the risks associated with unintended bias. Companies, both developers and deployers, should use impact assessments as a tool for the responsible development and use of high-risk AI systems.

V. Distinguishing Different Actors in the AI Ecosystem

BSA appreciates that SB 205 differentiates between different actors in the AI ecosystem, including AI developers and AI deployers. Much like privacy and security laws worldwide distinguish between different types of companies that handle consumers' personal data, AI laws should distinguish between developers and deployers to ensure that legal frameworks accurately assign obligations to a company based on its role in the AI ecosystem.

A developer is the company that designs, codes, or produces an AI system, such as a software company that develops an AI system for speech recognition. A deployer, in contrast, is the company that uses an AI system, such as a bank that uses an AI system either developed internally or by a third party to make loan determinations. Each type of company will have access to different types of information about an AI system and will be positioned to take different actions to mitigate the risks associated with the AI system. AI policies that distinguish between these roles can ensure that the appropriate company within the various real-world AI supply chains can identify and mitigate risks.

Distinguishing between these two types of entities based on of their role in the AI ecosystem can ensure companies are better able to fulfill their obligations and better protect consumers. For example, a developer would be able to describe the features of data used to train an AI system, but it generally would not have insight into how the AI system is used after another company has purchased and implemented the AI system. Instead, the deployer using the system is generally best positioned to understand how the AI system is being used, whether that use aligns with its intended use, whether and how to incorporate human oversight, the outputs from the AI system, any complaints received, and real-world factors affecting the system's performance.

VI. Enforcement

BSA commends SB 205 for granting exclusive enforcement authority to the Attorney General. Exclusive enforcement by the Attorney General helps ensure a consistent approach to enforcement. We also appreciate that the bill does not create a private right of action and expressly states that it does not create a private right of action under any other law.

Additionally, BSA understands that the Office of the Attorney General has significant experience conducting rulemaking processes, including to implement the state's consumer privacy law. However, we recommend policymakers prioritize creating clear statutory requirements in SB 205 that do not require a broad rulemaking process. Establishing strong and clear guardrails within the

[7] For example, three state privacy laws already require companies to conduct impact assessment for specific activities, including processing sensitive personal data, engaging in targeted advertising, or selling personal data; seven more state privacy laws will soon do so. Colorado, Connecticut, and Virginia already impose these requirements. See Colorado Privacy Act, Colo. Rev. Stat. Tit. 6, Art. 1, Pt. 13 §§ 6-1-1301–6-1-1313; Connecticut Data Privacy Act Conn. Gen. Stat. Tit. 42, Ch. 743jj, Sec. 42-515-525; Virginia Consumer Data Protection Act; Va. Code Tit. 59.1, Ch. 53, § 59.1-575-585. State privacy laws in California, Delaware, Florida, Indiana, Kentucky, Montana, Nebraska, New Hampshire, New Jersey, Oregon, Tennessee, and Texas will also require impact assessments for certain activities Globally, privacy and data protection laws worldwide use impact assessments as a tool for improving accountability.

legislation is important for businesses to understand their obligations and for consumers to know what to expect from companies.

* * *

Thank you for allowing us to provide the enterprise software sector's perspective. We welcome the opportunity to serve as a resource and further engage with you or a member of your staff on these important issues.

Sincerely,

Meghan Pensyl
Director, Policy

200 Massachusetts Avenue, NW
Suite 310
Washington, DC 20001

P 202-872-5500
W bsa.org

Testimony of the Center for Democracy & Technology on Senate Bill 24-205

Testimony for the Colorado Senate Judiciary Committee

April 24, 2024

The Center for Democracy & Technology (CDT) thanks the Judiciary Committee for considering our testimony on this legislation. CDT is a nonprofit, nonpartisan organization fighting to advance civil rights and civil liberties in the digital age.

Companies and government agencies are increasingly using artificial intelligence (AI) tools and other automated decision systems (ADSs) to make decisions that dramatically impact the lives of consumers and workers. CDT has been closely watching as policymakers across the country have grappled with how to manage the potential benefits and risks that AI and ADSs pose.

We greatly appreciate the effort that Sen. Rodriguez and his counterparts across the country have clearly put into crafting this legislation, which would cover algorithmic decision-making in settings as diverse as employment, housing, health care, and criminal justice. A great deal of effort and thought clearly went into assembling this ambitious bill. The suggestions below reflect the importance of the issues and how essential it is to get this legislation exactly right--or as close to exactly right as possible. We hope the Committee will take the below suggestions in that constructive spirit.

Take the time to consult carefully with consumer and worker advocates before moving this bill

Given the bill's broad scope, we urge the Committee to listen closely to advocates for consumers and workers with expertise in the various subject-matter areas that the bill covers. It is worth noting that, the testimony of some witnesses notwithstanding, the Connecticut legislation that this bill draws from was, in fact, the result of *industry* input far more than the input of consumer and worker groups. Notably:

- The Connecticut AI working group that was created last year and that helped shape the legislation included 4 representatives from the companies that develop and deploy AI systems and their industry groups, but zero from labor unions, civil rights groups, or other civil society organizations focused on advancing the rights and interests of workers and consumers.
- Key parts of sections 1-3 of the bill parallel HR software giant Workday's model legislation almost verbatim.

We hope, given that background, that the members of this Committee seek and consider the feedback of consumer and worker advocates as it considers this important bill.

Eliminate the requirement that a covered system be a "substantial factor" in a consequential decision.

The current language of the bill would only apply to systems that are "specifically developed and marketed, or specifically modified, to make, or be a substantial factor" in a consequential decision. This is inconsistent with existing laws. Civil rights protections (including in employment, education, and housing) generally apply not only to decisions in which unlawful bias plays a "substantial" role, but to any decision where impermissible factors influenced the outcome. This approach is essential because otherwise, a plaintiff's case becomes virtually impossible to prove. It would be inappropriate to subject AI systems to a lower standard.

Moreover, this requirement would have the practical effect of giving companies the ability to opt out of complying with the bill. In employment, criminal justice, housing, and many other contexts, vendors almost invariably say that their systems are designed merely as tools to *assist* humans, and deployers always say that humans have final say in decisions--even if, in reality, the tools' "recommendations" are decisive and human "reviewers" defer to AI outputs. It would be trivially easy for developers to avoid compliance by including a disclaimer in their marketing materials stating that a tool "is not designed to be a primary factor in any decision," and for deployers to avoid compliance by having a human rubber-stamp algorithmic recommendations.

Bear in mind that consumers and workers frequently are not even aware when companies are using AI in their decision processes. Consequently, the "substantial factor" requirement creates a catch-22 that effectively makes it impossible for anyone to challenge a deployer or developer's assertion that an AI system is exempt from the law. Once a company chooses to assert that a tool does not meet the "substantial factor" requirement, it need not even disclose the existence or use of the AI tool. In that case, consumers, workers, and regulators may not even be aware of the tool, and thus will not be able to challenge the employer's assertion that the tool is not being used in a manner that has a substantial impact on a decision. In effect, that means deployers and developers would have the unilateral ability to opt out of complying with the law.

Relatedly, a recent study by researchers at Cornell University, Consumer Reports, and Data & Society showed that few companies are conducting audits or making disclosures required by New York City's AI hiring law, which similarly is limited to tools that have a dominant effect on the decision-making process.

Recommended Revisions

- Revise the definition of "substantial factor" to align with the Connecticut bill's definition, so that it covers any use of AI that could alter the outcome of a consequential decision.

Ensure the bill does not undercut or cause confusion with respect to existing civil rights laws and other laws or regulation

In its current form, SB 24-205 threatens to confuse and potentially undermine existing civil rights laws. The bill would introduce the new term "algorithmic discrimination" to the legal code, giving it a definition and exceptions that don't correspond with existing civil rights laws. Moreover,

rather than firmly requiring companies to ensure that their AI systems do not cause violations of civil rights laws, the bill would simply require companies to take "reasonable care" to prevent algorithmic discrimination and grant them a rebuttable presumption that such care was taken if the employer adopts some limited safeguards. That could easily lead to different and potentially contradictory standards for algorithmic and non-algorithmic discrimination.

The bill compounds this problem by assigning enforcement authority solely to the Attorney General (AG) and district attorneys. The Colorado Civil Rights Division should, at a minimum, be given authority to issue rules interpreting the meaning of terms relating to discrimination and enforcement authority over claims involving algorithmic discrimination. Failure to do so will lead to errors and create confusion and inefficiency if an AI system ends up being subject to a discrimination claim. Similarly, the Attorney General should be required to confer with the heads of state offices and agencies with primary responsibility for regulating the various types of consequential decisions that the bill covers to ensure that any regulations regarding SB 24-205 do not undermine or confound existing laws or regulations.

Expand the notice requirements so that consumers and workers actually receive meaningful notice of how a tool works

Unlike with traditional decision-making processes, people frequently do not know when AI is evaluating them, much less how they will be evaluated. Without strong notice provisions, citizens will be unable to exercise their rights under existing laws. SB 24-205's requirement that deployers provide impacted individuals with a pre-evaluation notice that includes "a plain language description of" the system is unhelpfully vague, and the bill includes no post-assessment notice or explanation requirements whatsoever. The existing notice provisions would not provide consumers or workers with the information needed to determine whether a tool has violated or may violate their legal rights.

Consumers and workers need true transparency. This means, at a minimum, disclosure of each ADS a company uses on them, what types of decisions an ADS makes, what personal data and attributes an ADS uses, how it uses that information to make decisions, and explanations of adverse decisions. The bill should be amended to ensure individuals receive such notice.

<u>Recommended Revisions</u>

- Expand pre-evaluation notice requirements in section 6-1-1604(4) to give individuals the right to know what information the tool uses and how the tool uses that information to make a decision:

  [Deployer must provide the consumer with] a description, in plain language, of such high-risk artificial intelligence system, which description shall, at a minimum, include a description of

  (I) the personal characteristics or attributes that the system will measure or assess, the method by which the system measures or assesses those attributes

or characteristics, how those attributes or characteristics are relevant to the consequential decisions for which the system should be used,
(II) the system's outputs,
(III) The logic used in the system, including the key parameters that affect the output of the system;
(IV) the type(s) and source(s) of data collected from natural persons and processed by the system when it is used to make, or assists in making, a consequential decision,
(V) the results of the most recent impact assessment, or an active link to a webpage where a candidate can review those results,
(VI) any human components of such system, and
(VII) how any automated components of such system are used to inform such consequential decision.

Right to specific, accurate, and actionable explanation and barring tools whose output are not easily explainable

An algorithmic tool that is used to make consequential decisions about consumers should be able to produce specific and accurate explanations of why it generates the outputs it generates, such that individuals understand the output and whether that output is based on inaccurate information or inferences about them. When an ADS is used to make, or assist in making, a consequential decision, consumers should thus receive a simple, actionable explanation of why the decision was made. When a deployer cannot provide such an explanation, it should not use the ADS.

<u>Recommended Revisions</u>

- Add the following text to section 6-1-1604(4):

A deployer shall, within 14 days after an automated decision tool is used to make, or assists in making, a consequential decision, provide any natural person that was the subject of the consequential decision:

A simple and actionable explanation that identifies the principal factors, characteristics, logic and other information related to the individual that led to the consequential decision;

The role that the automated decision tool played in the decision-making process; and

A meaningful opportunity to submit corrections or otherwise provide supplementary information relevant to the consequential decision.

No deployer shall use an automated decision tool to make, or assist in making, a consequential decision if it cannot provide an accurate notice that satisfies the requirements of [the above paragraph].

Require notices/explanations to be prepared and presented in a manner that ensures effective disclosure

The present language of the bill allows deployers to choose the manner in which they provide notices, subject only to the requirement that they choose a method of disclosure that is "clear and readily accessible." This language is far too vague and permissive; it does not ensure that consumers or workers will receive actual notice of the information that the bill ostensibly requires deployers to provide them, much less ensure that it is presented in a manner that consumers and workers will easily understand. Instead, The bill should require deployers to take specific steps to ensure that Colorado residents subjected to consequential decisions by AI systems receive actual notice in languages and formats that clearly and effectively convey the required information.

Recommended Revisions

- Replace language in section 6-1-1604(4) allowing deployer notices to be presented "in a manner that is clear and readily accessible" with a requirement that notices be:
    - *Directly to the consumer;*
    - *In plain language;*
    - *In all languages in which such deployer, in the ordinary course of such deployer's business, provides contracts, disclaimers, sale announcements and other information to consumers; and*
    - *In a format that is accessible to consumers with disabilities.*

Require that deployer impact assessments be conducted by an independent third party

The bill seeks to address potential harms of AI systems via impact assessments, but the current permissive language will allow those impact assessments to be ineffective. There have already been multiple instances where vendors published misleading impact assessments, in which the company either conducted the impact assessment themselves and seemed to cherry-pick the data points to present or retained a third party that was only granted partial access to relevant data. Such an impact assessment is not reliable. Third-party auditors who have full access to AI systems and are free of conflicts of interest are more likely to analyze and publish truthful assessments.

Recommended Revisions

- Revise section 6-1-1603(3) so that deployer impact assessments must be conducted by an independent entity, using the definition of independent auditor from the Lawyers' Committee for Civil Rights Under Law's Model AI Bill:

    For purposes of this subsection, an "independent auditor" means an independent person or entity who exercises objective and impartial judgment on all issues within the scope of the impact assessment or review. Each deployer or developer of the automated decision tool shall provide the independent third party with all information and data regarding the

design, functionality, testing, and performance of the high-risk artificial intelligence system. A person or entity is not independent for purposes of this subsection if they:

Are, or at any point during the five years preceding the impact assessment or review were, involved in using, developing, offering, licensing, or deploying the high-risk artificial intelligence system;

Have, or at any point during the five years preceding the impact assessment or review had, an employment relationship with a developer or deployer that uses, offers, or licenses the high-risk artificial intelligence system; or

Have, or at any point during the five years preceding the impact assessment or review had, a direct financial interest or a material indirect financial interest in a developer or deployer that uses, offers, or licenses the high-risk artificial intelligence system.

