COMMENT

MODEL ACT FOR ALGORITHMIC MODELS: A REGULATORY SOLUTION FOR AI USED IN HIRING DECISIONS*

TABLE OF CONTENTS

I. INTRODUCTION 51

II. VIDEO INTERVIEW AI AND REGULATION LANDSCAPE 53
- *A. Video Interview AI Generally* 53
- *B. Video Interview AI Regulation Landscape* 54
  1. Illinois Artificial Intelligence Video Interview Act. 54
  2. New York City's Automated Employment Decision Tool Law (AEDT). 55
  3. *EEOC Guidance* 56
  4. OSTP's Blueprint for an AI Bill of Rights 57
  5. *AI Training Act of 2022* 58

III. THE DEMAND FOR EXPLAINABILITY 58
- *A. Explainable AI (XAI) Generally* 58
- *B. Different Forms of Explainability and Efforts by Private Actors* 60
  1. *Global v. Local Explainability.* 60
  2. *Model Cards* 60
  3. *Explainability Statements.* 63

IV. THE CASE AGAINST FEDERAL LEGISLATION ON AI USED IN EMPLOYMENT DECISIONS 64

* University of Houston Law Center, J.D. Candidate, 2023, and Senior HLR*e* Editor of the *Houston Law Review*.

Thank you to Professor Seth Chandler for his guidance on this Article and for sparking my interest in the intersection of AI and law.

A. The EEOC's Lack of Enforcement and the AAA's Overreach 64
B. The AAA's Flaws Show that Congress is Not Equipped to Regulate AI 65

V. AI SHOULD NOT BE HELD TO A HIGHER STANDARD THAN HUMANS 68

VI. A SOLUTION TO REGULATING AI TOOLS IN HIRING: A MODEL ACT 72
A. Why Have a Model Act 72
B. What Should be in it: The Model Act Should Move Away from an Explainability Requirement 76
C. What Should be in it: The Model Act Should be Limited in Scope 79
D. What Should be in it: The Model Act Should Contain Enforcement Procedures or Cause of Action 81

VII. CONCLUSION 81

I. INTRODUCTION

Over the past several years, concerns surrounding Artificial Intelligence (AI) have gone mainstream.[1] Many such concerns stem from the potential misuse or unintended consequences of AI, particularly with critical decisions like health outcomes or job

1. *See* Bernard Marr, *Is Artificial Intelligence Dangerous? 6 AI Risks Everyone Should Know About*, FORBES (Nov. 19, 2018, 12:20 AM), https://www.forbes.com/sites/bernardmarr/2018/11/19/is-artificial-intelligence-dangerous-6-ai-risks-everyone-should-know-about/?sh=78a49b882404; *see also* Melissa Heikkilä, *Who's going to save us from bad AI?*, MIT TECH. REVIEW (Oct. 10, 2022), https://www.technologyreview.com/2022/10/10/1061016/whos-going-to-save-us-from-bad-ai/; *see also* Lee Rainie et al., *How Americans Think About Artificial Intelligence*, PEW RESEARCH CENTER (March 17, 2022), https://www.pewresearch.org/internet/2022/03/17/how-americans-think-about-artificial-intelligence/; *see also* Reid Blackman & Beena Ammanath, *Ethics and AI: 3 Conversations Companies Need to Have*, HARVARD BUS. REV. (March 21, 2022), https://hbr.org/2022/03/ethics-and-ai-3-conversations-companies-need-to-be-having ("No one wants to push out discriminatory or biased AI. No one wants to be the object of a lawsuit or regulatory investigation for violations of privacy. But once we've all agreed that biased, black box, privacy-violating AI is bad, where do we go from here? The question most every senior leader asks is: How do we take action to mitigate those ethical risks?"); *see also* Margarita Simonova, *Top Nine Ethical Issues In Artificial Intelligence*, FORBES (Oct. 11, 2022, 8:30 AM), https://www.forbes.com/sites/forbestechcouncil/2022/10/11/top-nine-ethical-issues-in-artificial-intelligence/?sh=755df72d5bc8.

opportunities.[2] To combat such concerns, the concept of AI Ethics or "AI Accountability" has emerged and spread throughout the industry, led by AI developers and advocacy organizations.[3] State legislatures are also proposing and passing state-specific bills addressing the use of AI. Thus far, many of these laws have focused on employment and hiring decisions.[4] On the national stage just in 2022, the U.S. Department of Commerce formed the National Artificial Intelligence Advisory Committee, Congress proposed the Algorithmic Accountability Act (AAA), and the White House's Office of Science and Technology Policy released a Blueprint for AI Bill of Rights.[5] These measures presage a federal solution for "AI Accountability."