Addressing harms other than discrimination

AI systems can cause a wide range of harms beyond discrimination, including unfair or deceptive trade practices; emotional, physical, financial, and reputational injuries; and invasions of privacy. Additionally, employers have used AI in ways that violate workers' legal rights in other ways, such as by threatening their health and safety and depriving them of wages. The bill's impact assessment requirements should be expanded to require an analysis of whether the use of an AI system is likely to result in any of these other harms as well.

Recommended Revisions

- Revise section 6-1-1603(3)(b)(II) to require impact assessments to analyze whether the AI system poses reasonably foreseeable risks of these harms, in addition to discrimination:

  limits on accessibility for individuals who are pregnant, breastfeeding, or disabled, and, if so, what reasonable accommodations the deployer may provide that would mitigate any such limitations on accessibility;

  any unfair trade or deceptive trade practice under section 105 of part 1 of article 1 of title 6;

  any violation of state or federal labor laws, including laws pertaining to wages, occupational health and safety, and the right to organize;

  any emotional, financial, mental, physical or reputational injury to consumers that may be redressed under the laws of this state; and

  any physical or other intrusion upon the solitude or seclusion, or the private affairs or concerns, of consumers if such intrusion (i) would be offensive to a reasonable person, and (ii) may be redressed under the laws of this state

Establishing meaningful enforcement

The bill's enforcement and remedy provisions are currently far short of what is needed to deter companies from violating the law, a flaw that could render the bill's protections practically meaningless if left uncorrected. While the bill makes violations of the bill violations of Colorado's unfair trade practices law, it vests enforcement authority solely in the attorney general and district attorneys, understaffed public officials with many other demands on their scarce resources. That is a recipe for ineffective and inadequate enforcement.

Recommended Revisions

- Allow consumers and workers to have a private right of action so that individuals can seek remedies and vindicate their own rights, rather than relying on overstretched enforcement agencies;
- Impose penalties substantial enough that companies stand to lose more by ignoring the bill's notice requirement than they would gain by hiding a potentially discriminatory system's existence; and
- Raise the maximum fine for failing to conduct a compliant impact assessment to an amount that exceeds the expected cost of conducting such an impact assessment.

Eliminating loopholes and carve-outs that would greatly undermine the bill's effectiveness

The bill also contains numerous loopholes, exemptions, and carve-outs that would make it too easy for companies to avoid accountability or opt themselves out of compliance. In many cases, these provisions would negate the effectiveness of the bill's protections. For example, the current text includes:

- Novel presumptions and affirmative defenses in favor of developers and deployers that are both broad in scope and vague in their outer bounds and that would make it even easier for companies to escape the bill's already-insufficient penalties.
- A "right to cure" for the first year after the law goes into effect that seemingly would cover even willful violations of the law.

These loopholes and exceptions must be eliminated, or significantly narrowed and clarified so that they are precisely targeted to cover only clearly delineated types of inadvertent violations, lest they swallow the protections that the bill otherwise would afford consumers and workers.

Conclusion

It is not merely possible, but imperative, to do AI regulation well. To that end, we urge the Committee members to consider the recommendations in this document, as well as those submitted by other labor and consumer advocates, as it considers this bill. Please do not hesitate to contact Matthew Scherer, Senior Policy Counsel on CDT's Privacy & Data Project, at mscherer@cdt.org if you have any questions regarding this testimony or if we can provide

additional resources or information that will assist in your consideration of this legislation. Thank you for your attention.

Consumer Technology Association™

1919 S. Eads St.
Arlington, VA 22202
703-907-7600
CTA.tech

April 24, 2024

The Honorable Robert Rodriguez
Colorado Senate
200 E. Colfax Avenue
Denver, CO 80203

Re: Senate Bill 24-205, A Bill for An Act Concerning Consumer Protections in Interactions with Artificial Intelligence Systems – OPPOSE

Dear Majority Leader Rodriguez,

The Consumer Technology Association ("CTA")® respectfully presents its concerns and opposition to Senate Bill 24-205, titled "A Bill for An Act Concerning Consumer Protections in Interactions with Artificial Intelligence Systems," ("SB 205") now pending before the Colorado legislature.

As North America's largest technology trade association, CTA *is* the tech sector. Our members are the world's leading innovators – from startups to global brands – helping support more than 18 million American jobs. CTA owns and produces CES® – the largest, most influential tech event on the planet. CTA and its members thus have a substantial interest in the Colorado legislature's proposal to impose new regulations on artificial intelligence ("AI").

AI and associated technology (such as machine learning and neural networks) offer tremendous opportunities for human and societal development. AI is already demonstrating that it can promote inclusive growth, improve the welfare and well-being of people, and enhance global innovation and productivity. AI also has profound promise for our national interests.[1]

CTA is concerned that SB 205 would impose significant new duties on developers and deployers of AI which are serving national and international markets and would effectively mandate strict new compliance obligations that would reach far beyond Colorado. Further, we are also concerned that heavy-handed regulation of this still nascent technology will hamper innovation and investment in the market. Finally, the costs of complying with burdensome, top-down regulations like those

[1] Continued American leadership in AI is of paramount importance to maintaining the economic and national security of the United States and to shaping the global evolution of AI in a manner consistent with our Nation's values, policies, and priorities. Exec. Order No. 13859 84 C.F.R. 3967 (2019).

outlined in SB 205 will affect small and medium-sized enterprises the hardest, thereby further limiting potential innovation and competition in the market.

In addition to these concerns, there are several problematic provisions in SB 205:

- **Definition of High-Risk Artificial Intelligence System ("HRAI") is too broad.** SB 205 defines AI to include systems that make or are "a substantial factor in making" consequential decisions. Including HRAI systems that are "a substantial factor" in making consequential decisions under this definition is likely to cause confusion within the market, and ultimately lead to a chilling of American AI innovation.

- **Mandating disclosure of information available to the public could raise other risks.** SB 205 would require developers to make publicly available certain information, including how the developer "manages known or reasonably foreseeable risks of algorithmic discrimination." Such information could expose developers to liability if the information is misconstrued, and it could expose developers and deployers to security risks, if – in disclosing information about how the AI system operates – developers inadvertently disclose information that could be useful to bad actors.

- **The mandate for developers to disclose certain information to deployers of high-risk AI systems is ill conceived.** SB 205 would require developers to make extensive information available to users of high-risk AI systems, including information about known and reasonably foreseeable risks. This could force developers to disclose information that, if further disclosed, could be misconstrued by third parties. Further, certain elements of the disclosure mandate in Sec. 1602(2) may be outside the scope of the developer's knowledge, including the system's intended outputs. Thus, while deployers would have access to this information because they will be using these AI systems in their own business applications, developers may not.

- **Mandate for disclosures regarding General Purpose AI models are overbroad.** SB 205 would require developers of General Purpose AI systems to maintain extensive documentation, including "the tasks the General Purpose AI" model is intended to perform. Due to the nature of General Purpose AI models, specifically that they can be used to develop any number and type of specific applications, it is unlikely that developers of General Purpose AI models will have the information available to determine every possible task that their models can perform. Moreover, developers of General Purpose AI models would also be required to make extensive disclosures to deployers, including "key design choices" and "what the GPAI is designed to optimize for and the relevance of different parameters." Disclosing such detailed information may place developers at a competitive disadvantage.

- **The standard of care is vague and ambiguous.** SB 205 requires developers and deployers of AI systems to use "reasonable care" to protect consumers from certain harms. The standard articulated here is vague and ambiguous considering the rapidly evolving nature of the technology at issue, and the attendant strategies and practices for mitigating and reducing risks. What may be reasonable now may not be at a later point in time as technology evolves.

- **The duty to disclose certain information to the Attorney General, consumer protection authorities and all deployers is problematic**. Sec. 6-1-1602(5) requires developers to affirmatively disclose to the Attorney General and "all known deployers" of a high-risk AI system information about algorithmic discrimination caused, or likely to have been caused, by such system after testing or learning that a deployer's use has caused such harm. This provision is problematic because developers may not be able to assess the credibility of reports from deployers about such potential harms, which will be fact-specific and may involve confidential deployer or end-user information. Further, one deployer's use of an AI system that results in such harm does not necessarily indicate that all other deployers will use the system in the same or similar manner.

- **Disclosure requirements are too broad and likely to cause confusion.** SB 205 would require any person who provides *any AI system* – not just a high-risk AI system or generative AI system – that interacts with consumers to disclose to the consumer that he or she is interacting with an AI system. Because AI system is defined broadly to include a range of automated systems and processes – including many systems that do not present risks of consumer harm – this requirement risks unnecessarily confusing consumers who may falsely believe that they are at risk.

For the above reasons, CTA respectfully opposes Senate Bill 205.

Respectfully submitted,

/s/ Douglas K. Johnson
Douglas K. Johnson
Vice President, Emerging Technology Policy
djohnson@cta.tech

Notes on SB 24-205, A Bill For an ft Concerning Consumer Protections in Interactions with Artificial Intelligence Systems

Introduction

I originally set out to critique SB 24-205, which purported to be a consumer protection bill focused on algorithmic discrimination from AI, and this I have done. I am not a lawyer, but I have some legal training. I am not involved in building any of the current AI products, but I built my first neural network (an AI technology) over 30 years ago, so I have some useful technical experience.

After beginning this analysis, I saw a near-identical bill out of Connecticut[1]. This makes me wonder whether Colorado copied Connecticut or vice versa, or whether the two bills have a common origin. If common origin, what is that origin? It could be some third state, but it seems also highly likely that it is some undisclosed lobbying or interest group. The Colorado bill is sponsored by Senator Robert Rodriguez, and I am hopeful that he discloses the origin of the bill so that we can fairly evaluate the actual interests that Senator Rodriguez is representing by sponsoring this bill.

I am concerned about the source of the bill because the way that uniform laws among the states are often created is through the Uniform Law Commission.[2] ULC has been around since 1892, first as NCCUSL and then renamed as the Uniform Law Commission. Some of the 400+ laws that originated from ULC include the Uniform Commercial Code, Uniform Trade Secrets Act, Uniform Child Custody Jurisdiction and Enforcement Act, and many others. The reader will readily recognize the superiority of one-stop shopping for uniform statutes in all states by working through ULC as opposed to piecemeal ad hoc copying of statutes from Connecticut or random unnamed interest groups, or whatever. So why is Senator Rodriguez going the piecemeal ad hoc route with this bill?

Text of the Bill

This section of this document contains notes on the Bill organized under headings with names that mirror sections of the Bill. Where text is quoted without a citation, it refers to text in the

[1] https://www.cga.ct.gov/2024/FC/PDF/2024SB-00002-R000188-FC.PDF
[2] https://en.wikipedia.org/wiki/Uniform_Law_Commission

section of the Bill discussed in that section of this document. See, e.g., “The bill requires… system” in Bill Summary.

Bill Summary

“The bill requires a developer of a high-risk artificial intelligence system (high-risk system) to use reasonable care to avoid algorithmic discrimination in the high-risk system.”

6-1-1601. Definitions

“(1)(a) 'Algorithmic discrimination' means any condition in which an Artificial Intelligence System materially increases the risk of an unlawful differential treatment or impact that disfavors an individual or group of individuals on the basis of <list of factors>”

Many AI systems (and non-AI systems) will materially increase risk of unlawful discrimination as a side effect of performing otherwise permissible, socially valuable, and desirable operations. A system that provides complete information to a business on how to comply with the Americans with Disabilities Act for a potential employee with a disability can make it easier to hire that employee and efficiently achieve ADA compliance. In doing so, it can also expose large costs that the business will incur by hiring that employee, thus discouraging the business from hiring employees that this Bill and other statutes already in Colorado law[3] purport to protect. This obviously creates, in the language of the definition, “*any* condition in which an AI materially increases the risk” of unlawful discrimination.

Increasing the regulatory burden and compliance risks for taking steps that someone doing business in this state might take in a good faith effort to avoid discrimination discourages actions that this Bill is clearly trying to encourage. The use of “any” with no consideration of ways to use intent, giving greater weight to primary versus secondary effects, balancing equities, considering the interactions with existing anti-discrimination statutes[4], and so forth expands the definition of prohibited “algorithmic discrimination” to an unworkable scope.

[3] See, e;g., Colorado Anti-Discrimination Act at Colorado Revised Statutes § 24-34-301 et seq., C.R.S. § 24-34-301 et seq. and others.
[4] E.g., C.R.S. § 24-34-400.2 et seq.

"(1)(b)(I)(A)-(B) 'Algorithmic Discrimination' does not include the offer, license, or use of an Artificial Intelligence System by a developer or deployer for the sole purpose of (A) the developer's or deployer's self-testing to identify, mitigate, or prevent discrimination or otherwise ensure compliance with state and federal law; or (B) expanding an applicant, customer, or participant pool to increase diversity or redress historical discrimination"

There is no reasonable public or business purpose in requiring that the "sole purpose" be *either* testing, mitigation, or compliance; OR expanding an applicant, customer, or participant pool to increase diversity or redress discrimination. Surely someone licensing an AI System to ensure compliance with state and federal law cannot become liable for Algorithmic Discrimination by also using the system to expand an applicant or participant pool to include persons who had been discriminated against. If nothing else, one form of testing is verifying that a particular AI system will actually expand an applicant pool, and that requiring that a Deployer have the "sole purpose" of either testing *or* verifying that an AI System will expand an applicant pool defeats that legitimate purpose.

"(2) 'Artificial Intelligence System' means any machine-based system that, for any explicit or implicit objective, infers from the inputs the system received how to generate outputs, including content, decisions, predictions, and recommendations that can influence physical or virtual environments."

Note that this is the definition on which all else depends. If it is done poorly, nothing else matters because this defines what is and is not regulated by every other part of this Bill.

"'Artificial Intelligence System' means any machine-based system…"

As noted above, the focus on AI systems explicitly restricts the listed types of discrimination to only where it originates with the highly specific computer technology of artificial intelligence. Typically, regulation is intended to reduce harm regardless of the means by which said harm is realized or delivered[5]. What is the public interest in reducing discrimination, but only where said discrimination is done with artificial intelligence?

[5] See https://en.wikipedia.org/wiki/Taylor_Swift_deepfake_pornography_controversy for a description of a specific harm against Taylor Swift whose perpetrators happened to use AI, but who could have used non-AI tools to the same end. Would Ms. Swift and others similarly situated be satisfied by the Bill's focus on the *means* of AI rather than focusing on the *harm* from the creation and publication of the offending material by whatever means. Updated to respond that the definition of Algorithmic Discrimination in this Bill includes "... impact that disfavors an individual or group of individuals." Need I defend the idea that Ms. Swift has suffered a disfavoring impact?

“Infers from the inputs the system receives how to generate outputs”

Everything infers from its inputs how to generate outputs. When the user turns the knob to Dark, inserts bread, and presses the handle down, a toaster may apply 1020 watts to the heating elements for 3 minutes, then pop the bread up out of the toaster and emit an optional sound. The user does not explicitly specify the power level, the amount of time, what to do after the toaster has applied that level of power for that amount of time, or the volume or tone of the sound. The toaster infers all of these things and more about how to “generate outputs” of toasted bread. Assuming that the Bill is supposed to impose a substantial regulatory burden on every technological artifact that is at least as complex as a toaster, the most important definition in the Bill is competently written.

There will be those who argue that this objection is as frivolous and poorly-considered as the clause that it critiques. Such defenders may note that 6-1-1609 states that “the Attorney General and district attorneys have exclusive authority to enforce” this and surely the AG and various DAs will not take action against every toaster manufacturer doing business in Colorado. While likely true, it is also true that the AG and DAs look to the text of statutes to divine the intent of the legislature in enacting the statutes. Where they find narrow and limiting language, they often take that as a signal that the legislature intended that a statute would have narrow application, and judges are encouraged to give the language narrow construction. Where, as here, they find broad language encompassing every piece of technology at least as advanced as a toaster, they may and should take that as a signal from the legislature that the statute is intended to be broadly interpreted and applied.

This signal from the legislature that the statute is intended to have very broad application will be heard by others outside of the AG’s office. Businesses and technology innovators will see this and understand that Colorado is enshrining in its statutory law the principle that technology innovations in this area will be scrutinized closely and likely unfavorably by people who think that toasters have achieved intelligence. Decisions about investment, business and job creation, and much more will follow.

Outputs: “Content, decisions, predictions, and recommendations”

The outputs that the Bill regulates are defined as “content, decisions, predictions, and recommendations.” Notably absent are actions such as invocations of APIs. Also absent are electrical, optical, or other signals. These omissions mean that AIs that control other machines by invoking an exposed API endpoint or controlling attached hardware by outputting an electrical or optical signal are not part of the definition.

Under the legal principle of ejusdem generis (“of the same kind”), using a word and then stating specific instances of the word will limit the meaning of the word to things that are very similar to the specified instances. Here, text, audio, and video are similar enough to “content” in “content, decisions, predictions, and recommendations” that text, audio, and video are covered by the Bill, but actions such as invoking an API and electrical signals are likely not covered. The legal principle of noscitur a sociis (“it is known by its associates”) further supports this interpretation because “content, decisions, predictions, and recommendations” are all things that would

typically conveyed to a human, whereas electrical and optical signals would instead be conveyed to a machine in a form that cannot be directly interpreted by a human and therefore would not be part of the list.

This limitation of the scope of the AI systems covered by the Bill seems contrary to the intent of the Bill to protect against algorithmic discrimination. An AI system may be embedded as a component of a medical device where the device does not directly communicate the AI's outputs to a human but instead incorporates them into other data before conveying the device's output to a human user. This has obvious implications for potential algorithmic discrimination[6], but the language of the Bill goes out of its way to exclude coverage of such embedded systems by covering only AIs with outputs "including content, decisions, predictions, and recommendations."

"(5) 'Deploy' means to use a Generative Artificial Intelligence System or a High-Risk Artificial Intelligence System"

The term "use" in this definition is not defined elsewhere in the Bill, so one assumes that it has its customary meaning. That means that at this point, most of the people in Colorado have deployed (used) a Generative Artificial Intelligence System by virtue of having simply toyed with ChatGPT, done an AI-powered Google search, used an iPhone, taken a photograph with a digital camera having AI image-enhancement capabilities, or any of the countless acts that we take every day. Because of this broad definition, almost everyone in Colorado has deployed (merely used) an Artificial Intelligence System.

It is axiomatic that if a law is written so that nobody knows whether an activity is covered by the law, nobody can take care to follow the law. Under this definition (5), "Deploy means to use" an AI, so everyone who uses an AI also Deploys an AI. (But see below the discussion of whether one who Deploys (5) an AI is necessarily a Deployer (6) of the AI.) What AIs have you, the reader of this document, deployed already today? Have you used an iPhone and, if so, did you use any feature that contains AI or did you restrict yourself to only those features that do not use AI? If you used a word processor with a grammar checker, was it the kind of grammar checker that uses AI or the kind that does not? If you do not know then you do not know whether you are covered by this law, and so you cannot know whether to follow the law. See also the discussion above about the hazards of expanding the scope of the language and thereby expanding the scope of interpretations and applications.

If you operate a business and you use software that did not have AI in it when you acquired it, has the vendor done a subsequent silent background update so that the software now does use AI? If so, be warned that ordinary business activities that were once not covered by the Bill would now be covered by the Bill and you are immediately required to comply with every

---

[6]See https://foundation.mozilla.org/en/blog/facial-recognition-bias/, https://www.scientificamerican.com/article/fixing-medical-devices-that-are-biased-against-race-or-gender/, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9372891/, https://ainowinstitute.org/wp-content/uploads/2023/04/disabilitybiasai-2019.pdf, and many others for examples of why we may want to regulate AI interacting with devices also instead of only AI interacting with humans.

requirement of 6-1-1602 through 6-1-1610 of this Bill and also the modification to 6-1-105 of the Colorado Anti-Discrimination Act (CADA) and also those sections of CADA that are implicitly modified by needing to be read in pari materia[7] with the related areas of this Bill. It's unclear how you as a business owner might stay out of trouble if you use any vendors outside of Colorado and therefore outside of the jurisdiction of this Bill because they have no duty to notify you of such changes. Be sure to stay abreast of the latest changes in the proprietary trade secret source code implementations of Windows, Excel, Gmail, Google Search, and others to ensure that you don't suddenly switch from using ordinary software to Deploying (using) AI software due to an update from your vendor.

What does it mean to use (Deploy) an AI System? Do I use/Deploy ChatGPT by typing a prompt in Colorado, or does the actual *use* of the AI take place on the server on which ChatGPT is running and performing AI algorithms, likely (but you don't know) outside of Colorado?

"(6) 'Deployer' means a person doing business in this state that deploys a Generative Artificial Intelligence System or a High-Risk Artificial Intelligence System"

A Deployer must be "doing business in this state." Definition (5) of Deploy makes no mention of doing business in this state or even using the AI in this state. As a result, it is possible to Deploy (5) an AI without being the Deployer (6) of the AI. This makes the Bill harder to read than necessary because it can be unclear whether the root word of related words such as "deployed" and "deploys" is "Deploy" from definition (5) or "Deployer" from definition (6). More careful attention to writing definitions (5) and (6) would be helpful.

(7) 'Developer' means a person doing business in this state that develops or intentionally and substantially modifies a General Purpose Artificial Intelligence Model, a Generative Artificial Intelligence Model, a Generative Artificial Intelligence System, or a High-Risk Artificial Intelligence System.

Scope of the AI System

Absent a usable definition of Generative Artificial Intelligence System or most of the other AI-related words in this definition, it is impossible to tell what it might mean, so here I am limited

---

[7] In pari materia is a legal doctrine that means that two Acts that address similar areas must be read and interpreted as a unit rather than separately. Words or ideas that are interpreted one way in one Act ought to be interpreted and applied the same way in the other Act. This bill must be read in pari materia with the Colorado Anti-Discrimination Act (CADA) for several reasons. First, the Bill Summary for this Act has as its purpose to "require[] a developer of a high-risk artificial intelligence system… to avoid algorithmic discrimination...". In other words, this Act explicitly addresses anti-discrimination including some of the exact areas such as employment (6-1-1601(3)(c)) that CADA addresses. This Act also explicitly modifies 6-1-105 CADA. All of this combines to make this Act and CADA textbook examples of Acts that must be read in pari materia. Because of that, maintaining compliance with this Act also requires considering developments in anti-discrimination law that may manifest in CADA interpretations and actions.

to pointing out what might happen under various definitions of those terms and leave it to others to select a definition.

One definition of an AI System might encompass only the AI system itself as shipped from OpenAI or another AI vendor. In that case, it seems that a Developer must work for the AI vendor.

Another definition of AI System may encompass the original AI system plus the custom rules that a user enters for the purpose of applying to all subsequent prompts. Because these custom rules have the purpose of affecting the output generated by the AI, one can imagine them being treated as part of the AI System. In that case, the user who entered the custom rules may be considered a Developer. What if the user who enters the custom rules did not write the rules, but merely selected them from a set of one or more custom rules written by others? Is the person who wrote the original rules a Developer? Both the user and the author of the rules are Developers? Neither? It depends? We need an answer.

Configurations and Other Factors that May Unexpectedly Transform a Person into a Developer

A key question is what is meant by "modifies," as in "intentionally and substantially modifies" an AI System. Much software ships with settings, configuration options, and parameters, and AI Systems are no different. For example, one such parameter is called temperature, which controls the extent to which the outputs from an AI are reproducible. An AI System with a temperature parameter ships from the AI vendor with that parameter in existence and set to a default value. Changing that value is an operation that takes no more than a very few seconds and it requires none of the specialized training, expertise, or other characteristics that one would expect of a Developer. However, changing the temperature value leads to substantial alteration in the output from the AI System, which is an effect that typically would require the intervention of a Developer.

Is one who substantially changes the behavior of an AI by changing things like the temperature parameter a Developer, with all of the responsibilities that this Bill attaches to a Developer? What mental state is required of such a person? Must a person who makes such a change be aware of the Developer-level consequences of the act before inheriting the many responsibilities of a Developer, or is ignorance of the operation of such a complex system a defense against being classified as a Developer? Must that person understand and accept that they are in a Developer role for legal purposes, or is mere performance of certain activities sufficient? If the standard written procedure at one's employer states that one should configure an AI in a specific way before operating it, can someone who substantially changes the outputs from an AI by merely following the standard written procedure be considered a Developer simply because that person blindly followed instructions that resulted in substantial changes to the behavior of the AI System? If not, does that mean that changes that are ordinarily attributed to a Developer may in fact have no Developer at all because of the separation of the author of a policy and the actor who actually carries out the policy, or does it mean that an otherwise blameless employee becomes a Developer simply by doing a job?

There are obviously many other questions related to this, none of which can be answered until this Bill defines what the scope of an AI System is. This is nontrivial at best because AI Systems span the range from standalone to deeply integrated, possibly with other AI Systems (at which point it will be useful to ask how long two interacting AI Systems are distinct and when they merge for some legal purpose). However, such a detailed definition is plainly required because questions such as "who is a Developer under definition (7)" cannot be answered without it, and it is clearly critical to answer that and related questions.

"8(a) 'General Purpose Artificial Intelligence Model' means any form of Artificial Intelligence System that (I) Displays significant generality; (II) Is capable of competently performing a wide range of distinct tasks; and (III) can be integrated into a variety of downstream applications or systems"

Quite a lot of software that would not be considered AI outside of this Bill satisfies this definition for purposes of the Bill. For example, operating systems such as Windows 7 or Linux display significant generality, competently perform a wide variety of distinct tasks, and are integrated into a variety of applications and systems, but are not generally considered AI Systems. See the discussion in Infers from the Inputs the System Receives How To Generate Outputs for why 6-1-1609. Enforcement by Attorney General and District Attorneys does not save overly broad definitions in the Bill from causing harm. A reasonable practice in statutory construction by courts, attorneys general, and others is to construe broad language broadly and narrow language narrowly, and this broad language encourages broad application in legal contexts, which sows enormous uncertainty in business and other social contexts.

"(9) 'Generative Artificial Intelligence System' means any Artificial Intelligence System, including a General Purpose Artificial Intelligence Model, that is able to produce Synthetic Digital Content."

Combining this definition (9) of a Generative Artificial Intelligence System with definition (13) of Synthetic Digital Content tells us that a "Generative Artificial Intelligence System means any Artificial Intelligence System… that is able to produce digital content" ←(9) … (13)→ "[that is] produced by a Generative Artificial Intelligence System." Circular definitions are per se unhelpful.

"10(a) 'High-Risk Artificial Intelligence System' means any Artificial Intelligence System that has been specifically developed and marketed, or intentionally and substantially modified, to make, or to be a substantial factor in making, a consequential decision."

This section brands many Colorado-developed artificial intelligence systems as high-risk whereas comparable systems developed elsewhere are not required to brand themselves as

high-risk. One can readily imagine that this will create at least market confusion if not outright market hostility toward products from Colorado businesses.

See the discussion in definition 11, below, definition 7 of Developer, above, and specifically in Configurations and Other Factors that May Unexpectedly Transform a Person Into a Developer, above, to see the consequences, unanticipated and otherwise, of the use and specific definitions of the term "intentionally and substantially modified" in this definition. It is easy to see how a particular AI System may unexpectedly become or not become a High-Risk AI System depending on subtle choices in how this term is defined. Note, of course, that "subtle choices" is a shorthand way of saying "lethal poison for the predictability and consistent application of a rule that business and other types of decisions require."

"(11) 'Intentional and Substantial Modification' or 'Intentionally and Substantially Modifies' means a deliberate change made to: (a) a Generative Artificial Intelligence System, other than a change made to a Generative Artificial Intelligence System as a result of learning after the Generative Artificial Intelligence System has been Deployed, that: (I) affects compliance of the Generative Artificial Intelligence System; or (II) changes the purpose of the Generative Artificial Intelligence System"

See the discussion above for some of the implications of various plausible definitions of "Intentional and Substantial Modification." This bill has no usable definition of the scope of what is considered an AI system, especially as regards configuration parameters, boundaries where multiple AIs are incorporated into the same system, and many other considerations discussed herein. Absent a definition of exactly what an AI System is, there can be no discussion of the bounds of the system and whether a particular modification is within that bound (modifies an AI) or outside the bounds (does not modify an AI).