In this paper, I (1) discuss current regulations surrounding AI tools used in employment and hiring decisions, (2) demonstrate why federal legislation by Congress should not be pursued during the infancy of this industry, and (3) explore current "AI Accountability" solutions and their viability in place of regulations. Towards the end, I propose and recommend an alternative to federal legislation: a Model Act on AI which can be adopted by states as they see fit. Lastly, this paper is not intended to be fully comprehensive, but rather, is intended to be another contribution to the dialogue in this growing field.

---

2. *See* Lee Rainie et al., *Worries about Developments in AI*, PEW RESEARCH CENTER (June 16, 2021), https://www.pewresearch.org/internet/2021/06/16/1-worries-about-developments-in-ai/; *see also* Cynthia Dwork & Martha Louise Minow, *Distrust of Artificial Intelligence: Sources & Responses from Computer Science & Law*, 151 Daedalus 309, 310 (2022) ("Distrust is manifested in growing calls for regulation, the emergence of watchdog and lobbying groups, and the explicit recognition of new risks requiring monitoring by corporate audit committees and accounting firms").

3. *See* John McCormick, *AI Ethics Teams Bulk Up in Size, Influence at Tech Firms*, WALL STREET JOURNAL (May 27, 2021, 7:00 AM), https://www.wsj.com/articles/ai-ethics-teams-bulk-up-in-size-influence-at-tech-firms-11622113202.

4. *See Legislation Related to Artificial Intelligence*, NATIONAL CONFERENCE OF STATE LEGISLATURES (Aug. 26, 2022), https://www.ncsl.org/research/telecommunications-and-information-technology/2020-legislation-related-to-artificial-intelligence.aspx ("General artificial intelligence bills or resolutions were introduced in at least 17 states in 2022, and were enacted in Colorado, Illinois, Vermont and Washington. Colorado, Illinois and Vermont created task forces or commissions to study AI").

5. *See* Algorithmic Accountability Act, H.R. 6580, 117th Cong. (2022); *See Blueprint for an AI Bill of Rights*, OFFICE OF SCIENCE AND TECHNOLOGY POLICY, https://www.whitehouse.gov/ostp/ai-bill-of-rights/ (last visited Nov. 23, 2022).

II. VIDEO INTERVIEW AI AND REGULATION LANDSCAPE

A. Video Interview AI Generally

Companies specializing in AI tools in hiring have reported a surge in demand and business in the past few years.[6] Leading recruiting software developers like HireVue, MyInterview, Interview.AI, and Curious Thing use AI in various steps of recruiting, including the video interview process.[7] Many AI systems use computer vision video analysis, voice analysis, and natural language processing to assess and score candidates' video responses.[8]

The AI systems can detect an interviewee's facial expression, track eye contact, and capture speech pattern data.[9] These data points can help determine perceived "enthusiasm" and a score for each interviewee. In some video interview AI systems, each facial movement is considered a "facial action unit" and can make up a large percentage of an interviewee's score, while "audio features" like voice and tone can make up another significant percentage.[10] The generated score serves as a report card on the interviewee's "competencies and behaviors", which include their "willingness to learn" and "conscientiousness & responsibility."[11] Other systems categorize candidates based on four traits: "professionalism, energy levels, communication, and sociability."[12]

Hundreds of employers, including large corporations like Amazon, Target, Hilton, and Unilever, have adopted video interview AI systems.[13] Many HR executives praise these systems for saving

---

6. Sheridan Wall & Hilke Schellmann, *We Tested AI Interview Tools. Here's What We Found.*, MIT TECH. REVIEW (July 7, 2021), https://www.technologyreview.com/2021/07/07/1027916/we-tested-ai-interview-tools/.

7. *Id.*; *see* Neha Kaushik, *13 Best Video Interview Platforms in 2022 for Faster Hires*, GEEKFLARE (Feb 14, 2023), https://geekflare.com/video-interview-platforms/; INTERVIEWER.AI, https://interviewer.ai/ (last visited Nov. 23, 2022).

8. *See* Wall & Schellmann, *supra* note 6; *HireVue AI Explainability Statement*, HIREVUE 7-10 (2022), https://hirevue-api.dev-directory.com/wp-content/uploads/2022/04/HV_AI_Short-Form_Explainability_1pager.pdf; Drew Harwell, *A face-scanning algorithm increasingly decides whether you deserve the job*, WASHINGTON POST (Nov. 6, 2019, 12:21 PM), https://www.washingtonpost.com/technology/2019/10/22/ai-hiring-face-scanning-algorithm-increasingly-decides-whether-you-deserve-job/.

9. *See* Harwell, *supra* note 8.

10. *Id.*

11. *Id.*

12. *See What is Interviewer.AI's Critical Video Intelligence?*, INTERVIEWER.AI (June 13, 2022), https://interviewer.ai/help/what-is-interviewer-ais-video-intelligence/.

13. *See* Harwell, *supra* note 8; *Amazon + HireVue Amazon Drives Hiring Innovation and Experience with HireVue*, HIREVUE, https://www.hirevue.com/case-studies/amazon (last visited Feb. 19, 2023); Matt O'Brien, Associated Press, *Want a Job? Employers Say: Talk to the Computer*, TAMPA BAY TIMES (June 15, 2021), https://www.tampabay.com/news/business/2021/06/15/want-a-job-employers-say-talk-to-the-computer/.

thousands of hours and millions of dollars in hiring efforts while also helping to select quality candidates.[14]

B. Video Interview AI Regulation Landscape

The Video Interview AI regulation landscape is expanding rapidly. The following section lays out some such regulations.

1. *Illinois Artificial Intelligence Video Interview Act*. Illinois State Representative Jamie Andrade wrote and sponsored the Illinois Artificial Intelligence Video Interview Act (AIVIA) in 2019, the first law regulating AI's use in hiring decisions in the nation.[15] The Act requires employers to provide prior notice to applicants who are "based in Illinois" and obtain their affirmative consent to participate.[16] The notice must (a) inform applicants that "artificial intelligence analysis" may be used to evaluate their application, and (b) explain how the AI analysis works and "the general types of characteristics" used to evaluate the applicant.[17] The Act also restricts disclosure of the video interview to anyone other than "persons whose expertise or technology is necessary in order to evaluate an applicant's fitness for a position" and requires employers to ensure that all copies of the video interview, whether in the possession of the employer or any third party, are destroyed within 30 days of an applicant's request.[18]

Rep. Andrade proposed the AIVIA after he learned about how many job applicants were rejected at the "AI stage" of a job interview.[19] "What is the model employee? Is it a white guy? A white woman? Someone who smiles a lot?" said Rep. Andrade in an interview with the Washington Post.[20] He worried that accents, cultural differences, or people who declined to sit for the assessment could be unfairly punished by not being considered for the job.[21]

---

14. Robert Booth, *Unilever Saves on Recruiters by Using AI to Assess Job Interviews*, THE GUARDIAN (Oct. 25, 2019, 1:00 PM), https://www.theguardian.com/technology/2019/oct/25/unilever-saves-on-recruiters-by-using-ai-to-assess-job-interviews.

15. Illinois Artificial Intelligence Video Interview Act, 820 Ill. Comp. Stat. Ann. 42/1 (2020); Daniel Waltz et al., *Illinois Employers Must Comply with Artificial Intelligence Video Interview Act*, SHRM (Sept. 5, 2019), https://www.shrm.org/resourcesandtools/legal-and-compliance/state-and-local-updates/pages/illinois-artificial-intelligence-video-interview-act.aspx.

16. 820 Ill. Comp. Stat. Ann. 42/1.

17. 820 Ill. Comp. Stat. Ann. 42/5.

18. 820 Ill. Comp. Stat. Ann. 42/10; 820 Ill. Comp. Stat. Ann. 42/15.

19. *See* Harwell, *supra* note 8.

20. *Id.*

21. *Id.*

2. *New York City's Automated Employment Decision Tool Law (AEDT).* In a similar attempt to "reduce bias" by AI systems, New York City passed the AEDT which will become effective on April 15, 2023.[22] The term "automated employment decision tool" is broadly defined and lists "artificial intelligence" as one of the categories this law applies to.[23] The law requires, inter alia, employers who use an automated employment decision tool to (1) provide notice of the tool, (2) provide an alternative process or accommodation for applicants, (3) provide "information about the type of data collected for the automated employment decision tool", and (4) conduct an independent "bias audit" of the tool through an "independent auditor."[24] This independent bias audit must allow for a disparate impact analysis based on the Equal Employment Opportunity Commission (EEOC) framework.[25]

The AEDT defines "employment decision" as the "screen[ing of] candidates for employment or employees for promotion in NYC", which limits the scope and jurisdiction of this law.[26] Employers who do not comply may face fines of $500 for initial violations and up to $1,500 for each subsequent violation.[27] Presently, it remains unclear whether this law is limited to companies located in NYC or if it may be applied more broadly to any company hiring NYC residents.[28] Another ambiguity lies in which entities are to be considered "independent auditor[s]."[29]

---

22. *See* Sharon Goldman, *AI Bias Law Postponed Until April 15 As Unanswered Questions Remain*, VENTUREBEAT (Dec. 12, 2022), https://venturebeat.com/ai/for-nycs-new-ai-bias-law-unanswered-questions-remain/.