What does "deliberate change" mean here? If an action is taken by a person who has no ability to understand the implications of the action then law would typically not regard the action as deliberate and would not hold that person responsible for the action. For example, someone who deliberately hands me a glass of liquid believing that it is water is generally not responsible for what happens to me when the liquid turns out to be deadly poison. Here, if a person makes a change to an AI such as setting a configuration parameter or other apparently innocuous change made deliberately but without knowledge of the full consequences, has that person made a deliberate change within the meaning of this bill?

One could reasonably suggest that only a Developer or certain sophisticated Deployers can make an intentional and substantial modification because only one with the sophistication of a Developer or Deployer can be said to make deliberate changes as the term "deliberate" is typically used in law.

“(13) ‘Synthetic Digital Content’ means digital content, including audio, images, text, or videos, that is produced by a Generative Artificial Intelligence System.”

See the discussion at definition 9 of how this definition 13 of Synthetic Digital Content interacts with definition 9 of Generative Artificial Intelligence System in such a way as to render both definitions meaningless.

“6-1-1602. Developer duty to avoid algorithmic discrimination - required documentation”

6-1-1602(1)-(2) A Developer of a High-Risk Artificial Intelligence System shall use reasonable care to protect consumers from any known or reasonably foreseeable risks of Algorithmic Discrimination… and a Developer shall make available to a Deployer [various types of documentation]

A risk of Algorithmic Discrimination arises only in certain contexts. It is well understood that ChatGPT has a risk of Algorithmic Discrimination in making certain evaluations for employment, but it has little or no risk of Algorithmic Discrimination in describing the most common process for shearing sheep or recommending whether a person should flee from a rabid grizzly bear.

The context that determines the degree of risk of Algorithmic Discrimination, if any, arising from the use of a specific AI includes elements that are customer-specific. As noted in my comments about on the definition of Algorithmic Discrimination, an AI that helps a customer determine the steps that the Americans with Disabilities Act requires to support the hire of a specific employee with a covered disability may help a customer comply with the ADA or may show that compliance will be very expensive and cause the customer to not hire employees with certain types of disabilities. Most vendors or Developers simply cannot meet the standard announced in this section of protecting consumers from “*any* known or foreseeable risks” because those risks are beyond the control and even view of the vendor/Developer.

Because many of the risks of Algorithmic Discrimination are customer-specific rather than arising from any particular AI itself, the requirement of this section that “*any* known or reasonably foreseeable risks” be identified, mitigated, and otherwise handled by Developers instead of Deployers or other actors in the system is per se unworkable.

Because the risks of Algorithmic Discrimination are context-specific, an AI vendor seems to have only two natural ways to comply with this section, those being to provide customer-specific documentation or to provide customer-agnostic documentation. Each is described in the subsections that follow.

Customer-Specific Documentation

First, the vendor must know all purposes to which a potential customer will put the vendor's AI so that the vendor may tell each customer individually the risks that arise from the use of the AI in terms of that specific customer's use cases in terms of regulatory, safety, commercial, and other customer-specific context elements. This clearly precludes various customers from using Colorado-developed AI at all, including government customers who might use it in working with classified information that they cannot disclose to the vendor, companies who want to use the AI in working with trade secrets that the customer cannot disclose to the vendor without losing trade secret protection, and others.

This first alternative also seems to preclude the most common ways of selling or distributing software. The apparent requirement to gather detailed information about the customer environment as part of each sale seems to preclude mass distribution such as web-based sales and distribution. It also seems to invalidate the most common software support models because many software updates have the potential to introduce new risks for specific customers and customer types. This seems to require updating detailed customer use case information before each update that may introduce a risk of Algorithmic Discrimination. This is true even if the new risk may affect only a very small subset of customers because this section seems to require the vendor to ensure that it has identified and notified any affected customers and it is difficult to see how that can be done without auditing the use cases of every customer.

Finally, the vendor must secure and enforce a requirement that the customer allow ongoing auditing of how a given customer is actually using the AI so that the vendor may comply with its obligation to supply a customer with "*any* known or foreseeable risks of Algorithmic Discrimination" in the customer use environment as that environment and associated sets of use cases changes. It is beyond obvious that few customers will find this acceptable and most or all will surely choose to either forgo using AI altogether or choose an out-of-state vendor for their AI solutions.

Customer-Agnostic Documentation

In a second alternative, an AI vendor who is not able to give and maintain the specific individual advice required by the first alternative must publish all foreseeable risks of Algorithmic Discrimination that could affect any potential customer. Here arises a question of granularity and generality/specificity. A fine-grained list of very specific risks is so vast and ever-growing that the most likely result is that any useful information is buried in an avalanche of literally millions of disclosable risks. As the required list of disclosures moves away from specificity toward a coarse-grained list of general risks, it also moves toward a smaller list of meaningless generalities that helps no one.

Conclusion

In summary, 6-1-1602(1)-(2) requires vendors to place themselves on a continuum with the endpoints of imposing unacceptable requirements that seem designed to drive customers to

out-of-state vendors at one end and compliance that is as meaningless as it is useless at the other end.

6-1-1602(2)(c) A Developer of a High-Risk Artificial Intelligence System shall make available to the Deployer of the system... documentation describing (I) the type of data used to train the... System, (III) the data governance measures used to cover the training datasets and the measures used to examine the suitability of data sources, possible biases, and appropriate mitigation.

The process of training GPT4 exceeds $100M and future generations may cost up to $7 trillion. For those AI projects that need to continue to maintain access to current-generation models but that cannot raise $7 trillion[8], they must either build on top of AIs from one or more of a very few very well-funded AI vendors or not even try.

This dependence means that users of those large systems from wildly well-funded vendors do not have the documentation required by this section because those well-funded vendors do not release such data at any meaningful level of detail for the AIs that almost everyone needs to use for their own AI products. 6-1-1602(2)(c)(I) requires information about "the type of data used to train the System?" Forget it. 6-1-1602(2)(c)(III) wants documentation about "governance measures used to cover the training datasets and… measures used to examine the suitability of data sources, possible biases, and appropriate mitigation?" Too bad. You lose.

Because of the requirement to produce documentation and information about an entire AI system rather than just the portion that was added by a specific Developer for a specific product or purpose, compliance with 6-1-1602(2)(c) is impossible for the foreseeable future as written..

Having said that, there are parts of 6-1-1602(2)(c) that are within the control of an independent Developer or vendor of an AI System. (c)(II), (IV), and (V) as written do not necessarily depend on cooperation from the vendor of the base AI.

Also, just because 6-1-1602(2)(c) is impossible for the foreseeable future as written does not mean that it cannot be rewritten. Restricting subsection (c) to only data or training sets that are provided by the vendor of the end-user AI System can be done, although one can imagine some vendors legitimately wanting to raise trade secret arguments, and drafters of a revised section 6-1-1602(2)(c) ought to consider those concerns during the redrafting. Even adding the phrase "to the extent feasible" that appears now in 6-1-1602(3) could be an improvement, though drafters may want to consider other language if it is judged that a feasibility exception will simply serve as a universal release for recalcitrant vendors.

[8] All of them.

6-1-1602(4)(a) Each Developer shall make… readily available for public inspection… a statement summarizing the types of High-Risk Artificial Intelligence Systems that such Developer has developed or Intentionally and Substantially Modified, and currently makes available to Deployers

If a Developer is trying to sell or take to the general market an AI system of any sort, high-risk or otherwise, said Developer will already make this information available. Hiding their products from view is not how vendors convert potential customers into paying customers.

If a Developer is not publicly announcing an AI that it has developed and made available to a Deployer, it has a reason. Perhaps it is a custom development shop whose customers (Deplyers) do not want a Developer to announce the product that is conveying a substantial competitive advantage to said Deployer. In that case, this section amounts to imposing a substantial disadvantage on Colorado-based Developers of custom AI systems.

Perhaps a Developer is trying to end-of-life a product, so it does not want to publicly advertise it, but continues to make it available to existing customers/Deployers, thus triggering this section's requirement that the Developer must continue to make information about the obsolete product "readily available for public inspection." In that case, this section amounts to a requirement that a Developer present confusing information to the public about products that the public can no longer get.

Other scenarios with similarly poor outcomes arise from the requirements of this section. It is unclear whether the legislature wishes to disadvantage Colorado-based AI Developers, to require Developers to maintain information that will confuse the public, or other readily foreseeable negative outcomes. Regardless, it is hard to see how any of these make for even acceptable, let alone good, public policy goals.

6-1-1602(5) The Developer of a High-Risk Artificial Intelligence System shall disclose to the Attorney General and all known Deployers of the High-Risk Artificial Intelligence System any known or reasonably foreseeable risk of Algorithmic Discrimination arising from the intended uses of such High-Risk Artificial Intelligence System not later than ninety days after the date on which such Developer discovers (a) through such Developer's ongoing testing and analysis… (b) or receives from a Deployer a credible report that such High-Risk Artificial Intelligence System has caused, or is reasonably likely to have caused, Algorithmic Discrimination

Explicit Requirements

As the discussion in 6-1-1602(1)-(2) makes clear, the requirement to disclose *any* known or reasonably foreseeable risk requires either customer-specific or customer-agnostic disclosures.

The discussion concludes by noting that, while each approach imposes different harms, neither approach can work. Limiting this requirement to disclosing only a new harm that gave rise to the reporting requirement could possibly be made to work, but one would want to see the specific language for that different requirement before making a judgment.

As everyone knows, AI systems are often Deployed as a component in a larger system that may include other AI and/or non-AI components. It is entirely possible, even to be expected as a routine and mundane matter, that a problem will arise in a large system of which a specific vendor's AI is but one of many components. Consider a harm that arises only in a specific configuration of the vendor's AI together with other components from other vendors, possibly further limited to occurring only with a specific customer configuration.

As a non-AI example[9], note that the infamous Microsoft Windows Blue Screen of Death (BSOD) is caused (by an enormous margin) far more often by defective third-party software than by a failing in Microsoft's software, but it is Microsoft that gets the report. Note also that Windows is always Deployed with at least hundreds of third-party components. Does each report of a Windows problem that could credibly be due to a Microsoft failing require Microsoft to spin up an investigation with the resources to conclusively lay blame either with Windows or elsewhere within 90 days for every such report worldwide? If not, must Microsoft send a 6-1-1602(5) notification to "the Attorney General and every known Deployer" of Windows for each Windows fault that cannot be traced elsewhere? Has this bill given the Attorney General a multibillion dollar budget solely for the purpose of handling such an avalanche of incoming reports?

Note that one of the largest AI systems supported by one of the largest companies in the world, is Meta's Llama open source AI. Therefore, it is obvious that many AI projects will be created by open source Developers. Except not in Colorado, because no open source free software Developer has anywhere near the resources to meet even the explicit requirements of this section, let alone the implicit requirements discussed below. And does the Attorney General want to hear about every report filed with Meta from an almost endless army of Developers and Deployers? Is it even possible that the Attorney General will ever have the resources to do so?

But even Meta does not have infinite resources. Can even Meta respond within 90 days to every report that it gets from an almost endless army of Developers and Deployers about some problem that allegedly originates with Llama? Or would it be better for them to withdraw Llama from Colorado?

Is there any reasonable scenario in which complying with the requirements of this section is even possible?

Implicit Requirements?

In addition to the numerous regulatory and reporting burdens that this section loads onto Developers and the luckless Attorney General, it is unclear whether this section also seeks to

[9] A non-AI example for now, but not for long. The pace of Microsoft's adoption of AI into all of its products makes the realization of this example merely a matter of a very short time.

implicitly impose additional requirements on a Developer. A few of these implicit requirements are discussed in this subsection.

“The Developer… shall disclose to… all known Deployers”

Does this create an implicit requirement that all Developers must maintain a list of all Deployers and keep the contact information for said Deployers up-to-date at all times? Must a small business or free/open source Developer keep its Deployer list up-to-date even as Deployers are acquired by or merged into other entities? If a human Deployer contact leaves the company that did a Deployment, is there a further implicit requirement that a small business meet the requirement that it shall, absolutely, with no provision for best efforts or commercially reasonable efforts or any other standard, disclose to Deployer? At a bare minimum, understanding that the real requirement that is not stated in this requirement is for best efforts or commercially reasonable efforts or some other standard, does this standard intend to implicitly require that all Developers must retain legal counsel at their own expense to determine the actual requirement of this section? Wouldn’t it be better to draft the section correctly and tell us what the requirement really is, given that it is definitely not an unqualified “shall”?

“The Developer... shall disclose... not later than ninety days after the date on which such Developer discovers (a) through such Developer's ongoing testing and analysis… (b) or receives from a Deployer a credible report that such High-Risk Artificial Intelligence System has caused…”

Does this create a requirement that all Developers will maintain “ongoing testing and analysis?” Most Developers do not do this, nor would it make any sense, nor can they if they wanted to. Most Developers run a product through a QA process before release, release it, and then test the next version that is scheduled for release, but do not continue to test the product that has already been certified as release-grade. If this section does seek to create such an implicit requirement, what type and level and other characteristics of testing satisfy this secret requirement? This is a crucial question because 6-1-1602(1) creates a safe harbor, saying “there shall be a rebuttable presumption that a Developer used reasonable care as required under this subsection *if* the Developer complied with the provisions of this section.” What level of “ongoing testing and analysis” is required for a Developer to comply with this section and gain access to the safe harbor? Again, please consider drafting this section to tell us what the requirement for “ongoing testing and analysis” actually is, or whether there even is such a requirement, or whether it is not required but if we do it then we get into the safe harbor, or whether it’s something else entirely. Because I cannot run my business without knowing this.

Conclusion

Complying with even the explicit requirements of this section is near impossible for anyone. The situation is not improved by adding in a need to also comply with implied requirements under threat of not having access to the safe harbor if we do not divine what those secret requirements might be.

6-1-1602(7) The Attorney General may require that a developer disclose to the Attorney General, in a form and manner prescribed by the Attorney General

To be able to guarantee future compliance with this requirement, it is necessary for the Attorney General to publish in advance the “form and manner” that will be “prescribed by the Attorney General.” Absent that, those covered by this section do not know what records to keep and other information to retain so that they can be responsive to a demand from the Attorney General.

The Attorney General should also be under an affirmative requirement to keep the “form and manner” that may be required up-to-date and present in a publicly accessible location, preferably on a web site.

6-1-1607(3)(b). Deployer duty to disclose synthetic digital content to consumer - exemptions; a person holds editorial responsibility for the publication of the synthetic digital content

It is often possible to mitigate editorial responsibility by using disclaimers. Please consider whether it is sensible to make the editorial responsibility non-disclaimable.

6-1-1609(2). Enforcement by attorney general and district attorneys; During the period beginning on July 1, 2025, through June 30, 2026, the Attorney General or a District Attorney, prior to initiating any actioun for a violation... shall issue a notice of violation to the Developer or Deployer alleged to have committed the violation if the Attorney General or District Attorney determines that the opportunity to cure is warranted.

Why does this go away on June 30, 2026? Giving the opportunity to cure is optional based solely at the discretion of the Attorney General or District Attorney. Why take away that discretion?

TECHNET
THE VOICE OF THE INNOVATION ECONOMY

TechNet Central | Telephone 720.308.0842
P.O. Box 113, Littleton, CO 80160
www.technet.org | @TechNetCentral

ARTIFICIAL INTELLIGENCE: RESPONSIBLE DEVELOPMENT FOSTERS TECH INNOVATION

As the Colorado General Assembly considers omnibus AI regulation SB24-205, sponsored by Senate Majority Leader Robert Rodriguez, TechNet provides this additional background on the existing frameworks and laws in place to guide and provide guardrails for responsible AI development. AI is transformative technology that already provides solutions to consumers on a daily basis, and will continue to evolve to help us solve the most challenging problems we face. This memo also details only a few of the many AI solutions that our members' work provides to consumers and the general public. These examples show that responsible AI development provides immense benefits to Coloradans, including helping fight wildfires, detecting deepfakes, assisting individuals who are blind or have low vision, detecting financial fraud, and more. Please reach out to TechNet with any questions at rbarko@technet.org or learn more at AI4America.com.

Introduction to TechNet

TechNet is the national, bipartisan network of technology CEOs and senior executives that promotes the growth of the innovation economy. TechNet was founded in 1997 in Silicon Valley.

TechNet's diverse membership includes dynamic American businesses ranging from startups to the most iconic companies on the planet and represents over 4.2 million employees and countless customers in the fields of information technology, artificial intelligence, e-commerce, the sharing and gig economies, advanced energy, transportation, cybersecurity, venture capital, and finance. TechNet has offices in Austin, Boston, Chicago, Denver, Harrisburg, Olympia, Sacramento, Silicon Valley, and Washington, D.C. Our membership can be viewed here.

As the Voice of the Innovation Economy, TechNet advances public policies and private sector initiatives at the federal, state, and local levels that make the United States the world leader in innovation. We champion policies that foster a climate of innovation and competition, allowing America's tech industry to flourish. When policymakers are grappling with today's most transformative new technologies, they turn to us.

Recent Developments in Responsible AI Deployment

In January 2023, the National Institute for Standards and Technology released the Artificial Intelligence Risk Management Framework (AI RMF). The AI RMF aims to help designers, developers, deployers, users, and evaluators of AI systems better manage A risks that could affect individuals, organizations or society. The AI RMF is a sector

and use-case agnostic framework for managing AI risks. Many TechNet members utilize the AI RMF in evaluating their AI systems throughout their lifecycle.

- Currently, NIST is working to developing an addendum to the AI RMF focused on risks posed by generative AI.

In July 2023, the White House secured voluntary commitments from several leading AI developers to help move toward safe, secure, and transparent development of AI technology. These commitments included:

- Rigorously testing AI systems before release to ensure their safety and reliability.

- Collaborating with industry, government, and academic experts to share knowledge and identify potential risks.

- Allowing for independent review to uncover vulnerabilities.

- Prioritizing research on potential risks, such as bias and job displacement.

- Working to address society's greatest challenges, such as climate change and poverty.

In October 2023, President Biden released his executive order on the "Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence." The sweeping executive order includes 150 responsibilities for the Administration to undertake. Below are some of the key provisions from the Order that touch on the topics we will discuss during TechNet Day:

- Establishes a pilot program for the National AI Research Resource (NAIRR). TechNet is working to pass the *CREATE AI Act* to authorize this program.

- Includes several directives to develop benchmarks for AI with industry collaboration. For example, NIST is charged with developing best practices to "red-team" AI systems, which is a type of "white hat" hacking to test a system's vulnerabilities and capabilities before it is released to the public. The Order ultimately led to the creation of the U.S. AI Safety Institute, which is partnering with industry to establish benchmarks for responsible AI systems.

- The State Department is directed to establish a program to identify and attract top talent in AI and other critical and emerging technologies at universities, research institutions, and the private sector overseas. The Departments of State and Homeland Security are also directed to streamline and expedite visa petitions and applications for noncitizens who seek to travel to the U.S. to work on, study, or conduct research in AI. It also encourages the development of

American AI talent by providing AI training to federal employees and focused efforts to increase our supply of AI scientists and researchers.

In February 2024, Secretary of Commerce Raimondo announced the creation of the U.S. AI Safety Institute, which will support the government in ensuring AI systems are developed with trust and safety. Right now, there are not globally agreed upon metrics to test and evaluate AI systems. The USAISI will coordinate the development of American industry-leading metrics; having America's standard be the global standard is important for our continued leadership on responsible AI policy.

- The USAISI will partner with other U.S. government agencies on evaluating AI capabilities, limitations, risks, and impacts and coordinate on building testbeds. The institute will also work with organizations in ally and partner countries to share best practices, align capability evaluation, and red-team guidance and benchmarks.

- The AI Safety Consortium, which will be operated within the USAISI, includes over 200 companies partnering with NIST to bring expert perspectives for the development of responsible AI practices. Several TechNet members are a part of the consortium.

TechNet believes that media of candidates that is a "fraudulent misrepresentation" must be barred for the safety of our democracy. TechNet has urged the Federal Elections Commission (FEC) to use its existing authority to ban fraudulent AI-created misrepresentations of candidates by campaigns, candidates, PACs, and political parties. The technology industry is working together to set guardrails on synthetic media for the upcoming global elections and recently released a Tech Accord to Combat the Deceptive Use of AI in the 2024 Elections at the Munich Security Conference on February 16. Some of these commitments included:

- Flagging the Source (Provenance): Attaching signals to identify the origin of synthetic content where appropriate and technically feasible to allow it to be identified across platforms.

- Detection: Attempting to detect Deceptive AI Election Content or authenticated content, including with methods such as reading origin signals across different platforms.

    - This might include using detection technology, ingesting open standards-based identifiers created by AI-producing companies or using content moderation services, enabling creators to disclose their use of AI when they upload content, and/or providing pathways for the public to report suspected Deceptive AI Election Content.

- Responsive Protection: Providing swift and proportionate responses to incidents involving the creation and dissemination of Deceptive AI Election Content.

As policymakers consider new regulations for AI, it is important to note that there are already existing rules, regulations, and laws that prohibit unlawful behavior, including those perpetuated through AI. For example, The Food and Drug Administration (FDA) has authority to regulate medical devices and software, including AI-powered medical technologies and applications. The Consumer Financial Protection Bureau (CFPB) has authority to issue subpoenas to investigate any potential "unfair, deceptive, or abusive act" related to a transaction for a "consumer financial product or service," which includes discrimination. The Department of Justice's (DOJ) Office of Civil Rights enforces constitutional provisions and federal statutes prohibiting discrimination across many facets of life, including in education, the criminal justice system, employment, housing, lending, and voting. Any new AI-focused laws or regulations should focus on AI-specific harms that could result from gaps in existing law where there is a high risk to individuals.

Responsible AI Deployment Empowers Americans

Box:

Approximately 90% of the data within an enterprise is unstructured -- it's the sales presentations in a slide deck, finance documents in a spreadsheet, marketing assets in a document, and so on. All of these mission-critical business insights have been locked away within content--until now. With the power of AI, businesses will be able to tap into the wealth of knowledge already securely stored in Box. At Box, we know that content is crucial for businesses, and as the leading cloud content platform for managing unstructured data regardless of productivity suite or business application, we are intent on delivering tremendous value using the power of Box AI. Box AI is a new suite of capabilities that natively integrates advanced AI models into the Box Content Cloud while maintaining high standards for security, compliance, and privacy. Enterprises retain full control over their use of AI. Box AI will drive business insights instantly by allowing users to ask questions about a document, pull insights from a spreadsheet, or summarize a presentation, all with just one click. These tools will also help users create content in seconds by quickly drafting emails, newsletters, or blog posts from the ground up in different tones, lengths, and styles or developing agendas, manuals, and reports that build upon the information already in Box. Finally, Box AI will assist in making business processes more efficient by automating workflows, tasks, and metadata generation to drive faster business outcomes.

Chegg:

With technological advancements, it's crucial to enable students to learn how they want, when they want, and what they want in a personalized way. Chegg's

personalized learning assistant is thoughtfully designed to help students learn with confidence. Chegg breaks down tough concepts into easy-to-learn steps and meets each learner where they are, providing 24/7 on-demand support with personalized learning tools and high-quality learning content. Chegg's personalized learning assistant is powered by AI using 100 million+ pieces of proprietary learning content, includes more than a decade of data-driven insights from the millions of students it serves, and over 150,000 subject matter experts. Chegg has spent over a decade creating a robust, learning taxonomy. As a result, they are able to offer vertical language learning models that are built specifically for helping educational learning.

Google:

In 2021, Google released open sourced AlphaFold, our AI system to predict the 3D structure of a protein just from its 1D amino acid sequence, and created the AlphaFold Protein Structure Database to freely share this scientific knowledge with the world. Proteins are the building blocks of life; they underpin every biological process in every living thing. And, because a protein's shape is closely linked with its function, knowing a protein's structure unlocks a greater understanding of what it does and how it works. Google hoped this groundbreaking resource would help accelerate scientific research and discovery globally, and that other teams could learn from and build on the advances we made with AlphaFold to create further breakthroughs. By demonstrating that AI could accurately predict the shape of a protein down to atomic accuracy, at scale and in minutes, AlphaFold not only provided a solution to a 50-year grand challenge, it also became the first big proof point of our founding thesis: that artificial intelligence can dramatically accelerate scientific discovery, and in turn advance humanity. Just one year after release, AlphaFold had been accessed by more than half a million researchers from 190 counties and used to accelerate progress on important real-world problems ranging from plastic pollution to antibiotic resistance.

Google:

Wildfires affect hundreds of thousands of people each year and are increasing in frequency and size. The need for accurate information when wildfires occur has never been greater. Google has partnered with a number of governments to develop a wildfire tracker that detects wildfire boundaries using new AI models based on satellite imagery to show their real-time location in Search and Maps. The tracker provides updated fire boundary information every 10–15 minutes and incorporates information from local authorities, on Google Search and Google Maps, allowing people to keep safe and stay informed about potential dangers near them, their homes, or loved ones.

HireVue:

HireVue's Find My Fit uses AI to help easily identify what roles best match a candidate's potential. Candidates complete a brief assessment of their interests,

personality, and background. The results are then compared to an organization's open opportunities, recommending the roles that are the best fit for the candidate. By recommending roles based on the candidate's skills and interests, it drives candidates to those roles that might be a better fit— and expands the diversity of a company's talent pool. One of our customers used this specifically to increase the number of female hires.

Intuit:

Intuit Assist is Intuit's generative AI-powered assistant that will provide personalized, intelligent recommendations to help consumer and small business customers make smart financial decisions with less work and complete confidence, enabling them to put more money in their pockets. For example, in Mailchimp, small businesses can enhance their marketing practices by using generative AI to create marketing email content based on industry, intent, and brand voice. With Intuit Assist, a user can select prompts like lengthen, change tone, shorten, correct spelling/grammar, or even ask for things like "rewrite this as a product announcement and make it short," all within our email builder. In addition, in QuickBooks, Intuit Assist will surface personalized insights and recommendations for small businesses based on cash flow. This can help a small business determine whether they can weather a drop in business or buy a new piece of equipment that will help take them to the next level. In each instance, Intuit Assist puts AI in the hands of small businesses to help level the playing field. AI must also be used and deployed responsibly. All Intuit products are built in keeping with our strong commitment to data privacy, security, and responsible AI governance.

Mastercard:

Mastercard's new generative AI model, Decision Intelligence Pro, can help financial institutions improve their fraud detection rates by as much as 300%. The proprietary model is trained on data from the roughly 125 billion transactions that pass through Mastercard's network annually. Instead of focusing on textual inputs, Decision Intelligence Pro uses historical data to improve fraud detection rates by analyzing merchant relationships and predicting fraudulent transactions. The technology operates in real time and can potentially save financial institutions significant costs by eliminating much of the resources they'd typically devote to assessing illegitimate transactions.

Meta:

Meta AI has developed a library of AI models and data to help transform clinical trial eligibility criteria into a machine-readable format. Using this library, trials can be easily searched by their eligibility requirements, making it easier for developers and researchers to build tools that determine trial eligibility. This work will help communities provide better ways for patients from all backgrounds to access clinical trials.

OpenAI and Be My Eyes:

Since 2012, Be My Eyes has been creating technology for the community of over 250 million people who are blind or have low vision. The Danish startup connects people who are blind or have low vision with volunteers for help with hundreds of daily life tasks like identifying a product or navigating an airport. Be My Eyes recently teamed up with OpenAI to develop Be My AI, the first-ever digital visual assistant. With the power of OpenAI's GPT-4 language model, the Be My Eyes app is now able to generate the same level of context and understanding as a human volunteer. This new technology will have profound implications for global accessibility, providing the blind and low vision community with new and powerful tools and capabilities for a host of visual interpretation needs that will introduce a greater degree of independence in their lives.

Pindrop:

In a groundbreaking development during the 2024 US election cycle, a robocall imitating President Joe Biden was circulated. Several news outlets arrived at the right conclusion that this was an AI-generated audio deepfake that targeted multiple individuals across several U.S. states. However, many mentioned how hard it is to identify the text-to-speech (TTS) engine used ("It's nearly impossible to pin down which AI program would have created the audio" – NBC News). This is the challenge Pindrop focused on, and their deep fake analysis was able to identify the TTS engine used as a company headquartered in New York. Pindrop's deepfake engine analyzed the 39-second audio clip through a four-stage process: audio filtering & cleansing, feature extraction, breaking the audio into 155 segments of 250 milliseconds each, and continuously scoring all of the 155 segments of the audio. The 2024 Joe Biden deepfake robocall incident emphasizes the urgency of distinguishing real from AI-generated voices and the importance of Pindrop's research and tools.