23. *Id.*

24. *See* Nicole Price, *New York City's New Law Regulating the Use of Artificial Intelligence in Employment Decisions*, JDSUPRA (April 11, 2022), https://www.jdsupra.com/legalnews/new-york-city-s-new-law-regulating-the-3122279/.

25. Automated Employment Decision Tools, New York City, N.Y. Administrative Code § 20-870; K.C. Halm et al., *NYC Proposes Rules to Implement New AI Audit Law*, DAVIS WRIGHT TREMAINE LLP (Oct. 6, 2022), https://www.dwt.com/blogs/employment-labor-and-benefits/2022/10/nyc-ai-audit-law-employment-decisions. ("This means determining the 'selection' rate for each protected gender/race/ethnicity group and comparing it to the rate for the most-selected group").

26. *See* David O. Klein, *NYC's Use Of Automated Employment Decision Tools ("AEDTs") Law: The Crucial Details*, MONDAQ (Sept. 15, 2022), https://www.mondaq.com/unitedstates/employee-rights-labour-relations/1230608/nyc39s-use-of-automated-employment-decision-tools-aedts-law-the-crucial-details.

27. *Id.*

28. *Id.*

29. *See NYC Issues Proposed Regulations Regarding Automated Employment Decision Tools Law*, COOLEY LLP (Oct. 17, 2022), https://www.cooley.com/news/insight/2022/2022-10-17-nyc-issues-proposed-regulations-regarding-automated-employment-decision-tools-law.

3. *EEOC Guidance.* In 2022, the EEOC released technical guidance on how algorithmic hiring tools discriminate against people with disabilities.[30] The guidance addresses AI-based software, including "video interviewing software that evaluates candidates based on their facial expressions and speech patterns."[31]

The guidance provides examples of how algorithmic decision-making tools could screen out an individual because of a disability: "[f]or example, video interviewing software that analyzes applicants' speech patterns to reach conclusions about their ability to solve problems is not likely to score an applicant fairly if the applicant has a speech impediment that causes significant differences in speech patterns."[32]

The guidance recommends ways that employers using algorithmic decision-making tools can reduce the chances of screening out individuals with disabilities.[33] Such recommendations include providing job applicants with "alternative formats and alternative tests" and providing in advance of the assessment "as much information about the tool as possible" such as which "traits or characteristics" are measured, and which methods are used to measure them.[34]

At the end of the guidance, the EEOC encourages candidates to file formal charges of discrimination with the EEOC if a candidate feels they were discriminated against on the basis of disability by an algorithmic hiring process.[35]

---

30. *See The Americans with Disabilities Act and the Use of Software, Algorithms, and Artificial Intelligence to Assess Job Applicants and Employees*, EEOC (2022), https://www.eeoc.gov/laws/guidance/americans-disabilities-act-and-use-software-algorithms-and-artificial-intelligence.

31. *Id.*

32. *Id.*

33. *Id.*

34. *Id.*

35. *See* Alex Engler, *The EEOC wants to make AI hiring fairer for people with disabilities*, BROOKINGS INSTITUTION (May 26, 2022) https://www.brookings.edu/blog/techtank/2022/05/26/the-eeoc-wants-to-make-ai-hiring-fairer-for-people-with-disabilities/.

4. *OSTP's Blueprint for an AI Bill of Rights.* In October 2022, the White House Office of Science and Technology Policy (OSTP) published *Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People*, a white paper intended to "support the development of policies and practices that protect civil rights and promote democratic values in the building, deployment, and governance of automated systems."[36] The white paper identifies five guiding principles for the design, use, and deployment of AI systems.[37]

The first guiding principle (Safe and Effective Systems) recommends that AI systems "should be developed with consultation from diverse communities, stakeholders, and domain experts to identify concerns, risks, and potential impacts of the system."[38] The systems should also undergo pre-deployment testing and ongoing monitoring.[39]