Scale AI:

Scale AI, the leading test and evaluation (T&E) partner for frontier artificial intelligence companies, is proud to share that we are partnering with the U.S. Department of Defense's (DoD) Chief Digital and Artificial Intelligence Office (CDAO) to create a comprehensive T&E framework for the responsible use of large language models (LLMs) within the DoD.

Through this partnership, Scale will develop benchmark tests tailored to DoD use cases, integrate them into Scale's T&E platform, and support CDAO's T&E strategy for using LLMs. The outcomes will provide the CDAO a framework to deploy AI safely by measuring model performance, offering real-time feedback for warfighters, and creating specialized public sector evaluation sets to test AI models for military support applications, such as organizing the findings from after action reports.

This work will enable the DoD to mature its T&E policies to address generative AI by measuring and assessing quantitative data via benchmarking and assessing qualitative feedback from users. The evaluation metrics will help identify generative AI models that are ready to support military applications with accurate and relevant results using DoD terminology and knowledge bases.

Scale AI and Orchard Robotics:

Fruit farmers lose billions of dollars annually because they lack the data needed to manage crops precisely. Scale AI and Orchard Robotics developed an AI-driven precision crop management system that helps farmers prevent these losses by optimizing inputs like fertilizer, pesticides, and thinners for each tree to achieve maximum production. Using tractor-mounted, AI-powered camera systems, Orchard Robotics collects precision data about every tree single tree in an orchard and uses an AI model to extract insights from terabytes of image data on the Orchard OS software platform, allowing farmers to act on this data directly by integrating it with existing farm operations.

Zoom:

Zoom believes AI's true potential is allowing people to focus on what matters most: connections, collaboration, and communication. Zoom has harnessed the power of AI for years – background noise suppression, avatars, and virtual backgrounds are among the immersive experiences users already enjoy. Zoom's AI Companion is our generative AI assistant, which works across the Zoom platform. It helps individuals be more productive, connect and collaborate with teammates, and improve skills. Zoom's unique federated approach to generative AI is designed to deliver high-quality results by dynamically incorporating Zoom's large language model (LLM) in addition to third-party large language models, such as OpenAI. This allows Zoom to incorporate innovations in LLMs while getting the benefits of improved quality and performance. Zoom has an unwavering commitment to developing and deploying AI responsibly. We design software with user security, safety, and trust at the heart.

CO Senate Bill 24-205 Comments

Logan Cerkovnik
CEO / Founder of Thumper AI
logan@thumper.ai
04/24/2024

Subject: Comments and Feedback on Colorado Bill 24-205

Hi, my name is Logan Cerkovnik, founder and CEO of Thumper AI, a local generative AI company and one of the few Colorado companies to release an open source foundation model this year.

I founded Thumper so artists would have a way to compete with generative AI models by fine tuning open source models on their own work and allowing them to sell their models to their audience.

Today's bill would stop artist's and anyone other than large companies from finetuning their own models in Colorado and ban our company's platform while benefiting the very companies who unfairly exploited artists initially by scraping their sites without permission.

Furthermore, this bill is a de facto ban on releasing open source models and the open source model ecosystem for Colorado.

This bill would force all the generative AI companies and most corporate IT groups to leave Colorado while failing to stop algorithmic discrimination due to loopholes.

I have 4 points to make here with more details in my written response :

1) This isn't a Colorado bill it's a copypasted version of the Connecticut bill authored by Connecticut State Senator Maroney who has recognized how flawed parts of his bill are and already admitted that at minimum the bill needs major changes if not another legislative session to fix the bill there. If the original author of this legislation doesn't fully support this than nobody in Colorado should either

2) Our legislature should work with Colorado experts and leaders to draft legislation that would regulate AI based on application area, risk level, and type of content generated for generative AI. We need precision regulations that won't be misapplied to

unrelated industries and domains and make sure that the costs of these regulations don't kill our AI startup ecosystem or our open source AI model ecosystem.

3) This bill has too many loopholes due to vague and unnecessary exemptions to discrimination practices such as procedural, or review based exemptions. For example, If you were hoping for a bill that would definitely ban the use of discriminatory review of resumes by AI systems this isn't it. Someone could easily argue that due to their HR process they are considered exempt in most circumstances.

4) Everyone should have the legal right to use, fine tune, and release open source generative AI models. Any bill such as this that would incidentally ban the consumer use of open source AI through regulation is inherently anti-consumer. The future of AI in Colorado is too important to be banned by poorly drafted regulations and we ask that you give Colorado time to create its own bill to bring up in the next legislative session.

There are a large number of concerns with the bill. I will separate these concerns into two areas: high risk AI / Discrimination and Generative AI and believe that these topics should each receive their own bill due to different concerns for each area.

Generative AI:

Open Source Generative AI is the Digital Infrastructure of the Future- Generative AI is the most important technological development since the internet. In ways similar to how internet access was monopolized by telecommunications companies who lobbied states to ban municipal internet at state level throughout the country, today large closed source AI companies are lobbying to ban open source AI models that compete with their closed source products. We must ensure that the open source AI foundation model system is not destroyed by either intentional regulatory capture or poorly written laws that don't understand and appreciate the treasure of our open source software and open source generative AI model ecosystem. Fine Tuning allows every artist, company, and individual in Colorado to benefit from being able to personalize their own model. The only way for smaller companies and artists to compete against big Closed AI companies is to finetune models on their own work and domain. To ban open source generative AI is to ban Colorado companies from being able to compete using AI in their own industry.

Flawed Definitions of AI Developers - The AI developer definition would inherently include many different business and consumer users of AI applications. In particular, the failure to distinguish between generative ai model fine tuning models means that any application that includes fine tuning such as apps like Lensa or loratrainer.com that are intended for users to fine tune models on themselves or their work would be banned because the users would be deemed AI developers. The developer definition should be limited to only companies who have trained a model from scratch. The developer piece of the Connecticut bill has been removed by the original author due to the massive amount of problems that it creates. In general, the model developer should be given safe harbor from how a 3rd party deployer may use or fine-tune their model.

Flawed Definition of AI Deployer - The AI Deployer definition is also similarly flawed in that it also can encompass individual application users as well at times especially for local run ai applications. The AI deployer should be limited to either large public deployments of AI services or private deployments done inside of large corporations and should exempt individual users running applications running AIi models on their laptops or desktops.

Threshold for Exempting Small Businesses and Startups There are several ways to exempt small businesses and startups from AI regulations. It would be reasonable to consider exempting companies with annual global revenue less than 25M USD or 50 employees from generative AI regulations.

Threshold for Exempting Smaller Generative AI Models The White House Executive Order 10^26 compute threshold is flawed because it doesn't adapt to changes in improvements to GPU performance. It would be more fair to exempt regulation of generative AI training models from scratch based on the dollar spent on training from scratch and put more filing requirements on companies spending more than 10M USD on training a particular model than to set an arbitrary flop number or benchmark number that can become outdated fairly quickly.

Failure to Distinguish Between Fine-Tuning and Training from Scratch - There is a huge difference in the investment required to train a foundation model versus fine tuning a model. Foundation model training from scratch costs anywhere from 40,000 USD (PixArt Sigma Text to Image Model) to 300M USD (Llama 3). Fine-tuning models can be done either on local machines with consumer GPUS or be done relatively cheaply ($5 - $10,000 USD) in comparison. Most foundation models are trained on extremely large public or synthetic datasets. Fine-tuning is generally done to focus a foundation model on a specific domain or task and fine tuning is often done on someone's own internal work or data. Only researchers, AI companies, and large tech companies are training foundation models today. Artists, consumers at large, and companies across every industry are fine tuning the open source foundation models that have been released. Due to the vast differences in training costs, data sources, and applications, it makes absolutely no sense to regulate fine-tuning and foundation models equivalently.

Presumption of AI Research Being Harmful - the bill would ban AI research outside of IRB's at large universities and kill startup innovation with its research ban. IRB's are usually considered necessary only for work that is likely to medically or physiologically harm humans or animals in some way. Most AI research benefits humans and poses little to no risk outside of specific areas like Medicine. It is wrong to make the presumption that all AI research is harmful and deserves the same scrutiny as medical research studies on humans or animals.

Lack of Investment in Public AI Infrastructure - California Frontier AI Bill and Connecticut AI bills both create a public computing cluster for public research and educational purposes. Colorado should consider doing the same.

Forfeiture of AI Company Intellectual Property- It is not fair to force AI companies to turn over intellectual property around their AI models in order to meet regulatory requirements. Thumper open sourced its foundation model that was trained on creative commons licensed

data and its training code, but many companies do not do this. In fact for many companies who open source models the piece they do not release is even that much more important to them. It is important that companies not be forced to forfeit their intellectual property to train or deploy models in Colorado.

AI Content Generation Labeling - AI content will soon outnumber authentic content on the internet. Text-based AI content is impossible to label effectively especially if used as a part of text completion or hybrid content that is a mixture of human and AI generated. It is impossible to label AI generated images in a way that cannot be trivially removed by consumers often through operations as easy as taking a screenshot or resizing an image. The future of AI content labeling will have to focus on labeling and protecting the small amount of content that is not authentic content instead. We would ask that the legislature consider choosing a system for labeling authentic content and create penalties for misrepresenting ai generated content as human generated content instead to address this issue. We need to start training society that they cannot trust video and images from unreliable sources so they will not be fooled by ai generated misinformation in the future.

Agreeing to Follow US Copyright Law - Most AI companies today genuinely believe that they have a Fair Use Exemption to Copyright for training AI models on content that is publicly available on the internet. Most AI copyright lawsuits against generative AI companies such as Sarah Silverman's lawsuit have failed because copyright holders are unable to demonstrate the model creates new works that are substantially similar to the original work. The cases where substantially similar works are created generally require some piece of information from the original work to generate results substantially similar to the original work. Congress is the only one who can update the DMCA Act and change the definition of Fair Use with copyright. We recommend that state legislature's consider non-copyright based approaches to protect artists and other copyright holders from having their content scraped from their sites without their consent. We believe that it would be more effective right now to mandate that AI companies training generative AI models be forced to respect "do not train" headers that are present on websites and face a penalty for ignoring "do not train" headers.

Reporting Copyrighted Works Used During Training - Reporting copyrighted works used during training is unlikely to be a very useful regulation because:

- it will encourage many lawsuits by copyright holders that are unlikely to prevail due to the fair use exemption.
- The failed copyright lawsuits are likely to disproportionately hurt small companies and benefit the large companies capable of defending those lawsuits
- Generative AI model training is shifting to synthetic datasets for training models to improve model performance. Phi3, llama3 and Pixart Sigma are examples of this.
- Generative AI Training Datasets are often very large, sometimes even reaching the petabyte range and even providing the links to content can result in files that are in the terabyte range. The state is not really equipped to store or receive that data for a large number of companies. Storing the popular LAON-4B image dataset can cost around 15-20K per month in storage fees on AWS per model. If startups have to pay for the state to store a copy of all the generative AI training data for even a 10 year period of time you would prevent startups from being able to train foundation models. If the reporting is just links to where data was scraped from sites then the data is

obviously smaller, but it also begs the question of who is really going to use this huge regulatory dataset. This dataset would also be very difficult to offer a public search functionality due to the cost of maintaining a search service on that data especially for a reverse image search.

- If full datasets should be submitted companies should be allowed to submit them on hard drives that are stored offline in order to reduce regulatory costs for companies
- It may be better to have companies simply cite the name of large public datasets that they used subsets of rather than reporting entire datasets or dataset links

Safe Harbor for Open Source Generative AI models - Most AI generative AI model developers who release open source models should not be liable for how unrelated 3rd parties use these models provided the original model developers had a legitimate fair use exemption or license to all copyrighted materials used in training. Open source generative AI models should at least be exempt from any new reporting requirements after they are released and the only reporting requirements for an open source generative AI model should be ones that can be done quickly in an automated fashion at the time of release.. Model developers should be able to release a model as being some AI equivalent of for research use only or for general use only without facing legal liability for them.

Safe Harbor for Generative AI Model Fine Tuning - Artists, companies, and consumers who fine tune generative AI models on their own work should have some liability protections that shield them from liabilities related to their use of a particular open source foundation model especially if the fine tune model reduces the general purpose functionality of the original foundation model significantly.

Presumption that general purpose LLMs will discriminate against protected groups It would be worth considering to create a legal presumption that general purpose LLMs will discriminate against protected groups and that the burden of proving that they do not should fall to AI application developers in high risk areas. The state should seek to strongly communicate that developers need to exercise extreme caution with general purpose LLMs and consider avoiding them for high risk applications without the use of fine-tuning and further study on a case by case basis. Additionally, the state should make it clear to everyone that no one is allowed to use general purpose LLMs in high risk areas without demonstrating clear and convincing evidence that they are not discriminatory before deploying these applications.

Consider different regulations for Generative AI by Category of Content Generated Image generating models have very different concerns from LLMs and text generating models. If we try to regulate all forms of generative AI the same then we will end up with regulations that don't make sense being applied to domains that they shouldn't. llm/s and text based models do not share the same deep fake concerns present in image, video, and audio models. Image, video, and audio models also are much less likely to be misused to cause cybersecurity harm like LLMs can. We need generative AI legislation to be tailored appropriately to the application and domain for it to be effective.

Generative AI is only high risk if applied to a high risk areas
Generative AI is mostly used for extremely low risk applications and generative AI as a whole should not be elevated to a high risk status based on the class of model. Artists should not have their text to image models regulated the same way as LLMs.

High Risk / Discriminatory AI

AI Developers and Companies seek to remove algorithmic discrimination - Almost everyone building high risk models AI alreadying intends to make sure that algorithmic discrimination is removed from high risk models. The main points of contention algorithmic discrimination are likely to be around how companies should deal with missing data for protected classes, how DEI programs should be protected, what liability standards and shields should be used for companies, should there be a Colorado AI regulatory commission, is there a reasonable compliance cost for algorithmic AI programs, will companies be incentivized to remove high risk algorithm costs due to compliance costs or legal liability concerns , how should consumers of high risk AI models be notified of or challenge a high risk algorithmic decision, and should model explainability be a hard requirement of high models.

Giant loopholes in this bill exempt too many forms of algorithmic discrimination 3 (b) "HIGH-RISK ARTIFICIAL INTELLIGENCE SYSTEM" DOES NOT 4 INCLUDE AN ARTIFICIAL INTELLIGENCE SYSTEM, AS DEFINED IN 5 SUBSECTION (2) OF THIS SECTION, IF THE ARTIFICIAL INTELLIGENCE 6 SYSTEM IS INTENDED TO:
7 (I) PERFORM A NARROW PROCEDURAL TASK;
8 (II) IMPROVE THE RESULT OF A PREVIOUSLY COMPLETED HUMAN 9 ACTIVITY;
10 (III) DETECT DECISION-MAKING PATTERNS OR DEVIATIONS FROM 11 PRIOR DECISION-MAKING PATTERNS AND IS NOT INTENDED TO REPLACE OR 12 INFLUENCE A PREVIOUSLY COMPLETED HUMAN ASSESSMENT WITHOUT 13 SUFFICIENT HUMAN REVIEW; OR
14 (IV) PERFORM A PREPARATORY TASK TO AN ASSESSMENT THAT IS 15 RELEVANT TO A CONSEQUENTIAL DECISION. “

1. A narrow procedural task is very poorly defined and could be applied to many different application processes. For instance consider an application that uses aI to summarize a resume for a job application pool but does so in a discriminatory way. This intermediate result can be used by HR or a hiring manager to determine what applicants to interview rather than reading whole resumes. Because this was arguably a narrow procedural task this may not be considered a high risk area and not be covered by this bill.
2. “Improve the result of previously completed human activity”, an llm model can still amplify bias present in a human activity in a way that increases the harmful behavior or hallucinate harmful output from human input that may not readily appear problematic
3. “Detect Decison-making Patterns or Deviations from prior decision making patterns and is not intended to replace or influence a previously completed human assessment without sufficient human review - again this also creates problems by assuming that the human review in many semi-automated applications will

4. “Perform a preparatory task to an assessment that is relevant to a consequential decision” - Again if this preparatory task is performing something in a biased way that impacts downstream decision making that is still a problem. Imagine a credit or risk scoring system that tried to argue the risk scoring system was just a preparatory task and that a loan officer made decisions based on other components as well. This discriminatory model behavior should have been addressed because it's a model that impacts granting a loan even though it's not the singular decision factor.

If most AI software applications have multiple steps then allowing intermediate steps to be exempted from being regulated for discriminatory impact will result allow almost all applications to be unregulated or intentionally designed to sandwich AI processes between steps

This bill expects the attorney general and district attorney to be the regulators of AI
Other bills such as the frontier bill in California have proposed creating an AI commission to help regulate AI companies. AI companies and consumers would be better served by an AI commission or separate regulatory division of AI experts to take the majority of the responsibility for regulating AI in Colorado. The vague language of this bill coupled with the lack of a regulatory body sets up a future conflict between AI companies and regulators who may have very different interpretations of this bill on many different levels. AI is moving at an extremely rapid pace right now and a Colorado AI commission would be able to better respond more flexibly to any new AI developments. We advise that future legislation strongly consider creating a group of AI experts in some way to make AI regulation more effective and nimble.

Algorithmic Diversity programs could undermine trust without human oversight - The bill deliberately exempts algorithmic diversity efforts from being considered algorithmic discrimination. It is unclear whether this is intended to simply exempt currently existing human DEI programs or allow a future class of algorithmic DEI programs. A better approach might be to exempt diversity programs that are run by humans from the algorithmic discrimination definition instead. Removing human oversight from diversity programs could undermine the public’s confidence in these important programs. Diversity programs have come under increasing criticism from some political leaders over the last few years and this criticism has unfairly hurt the public perception of these programs. It is worth considering requiring human oversight of any algorithm diversity programs for some period of time until we can ensure that they can function as they are intended and do not have adverse impacts on the disadvantaged groups they are intended to help.

Lack of Data on All Protective Classes - Many categories for protected classes are not available to companies building high risk models. The most effective way to train an AI model that doesn’t discriminate against a protected class is to use the projected class to slightly decrease overall model performance to force the model to perform equivantly across all protected classes. However, if you don’t have access to that protected class information it is not possible for an AI model developer to ever guarantee that that model doesn’t discriminate against protected classes that weren’t available to the company. Companies should be provided some sort of safe harbor from claims of AI model discrimination that they are not aware of due to missing or unavailable data. As companies gain access to more data they should be given a reasonable period of time to create and introduce a new model that

balances performance across new protected classes. Companies may well be able to anticipate that a model is likely to discriminate against a particular group but be unable to do anything about it due to privacy concerns or lack of access to protected class data for an application. We believe that companies should only be legally penalized for algorithmic discrimination against protected classes that companies have access to or data features that are historically known to correlate with protected classes such as the relationship between zip code and racial demographics.

Choice of Risk Framework- Allowing companies to choose their own framework and making the NIST framework a default framework is an inherently bad idea because the framework created by NIST was never intended to be used this way and NIST is not a policy or regulatory focused group that is well-situated to understand public policy and economic matters. The regulations should be clear, consistent and minimally sufficient to accomplish the intended goal. Furthermore the bill is vague on what would qualify as being recognized enough to be an acceptable risk framework other than NIST. Instead Colorado needs to create its own AI framework.

Lack of Clarity on Explainability as a Requirement of High Risk Models- the bill mentions explainability of high risk models for reporting requirements, "HOW THE HIGH-RISK ARTIFICIAL INTELLIGENCE SYSTEM WAS EVALUATED FOR PERFORMANCE AND RELEVANT INFORMATION RELATED TO EXPLAINABILITY BEFORE THE HIGH-RISK ARTIFICIAL INTELLIGENCE "but doesn't clarify whether explainability is a requirement. One of the biggest questions facing an AI developer building a high risk model is what type of model features are necessary for the given application. The legislation must clarify whether explainability is a hard requirement, recommended, or not a requirement for AI developers to understand where they stand. If explainability were to be mandated most deep learning and all generative ai models would be banned from use in high risk applications. Most general purpose generative ai models are not fit for use in high risk or discriminative applications.

Unintentional Regulation of Legacy Algorithms - the bill doesn't exempt older models and algorithms used by various companies. A rules-based recommendation system or linear regression model is probably not the intended target of this legislation, but would be impacted. It may well be worth grandfathering in older applications or giving companies a period of time to switch these models or exempt this class of models from regulations given their lower risk profile, easy explainability, and higher transparency than machine learning models.

Overlapping Coverage with Federal Regulations- Medical Device and Healthcare companies already have regulation mechanisms in place that are sufficient for most applications. Colorado should create a limited focus on areas that already have sufficient Federal oversight. Similarly, Colorado experienced what many believe to be the first Coloradan death due to failing self driving car technology in Evergreen in 2022 and yet self driving car algorithms are not included as a high risk algorithm for purposes in this bill.

In summary, this bill is too flawed to be easily salvaged and I would ask that the committee partner with Colorado AI experts and leaders to create a working group to collaborate on

drafting future legislation that will actually protect Coloradans from algorithmic discrimination without destroying the Colorado AI startup and open source ecosystem.

CO Senate Bill 24-205 Comments

Logan Cerkovnik
CEO / Founder of Thumper AI
logan@thumper.ai
04/24/2024

Subject: Comments and Feedback on Colorado Bill 24-205

Hi, my name is Logan Cerkovnik, founder and CEO of Thumper AI, a local generative AI company and one of the few Colorado companies to release an open source foundation model this year.

I founded Thumper so artists would have a way to compete with generative AI models by fine tuning open source models on their own work and allowing them to sell their models to their audience.

Today's bill would stop artist's and anyone other than large companies from finetuning their own models in Colorado and ban our company's platform while benefiting the very companies who unfairly exploited artists initially by scraping their sites without permission.

Furthermore, this bill is a de facto ban on releasing open source models and the open source model ecosystem for Colorado.

This bill would force all the generative AI companies and most corporate IT groups to leave Colorado while failing to stop algorithmic discrimination due to loopholes.

I have 4 points to make here with more details in my written response :

1) This isn't a Colorado bill it's a copypasted version of the Connecticut bill authored by Connecticut State Senator Maroney who has recognized how flawed parts of his bill are and already admitted that at minimum the bill needs major changes if not another legislative session to fix the bill there. If the original author of this legislation doesn't fully support this than nobody in Colorado should either

2) Our legislature should work with Colorado experts and leaders to draft legislation that would regulate AI based on application area, risk level, and type of content generated for generative AI. We need precision regulations that won't be misapplied to

unrelated industries and domains and make sure that the costs of these regulations don't kill our AI startup ecosystem or our open source AI model ecosystem.

3) This bill has too many loopholes due to vague and unnecessary exemptions to discrimination practices such as procedural, or review based exemptions. For example, If you were hoping for a bill that would definitely ban the use of discriminatory review of resumes by AI systems this isn't it. Someone could easily argue that due to their HR process they are considered exempt in most circumstances.

4) Everyone should have the legal right to use, fine tune, and release open source generative AI models. Any bill such as this that would incidentally ban the consumer use of open source AI through regulation is inherently anti-consumer. The future of AI in Colorado is too important to be banned by poorly drafted regulations and we ask that you give Colorado time to create its own bill to bring up in the next legislative session.

There are a large number of concerns with the bill. I will separate these concerns into two areas: high risk AI / Discrimination and Generative AI and believe that these topics should each receive their own bill due to different concerns for each area.

Generative AI:

Open Source Generative AI is the Digital Infrastructure of the Future- Generative AI is the most important technological development since the internet. In ways similar to how internet access was monopolized by telecommunications companies who lobbied states to ban municipal internet at state level throughout the country, today large closed source AI companies are lobbying to ban open source AI models that compete with their closed source products. We must ensure that the open source AI foundation model system is not destroyed by either intentional regulatory capture or poorly written laws that don't understand and appreciate the treasure of our open source software and open source generative AI model ecosystem. Fine Tuning allows every artist, company, and individual in Colorado to benefit from being able to personalize their own model. The only way for smaller companies and artists to compete against big Closed AI companies is to finetune models on their own work and domain. To ban open source generative AI is to ban Colorado companies from being able to compete using AI in their own industry.

Flawed Definitions of AI Developers - The AI developer definition would inherently include many different business and consumer users of AI applications. In particular, the failure to distinguish between generative ai model fine tuning models means that any application that includes fine tuning such as apps like Lensa or loratrainer.com that are intended for users to fine tune models on themselves or their work would be banned because the users would be deemed AI developers. The developer definition should be limited to only companies who have trained a model from scratch. The developer piece of the Connecticut bill has been removed by the original author due to the massive amount of problems that it creates. In general, the model developer should be given safe harbor from how a 3rd party deployer may use or fine-tune their model.

Flawed Definition of AI Deployer - The AI Deployer definition is also similarly flawed in that it also can encompass individual application users as well at times especially for local run ai applications. The AI deployer should be limited to either large public deployments of AI services or private deployments done inside of large corporations and should exempt individual users running applications running AIi models on their laptops or desktops.

Threshold for Exempting Small Businesses and Startups There are several ways to exempt small businesses and startups from AI regulations. It would be reasonable to consider exempting companies with annual global revenue less than 25M USD or 50 employees from generative AI regulations.

Threshold for Exempting Smaller Generative AI Models The White House Executive Order 10^26 compute threshold is flawed because it doesn't adapt to changes in improvements to GPU performance. It would be more fair to exempt regulation of generative AI training models from scratch based on the dollar spent on training from scratch and put more filing requirements on companies spending more than 10M USD on training a particular model than to set an arbitrary flop number or benchmark number that can become outdated fairly quickly.

Failure to Distinguish Between Fine-Tuning and Training from Scratch - There is a huge difference in the investment required to train a foundation model versus fine tuning a model. Foundation model training from scratch costs anywhere from 40,000 USD (PixArt Sigma Text to Image Model) to 300M USD (Llama 3). Fine-tuning models can be done either on local machines with consumer GPUS or be done relatively cheaply ($5 - $10,000 USD) in comparison. Most foundation models are trained on extremely large public or synthetic datasets. Fine-tuning is generally done to focus a foundation model on a specific domain or task and fine tuning is often done on someone's own internal work or data. Only researchers, AI companies, and large tech companies are training foundation models today. Artists, consumers at large, and companies across every industry are fine tuning the open source foundation models that have been released. Due to the vast differences in training costs, data sources, and applications, it makes absolutely no sense to regulate fine-tuning and foundation models equivalently.

Presumption of AI Research Being Harmful - the bill would ban AI research outside of IRB's at large universities and kill startup innovation with its research ban. IRB's are usually considered necessary only for work that is likely to medically or physiologically harm humans or animals in some way. Most AI research benefits humans and poses little to no risk outside of specific areas like Medicine. It is wrong to make the presumption that all AI research is harmful and deserves the same scrutiny as medical research studies on humans or animals.

Lack of Investment in Public AI Infrastructure - California Frontier AI Bill and Connecticut AI bills both create a public computing cluster for public research and educational purposes. Colorado should consider doing the same.

Forfeiture of AI Company Intellectual Property- It is not fair to force AI companies to turn over intellectual property around their AI models in order to meet regulatory requirements. Thumper open sourced its foundation model that was trained on creative commons licensed

data and its training code, but many companies do not do this. In fact for many companies who open source models the piece they do not release is even that much more important to them. It is important that companies not be forced to forfeit their intellectual property to train or deploy models in Colorado.

AI Content Generation Labeling - AI content will soon outnumber authentic content on the internet. Text-based AI content is impossible to label effectively especially if used as a part of text completion or hybrid content that is a mixture of human and AI generated. It is impossible to label AI generated images in a way that cannot be trivially removed by consumers often through operations as easy as taking a screenshot or resizing an image. The future of AI content labeling will have to focus on labeling and protecting the small amount of content that is not authentic content instead. We would ask that the legislature consider choosing a system for labeling authentic content and create penalties for misrepresenting ai generated content as human generated content instead to address this issue. We need to start training society that they cannot trust video and images from unreliable sources so they will not be fooled by ai generated misinformation in the future.