The second guiding principle (Algorithmic Discrimination Protections) expands on when algorithmic discrimination can occur, similar to the EEOC guidance.[40] It provides that this discrimination may happen when AI systems "contribute to unjustified different treatment or impacts disfavoring people based on their race, color, ethnicity, sex (including pregnancy, childbirth, and related medical conditions, gender identity, intersex status, and sexual orientation), religion, age, national origin, disability, veteran status, genetic information, or any other classification protected by law."[41] It recommends that AI system "[d]esigners, developers, and deployers should take proactive and continuous measures" to use and design systems "in an equitable way."[42]

The third guiding principle (Data Privacy) provides that end users should be protected from "abusive data practices" and "should have agency over how data about [end users] is used." It suggests that AI system designers, developers, and deployers should seek permission in "understandable" plain language and "respect [end users'] decisions" regarding the "collection, use, access, transfer, and deletion" of data.[43]

The fourth guiding principle (Notice and Explanation) recommends that AI system designers, developers, and deployers should

---

36. *Blueprint for an AI Bill of Rights*, *supra* note 5.
37. *Id.*
38. *Id.*
39. *Id.*
40. *Id.*
41. *Id.*
42. *Id.*
43. *Id.*

provide "accessible plain language documentation", which entails descriptions of the overall system, notice that such system is in use, the "individual or organization responsible" for the system, and that any explanation of outcomes are "clear, timely, and accessible."[44] The explanations must be "technically valid, meaningful, and useful to [end users]" and must be "calibrated to the level of risk based on the context."[45]

The fifth and final guiding principle (Human Alternatives, Consideration, and Fallback) recommends AI system designers, developers, and deployers provide an "opt-out" and access to a human person who can quickly consider and remedy any problems with the AI system.[46]

While these principles lay out important concerns surrounding AI use, this "blueprint" still only remains as mere suggestions. The blueprint does not establish any enforcement authority or even any oversight authority.

5. *AI Training Act of 2022.* Also in October 2022, Congress passed the AI Training Act, which requires the White House's Office of Management and Budget to establish an AI training program for federal employees. The act is intended to "ensure that [AI] is used in a way that is consistent with [the] nation's values."[47] The AI training program will include information on "how AI works", how AI can benefit the federal government, and "risks posed by AI."[48]

III. THE DEMAND FOR EXPLAINABILITY

A. Explainable AI (XAI) Generally

Many of the above regulations or guidance on AI tools in employment aim at limiting discriminatory effects.[49] They demand some level of explanation or information on how the AI system works.[50] For example in the AIVIA, the law requires that employers provide an explanation to applicants on how the AI system

---

44. *Id.*

45. *Id.*

46. *Id.*

47. *See* Natalie Alms, *AI training bill becomes law*, FCW (Oct. 18, 2022), https://fcw.com/workforce/2022/10/ai-training-bill-becomes-law/378587/.

48. AI Training Act, S.2551, 117th Cong. (2022).

49. EEOC, *supra* note 30; Illinois Artificial Intelligence Video Interview Act, *supra* note 15; Automated Employment Decision Tool, *supra* note 25.

50. EEOC, *supra* note 30; Illinois Artificial Intelligence Video Interview Act, *supra* note 15; Automated Employment Decision Tool, *supra* note 25.

works and what characteristics are used to evaluate the applicant.[51] This concept falls under the category of Explainable AI (XAI). XAI refers to a process that allows human users to understand and trust the results or output created by AI systems.[52] XAI can be used to assess model accuracy, fairness, transparency, and outcomes in decision-making.[53] The scope and "importance" of XAI varies—while XAI could be useful in public safety, legal, educational, and geopolitical applications, XAI need not be as important in games or email filters.[54]

Generally, users and other external stakeholders are afforded little insight into the behind-the-scenes workings of AI systems that impact their lives and opportunities.[55] Some AI developers argue that XAI is crucial in building trust and confidence in AI systems as they are put into production and operation.[56] Other industry analysts highlight that explanations can help provide "transparency" but are often incomplete and don't provide end users with any meaningful perspective.[57]

---

51. Illinois Artificial Intelligence Video Interview Act, *supra* note 15.

52. *See Explainable AI*, IBM (last visited Nov. 23, 2022), https://www.ibm.com/watson/explainable-ai.

53. *Id.*

54. *See* Guy Pearce, *Explainable Artificial Intelligence (XAI): Useful But Not Uncontested*, INFORMATION SYSTEMS AUDIT AND CONTROL ASSOCIATION (April 6, 2022), https://www.isaca.org/resources/news-and-trends/newsletters/atisaca/2022/volume-14/explainable-artificial-intelligence-useful-but-not-uncontested.