Agreeing to Follow US Copyright Law - Most AI companies today genuinely believe that they have a Fair Use Exemption to Copyright for training AI models on content that is publicly available on the internet. Most AI copyright lawsuits against generative AI companies such as Sarah Silverman's lawsuit have failed because copyright holders are unable to demonstrate the model creates new works that are substantially similar to the original work. The cases where substantially similar works are created generally require some piece of information from the original work to generate results substantially similar to the original work. Congress is the only one who can update the DMCA Act and change the definition of Fair Use with copyright. We recommend that state legislature's consider non-copyright based approaches to protect artists and other copyright holders from having their content scraped from their sites without their consent. We believe that it would be more effective right now to mandate that AI companies training generative AI models be forced to respect "do not train" headers that are present on websites and face a penalty for ignoring "do not train" headers.

Reporting Copyrighted Works Used During Training - Reporting copyrighted works used during training is unlikely to be a very useful regulation because:

- it will encourage many lawsuits by copyright holders that are unlikely to prevail due to the fair use exemption.
- The failed copyright lawsuits are likely to disproportionately hurt small companies and benefit the large companies capable of defending those lawsuits
- Generative AI model training is shifting to synthetic datasets for training models to improve model performance. Phi3, llama3 and Pixart Sigma are examples of this.
- Generative AI Training Datasets are often very large, sometimes even reaching the petabyte range and even providing the links to content can result in files that are in the terabyte range. The state is not really equipped to store or receive that data for a large number of companies. Storing the popular LAON-4B image dataset can cost around 15-20K per month in storage fees on AWS per model. If startups have to pay for the state to store a copy of all the generative AI training data for even a 10 year period of time you would prevent startups from being able to train foundation models. If the reporting is just links to where data was scraped from sites then the data is

obviously smaller, but it also begs the question of who is really going to use this huge regulatory dataset. This dataset would also be very difficult to offer a public search functionality due to the cost of maintaining a search service on that data especially for a reverse image search.

- If full datasets should be submitted companies should be allowed to submit them on hard drives that are stored offline in order to reduce regulatory costs for companies
- It may be better to have companies simply cite the name of large public datasets that they used subsets of rather than reporting entire datasets or dataset links

Safe Harbor for Open Source Generative AI models - Most AI generative AI model developers who release open source models should not be liable for how unrelated 3rd parties use these models provided the original model developers had a legitimate fair use exemption or license to all copyrighted materials used in training. Open source generative AI models should at least be exempt from any new reporting requirements after they are released and the only reporting requirements for an open source generative AI model should be ones that can be done quickly in an automated fashion at the time of release.. Model developers should be able to release a model as being some AI equivalent of for research use only or for general use only without facing legal liability for them.

Safe Harbor for Generative AI Model Fine Tuning - Artists, companies, and consumers who fine tune generative AI models on their own work should have some liability protections that shield them from liabilities related to their use of a particular open source foundation model especially if the fine tune model reduces the general purpose functionality of the original foundation model significantly.

Presumption that general purpose LLMs will discriminate against protected groups It would be worth considering to create a legal presumption that general purpose LLMs will discriminate against protected groups and that the burden of proving that they do not should fall to AI application developers in high risk areas. The state should seek to strongly communicate that developers need to exercise extreme caution with general purpose LLMs and consider avoiding them for high risk applications without the use of fine-tuning and further study on a case by case basis. Additionally, the state should make it clear to everyone that no one is allowed to use general purpose LLMs in high risk areas without demonstrating clear and convincing evidence that they are not discriminatory before deploying these applications.

Consider different regulations for Generative AI by Category of Content Generated Image generating models have very different concerns from LLMs and text generating models. If we try to regulate all forms of generative AI the same then we will end up with regulations that don't make sense being applied to domains that they shouldn't. llm/s and text based models do not share the same deep fake concerns present in image, video, and audio models. Image, video, and audio models also are much less likely to be misused to cause cybersecurity harm like LLMs can. We need generative AI legislation to be tailored appropriately to the application and domain for it to be effective.

Generative AI is only high risk if applied to a high risk areas
Generative AI is mostly used for extremely low risk applications and generative AI as a whole should not be elevated to a high risk status based on the class of model. Artists should not have their text to image models regulated the same way as LLMs.

High Risk / Discriminatory AI

AI Developers and Companies seek to remove algorithmic discrimination - Almost everyone building high risk models AI alreadying intends to make sure that algorithmic discrimination is removed from high risk models. The main points of contention algorithmic discrimination are likely to be around how companies should deal with missing data for protected classes, how DEI programs should be protected, what liability standards and shields should be used for companies, should there be a Colorado AI regulatory commission, is there a reasonable compliance cost for algorithmic AI programs, will companies be incentivized to remove high risk algorithm costs due to compliance costs or legal liability concerns , how should consumers of high risk AI models be notified of or challenge a high risk algorithmic decision, and should model explainability be a hard requirement of high models.

Giant loopholes in this bill exempt too many forms of algorithmic discrimination 3 (b) "HIGH-RISK ARTIFICIAL INTELLIGENCE SYSTEM" DOES NOT 4 INCLUDE AN ARTIFICIAL INTELLIGENCE SYSTEM, AS DEFINED IN 5 SUBSECTION (2) OF THIS SECTION, IF THE ARTIFICIAL INTELLIGENCE 6 SYSTEM IS INTENDED TO:
7 (I) PERFORM A NARROW PROCEDURAL TASK;
8 (II) IMPROVE THE RESULT OF A PREVIOUSLY COMPLETED HUMAN 9 ACTIVITY;
10 (III) DETECT DECISION-MAKING PATTERNS OR DEVIATIONS FROM 11 PRIOR DECISION-MAKING PATTERNS AND IS NOT INTENDED TO REPLACE OR 12 INFLUENCE A PREVIOUSLY COMPLETED HUMAN ASSESSMENT WITHOUT 13 SUFFICIENT HUMAN REVIEW; OR
14 (IV) PERFORM A PREPARATORY TASK TO AN ASSESSMENT THAT IS 15 RELEVANT TO A CONSEQUENTIAL DECISION. “

1. A narrow procedural task is very poorly defined and could be applied to many different application processes. For instance consider an application that uses aI to summarize a resume for a job application pool but does so in a discriminatory way. This intermediate result can be used by HR or a hiring manager to determine what applicants to interview rather than reading whole resumes. Because this was arguably a narrow procedural task this may not be considered a high risk area and not be covered by this bill.
2. “Improve the result of previously completed human activity”, an llm model can still amplify bias present in a human activity in a way that increases the harmful behavior or hallucinate harmful output from human input that may not readily appear problematic
3. “Detect Decison-making Patterns or Deviations from prior decision making patterns and is not intended to replace or influence a previously completed human assessment without sufficient human review - again this also creates problems by assuming that the human review in many semi-automated applications will

4. “Perform a preparatory task to an assessment that is relevant to a consequential decision” - Again if this preparatory task is performing something in a biased way that impacts downstream decision making that is still a problem. Imagine a credit or risk scoring system that tried to argue the risk scoring system was just a preparatory task and that a loan officer made decisions based on other components as well. This discriminatory model behavior should have been addressed because it's a model that impacts granting a loan even though it's not the singular decision factor.

If most AI software applications have multiple steps then allowing intermediate steps to be exempted from being regulated for discriminatory impact will result allow almost all applications to be unregulated or intentionally designed to sandwich AI processes between steps

This bill expects the attorney general and district attorney to be the regulators of AI
Other bills such as the frontier bill in California have proposed creating an AI commission to help regulate AI companies. AI companies and consumers would be better served by an AI commission or separate regulatory division of AI experts to take the majority of the responsibility for regulating AI in Colorado. The vague language of this bill coupled with the lack of a regulatory body sets up a future conflict between AI companies and regulators who may have very different interpretations of this bill on many different levels. AI is moving at an extremely rapid pace right now and a Colorado AI commission would be able to better respond more flexibly to any new AI developments. We advise that future legislation strongly consider creating a group of AI experts in some way to make AI regulation more effective and nimble.

Algorithmic Diversity programs could undermine trust without human oversight - The bill deliberately exempts algorithmic diversity efforts from being considered algorithmic discrimination. It is unclear whether this is intended to simply exempt currently existing human DEI programs or allow a future class of algorithmic DEI programs. A better approach might be to exempt diversity programs that are run by humans from the algorithmic discrimination definition instead. Removing human oversight from diversity programs could undermine the public’s confidence in these important programs. Diversity programs have come under increasing criticism from some political leaders over the last few years and this criticism has unfairly hurt the public perception of these programs. It is worth considering requiring human oversight of any algorithm diversity programs for some period of time until we can ensure that they can function as they are intended and do not have adverse impacts on the disadvantaged groups they are intended to help.

Lack of Data on All Protective Classes - Many categories for protected classes are not available to companies building high risk models. The most effective way to train an AI model that doesn’t discriminate against a protected class is to use the projected class to slightly decrease overall model performance to force the model to perform equivantly across all protected classes. However, if you don’t have access to that protected class information it is not possible for an AI model developer to ever guarantee that that model doesn’t discriminate against protected classes that weren’t available to the company. Companies should be provided some sort of safe harbor from claims of AI model discrimination that they are not aware of due to missing or unavailable data. As companies gain access to more data they should be given a reasonable period of time to create and introduce a new model that

balances performance across new protected classes. Companies may well be able to anticipate that a model is likely to discriminate against a particular group but be unable to do anything about it due to privacy concerns or lack of access to protected class data for an application. We believe that companies should only be legally penalized for algorithmic discrimination against protected classes that companies have access to or data features that are historically known to correlate with protected classes such as the relationship between zip code and racial demographics.

Choice of Risk Framework- Allowing companies to choose their own framework and making the NIST framework a default framework is an inherently bad idea because the framework created by NIST was never intended to be used this way and NIST is not a policy or regulatory focused group that is well-situated to understand public policy and economic matters. The regulations should be clear, consistent and minimally sufficient to accomplish the intended goal. Furthermore the bill is vague on what would qualify as being recognized enough to be an acceptable risk framework other than NIST. Instead Colorado needs to create its own AI framework.

Lack of Clarity on Explainability as a Requirement of High Risk Models- the bill mentions explainability of high risk models for reporting requirements, "HOW THE HIGH-RISK ARTIFICIAL INTELLIGENCE SYSTEM WAS EVALUATED FOR PERFORMANCE AND RELEVANT INFORMATION RELATED TO EXPLAINABILITY BEFORE THE HIGH-RISK ARTIFICIAL INTELLIGENCE "but doesn't clarify whether explainability is a requirement. One of the biggest questions facing an AI developer building a high risk model is what type of model features are necessary for the given application. The legislation must clarify whether explainability is a hard requirement, recommended, or not a requirement for AI developers to understand where they stand. If explainability were to be mandated most deep learning and all generative ai models would be banned from use in high risk applications. Most general purpose generative ai models are not fit for use in high risk or discriminative applications.

Unintentional Regulation of Legacy Algorithms - the bill doesn't exempt older models and algorithms used by various companies. A rules-based recommendation system or linear regression model is probably not the intended target of this legislation, but would be impacted. It may well be worth grandfathering in older applications or giving companies a period of time to switch these models or exempt this class of models from regulations given their lower risk profile, easy explainability, and higher transparency than machine learning models.

Overlapping Coverage with Federal Regulations- Medical Device and Healthcare companies already have regulation mechanisms in place that are sufficient for most applications. Colorado should create a limited focus on areas that already have sufficient Federal oversight. Similarly, Colorado experienced what many believe to be the first Coloradan death due to failing self driving car technology in Evergreen in 2022 and yet self driving car algorithms are not included as a high risk algorithm for purposes in this bill.

In summary, this bill is too flawed to be easily salvaged and I would ask that the committee partner with Colorado AI experts and leaders to create a working group to collaborate on

drafting future legislation that will actually protect Coloradans from algorithmic discrimination without destroying the Colorado AI startup and open source ecosystem.

TAXPAYERS PROTECTION ALLIANCE

April 23, 2024

Senate Judiciary Committee
Colorado General Assembly
200 E Colfax Avenue
Denver, CO 80203

Dear Chair Gonzales, Vice Chair Roberts, and Members of the Committee,

On behalf of the millions of taxpayers and consumers we represent (including many in the state of Colorado), the Taxpayers Protection Alliance (TPA) would like to express its opposition against bill SB24-205. While well-intentioned, the bill (as currently drafted) would significantly hamper the development of the up-and-coming artificial intelligence (AI) industry due to the onerous costs it would introduce.

Novel technologies, such as AI, bring valuable market disruption, technological progress, and growth to the American economy. Oftentimes, the emergence of these new technologies spark fears of potential harm that could stem from their widespread use. Such is the case with AI technology. While this bill attempts to tackle valid concerns, its approach would install mandates of dubious effectiveness at the expense of technological progress.

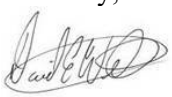
In an industry that is mostly driven by start-ups and small businesses (like AI), the compliance costs that would emanate from this bill could prove to be extremely prohibitive. For resource-constrained companies such as a startup, complying with the bill's requirements, such as annual impact assessments, would significantly slow their growth and profitability. If this bill is enacted small businesses will often have to divert resources from productive uses to comply with the bill's numerous requirements.

Colorado and the United States cannot afford to be fall behind in the race for leadership in AI. In a highly contested market with various foreign adversaries investing millions in hopes of contest the U.S.'s leadership in the space, this bill could enact unsurmountable barriers for the domestic AI industry.

Before enacting new regulations, legislators should evaluate existing regulations and adapt them if needed. Most of the sanctionable behaviors that the bill attempts to curtail is usually already regulated by a wide corpus of antidiscrimination laws. Thus, passing this bill will result in overlapping and redundant regulations that are likely to yield the same results, but with one of them adding significant costs in the process.

TPA commends the Senate's efforts to listen and inform themselves on the matter.

Sincerely,



David Williams
President