55. *See* Jessica Newman, *Explainability won't save AI*, BROOKINGS INSTITUTION (May 19, 2021), https://www.brookings.edu/techstream/explainability-wont-save-ai/; *See* Abhinav Raghunathan, *The Ethical AI Startup Ecosystem 03: ModelOps, Monitoring, and Observability*, MONTREAL AI ETHICS INSTITUTE (Aug. 31, 2022), https://montrealethics.ai/the-ethical-ai-startup-ecosystem-03-modelops-monitoring-and-observability/?utm_source=substack&utm_medium=email ("Lawyers and businesspeople have no insight into how the models work. Stakeholders are left in the dark and the barrier to bring them in is the communication of context via a shared language, which doesn't currently exist").

56. *Explainable AI*, *supra* note 52.

57. Newman, *supra* note 55.

B. Different Forms of Explainability and Efforts by Private Actors

1. *Global v. Local Explainability.* Researchers are exploring ways to evaluate how useful AI explanations are for end users.[58] As of present day, there is no standardized set of explainability metrics. However, two broad categories have emerged: global and local explainability. Global explanations provide all outcomes of a given process and "describe the rules or formulas specifying relationships among input variables."[59] In contrast, local explainability provides the rationale behind a specific result, output, or decision.[60]

Global explainability can be important for broader-scale scientific understanding or bias detection.[61] It can also help spot-check how features (like characteristics in a video interview) in the model contribute to the overall predictions or outcomes of the model.[62]

2. *Model Cards.* AI developers have attempted an industry-wide solution for XAI: Model cards.[63] A model card is "a short document that provides key information about a machine learning model."[64] Some AI developers have analogized model cards to nutrition facts labels on the back of food products.[65] Such AI developers also herald that model cards are designed to be read and used by experts and non-experts.[66] While there is no standardized template for model cards, many include the following sections:[67]

- Model details: Developer information, Model Version

---

58. *See* Hoffman et al., *Metrics for Explainable AI: Challenges and Prospects*, ARXIV (Feb. 1, 2019), https://arxiv.org/abs/1812.04608.

59. Candelon et al., *AI Regulation Is Coming*, HARVARD BUS. REV. (Oct., 2021), https://hbr.org/2021/09/ai-regulation-is-coming.

60. *Id.*

61. *See* Doshi-Velez et al., *Towards A Rigorous Science of Interpretable Machine Learning*, ARXIV (March 2, 2017), https://arxiv.org/abs/1702.08608.

62. *See* Aparna Dhinakaran, *A Look Into Global, Cohort and Local Model Explainability*, TOWARDS DATA SCIENCE (Sept. 21, 2021), https://towardsdatascience.com/a-look-into-global-cohort-and-local-model-explainability-973bd449969f.

63. *Model Cards*, GOOGLE, https://modelcards.withgoogle.com/about (last visited Nov. 24, 2022).

64. *Id.*; *Model Cards*, KAGGLE, https://www.kaggle.com/code/var0101/model-cards (last visited Nov. 24, 2022).

65. *Model Cards*, *supra* note 64.

66. *Model Cards*, *supra* note 63.63

67. *Model Cards*, *supra* note 64.

- Intended Use: Use cases, scope, intended user description.
- Factors: These vary by model. For example, a smiling detection model's results may vary by demographic factors like age, gender, ethnicity, etc.
- Metrics: Whether it is a classification system (where the output is a class label) or a score-based analysis (where the output is a score).
- Evaluation Data: Which datasets were used for evaluating model performance and why.
- Training Data: Which datasets were used for training the model.
- Quantitative Analyses: How did the model perform on the chosen metrics and explain factor intersections. For example:
    - In the smiling detection example, performance is broken down by age (young, old), gender (female, male), and then both (old-female, old-male, young-female, young-male).
- Ethical Considerations: This varies by model but can include whether sensitive data was used to train the model, whether the model has implications for "human life, health, or safety", and how risk was mitigated. Examples of sensitive data include payment information, personally identifiable information, and protected health information.[68]

68. *See What is Sensitive Data & How to Determine it?*, SECURITI.AI (May 17. 2021), https://securiti.ai/blog/what-is-sensitive-data/; *See also* Cameron F. Kerry, *Protecting privacy in an AI-driven world*, BROOKINGS INSTITUTION (Feb. 10, 2020), https://www.brookings.edu/research/protecting-privacy-in-an-ai-driven-world/.

Model Card for CTRL - A Conditional Transformer Language Model for Controllable Generation

The CTRL Language Model analyzed in this card generates text conditioned on control codes that specify domain, style, topics, dates, entities, relationships between entities, plot points, and task-related behavior.

On this model card, you can learn more about how this model was trained, its capabilities, its intended use, and its limitations.

Model Details

Organization
Salesforce Research

Model date	*Model type*	*Input*
September 10, 2019	Language model	Text

Information about parameters
1.63 billion-parameter transformer language model

Output
The model can generate text conditioned on control codes that specify domain, style, topics, dates, entities, relationships between entities, plot points, and task-related behavior

Read the full paper here:
https://arxiv.org/abs/1909.05858

Access the public code here:
https://github.com/salesforce/ctrl

Citation details
Title: CTRL - A Conditional Transformer Language Model for Controllable Generation
Authors: Keskar, Nitish Shirish and McCann, Bryan and Varshney, Lav and Xiong, Caiming and Socher, Richard
Journal: arXiv preprint arXiv:1909.05858
Year: 2019

License: BSD 3-Clause (https://github.com/salesforce/ctrl/blob/master/LICENSE.txt)

Metrics

Model performance measures
Performance evaluated on qualitative judgments by humans as to whether the control codes lead to text generated in the desired domain

Training Data

This model is trained on 140 GB of text drawn from a variety of domains: Wikipedia (English, German, Spanish, and French), Project Gutenberg, submissions from 45 subreddits, OpenWebText, a large collection of news data, Amazon Reviews, Europarl and UN data from WMT (En-De, En-Es, En-Fr), question-answer pairs (no context documents) from ELI5, and the MRQA shared task, which includes Stanford Question Answering Dataset, NewsQA, TriviaQA, SearchQA, HotpotQA, and Natural Questions. See the paper for the full list of training data

Intended Use

Primary intended use
1. Generating artificial text in collaboration with a human, including but not limited to:
   - Creative writing
   - Automating repetitive writing tasks
   - Formatting specific text types
   - Creating contextualized marketing materials
2. Improvement of other NLP applications through fine-tuning (on another task or other data, e.g. fine-tuning CTRL to learn new kinds of language like product descriptions)
3. Enhancement in the field of natural language understanding to push towards a better understanding of artificial text generation, including how to detect it and work toward control, understanding, and potentially combating potentially negative consequences of such models

Primary intended users
General audiences
NLP researchers

Out-of-scope use cases
- CTRL should not be used for generating artificial text without collaboration with a human.
- It should not be used to make normative or prescriptive claims.
- This software should not be used to promote or profit from:
  - violence, hate, and division;
  - environmental destruction;
  - abuse of human rights; or
  - the destruction of people's physical and mental health.

Ethical Considerations

We recognize the potential for misuse or abuse, including use by bad actors who could manipulate the system to act maliciously and generate text to influence decision-making in political, economic, and social settings. False attribution could also harm individuals, organizations, or other entities. To address these concerns, the model was evaluated internally as well as externally by third parties, including the Partnership on AI, prior to release.

To mitigate potential misuse to the extent possible, we stripped out all detectable training data from undesirable sources. We then red-teamed the model and found that negative utterances were often placed in contexts that made them identifiable as such. For example, when using the 'News' control code, hate speech could be embedded as part of an apology (e.g. "the politician apologized for saying [insert hateful statement]"), implying that this type of speech was negative. By pre-selecting the available control codes (omitting, for example, Instagram and Twitter from the available domains), we are able to limit the potential for misuse.

In releasing our model, we hope to put it into the hands of researchers and prosocial actors so that they can work to control, understand, and potentially combat the negative consequences of such models. We hope that research into detecting fake news and model-generated content of all kinds will be pushed forward by CTRL. It is our belief that these models should become a common tool so researchers can design methods to guard against malicious use and so the public becomes familiar with their existence and patterns of behavior.

Caveats and Recommendations

- A recommendation to monitor and detect use will be implemented through the development of a model that will identify CTRL-generated text.
- A second recommendation to further screen the input into and output from the model will be implemented through the addition of a check in the CTRL interface to prohibit the insertion into the model of certain negative inputs, which will help control the output that can be generated.
- The model is trained on a limited number of languages: primarily English and some German, Spanish, French. A recommendation for a future area of research is to train the model on more languages.

Send questions or comments to ctrl-monitoring@salesforce.com

Model cards can provide end users or applicants with some information on the AI system impacting their employment opportunities. However, model cards remain a fully voluntary industry practice, with no governing entity or regulation that tracks, enforces, or reviews such model cards.[69] The analogy of model cards to the nutrition labels misleads: nutrition facts labels are regulated and mandatory by law.[70] Unlike nutrition fact labels, there is no liability for creating false or erroneous model cards.[71]

While model cards may contain a section on "ethical considerations", they do not deter misuse. For example, Salesforce's CTRL model card acknowledges the "potential for misuse or abuse" and addressed this concern by evaluating the model internally and by third parties.[72] But again, these internal and external evaluations do not provide any action or remedy should the model cause harm in any capacity. Moreover, it is entirely unclear what internal and external evaluation was done or the results of any such evaluations. Thus, while model cards could provide a way for AI developers to share their models' functionality, they cannot serve as a meaningful substitution for regulation or oversight.

---

69. *Model Cards*, *supra* note 64.

70. U.S. FOOD & DRUG ADMIN., GUIDANCE FOR INDUSTRY: FOOD LABELING GUIDE (2013).

71. *FDA: Foods Must Contain What Label Says*, U.S. FOOD AND DRUG ADMIN. (Feb. 4, 2013), https://www.fda.gov/consumers/consumer-updates/fda-foods-must-contain-what-label-says.

72. *Model Card for CRTL*, SALESFORCE (April 20, 2020), https://github.com/salesforce/ctrl/blob/master/ModelCard.pdf.

3. *Explainability Statements.* Certain AI developers publish explainability statements instead of model cards. For example: HireVue, one of the leading video interview AI developers, published an explainability statement for its customers (i.e., employers) and their candidates.[73] This thirty-page document provides an in-depth dive into different models, data collection methods, bias monitoring methods, and assessment scoring methods.[74] While explainability statements can be much more thorough than model cards, they suffer from the same problems: no oversight, no deterrent, and no repercussions for false or misleading information.

Moreover, the majority of customers who lack advanced degrees in AI likely fail to assess in any critical way statements such as these in the document.[75] For example, according to a 2020 study on explainable AI deployments from 20 different organizations, most ML engineers used AI explanations as "sanity checks" for other engineers or research scientists, and were not used to "directly inform end users."[76] Further, studies that examine trust and use of explainability tend to use subjects already familiar with machine learning, signaling that some foundational exposure is necessary for meaningful assessment of explanations.[77]

In 2021, 23.5% of individuals ages 25 and older in the U.S. had a bachelor's degree as their highest degree, and that number dwindles to 14.4% when looking at advanced degrees like master's, professional, or doctoral degrees.[78] Although education level is not a direct indicator of whether a job applicant would understand an

---

73. *HireVue AI Explainability Statement*, *supra* note 8.

74. *Id.*

75. Candelon et al., *supra* note 59; *See* Preece et al., *Stakeholders in Explainable AI*, ARXIV (Sep. 29, 2018), https://arxiv.org/abs/1810.00184 ("Providing detailed transparency-based explanations may also overwhelm the recipient — more information is not necessarily better for user performance.").

76. *See* Bhatt et al., *Explainability in Machine Learning Deployment*, ARXIV (July 10, 2020), https://arxiv.org/abs/1909.06342.

77. Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. *"Why should I trust you?" Explaining the predictions of any classifier*. PROCEEDINGS OF THE 22ND ACM SIGKDD INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING 1135 (2016), ("Since this task requires some familiarity with the notion of spurious correlations and generalization, the set of subjects for this experiment were graduate students who have taken at least one graduate machine learning course."); Min Kyung Lee, *Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management*, BIG DATA & SOCIETY, Jan.–June 2018, at 5 (Participants were recruited from Amazon Mechanical Turks and "were fairly well educated").

78. *Census Bureau Releases New Educational Attainment Data*, U.S. CENSUS BUREAU (Feb. 24, 2022), https://www.census.gov/newsroom/press-releases/2022/educational-attainment.html#:~:text=10.9%25%20in%202011.-,Sex,women%20and%2046.9%25%20were%20men.

explanation for AI, it does serve as a proxy to show that an average American job applicant may not be fully equipped to understand an AI explanation, when or if presented with one.

IV. THE CASE AGAINST FEDERAL LEGISLATION ON AI USED IN EMPLOYMENT DECISIONS

A. The EEOC's Lack of Enforcement and the AAA's Overreach

With the EEOC's guidance for employers using automated decision-making tools, some groups argue that federal legislation on AI should be the direction forward.[79] Certain members of Congress also share this view and have even introduced a new bill titled the Algorithmic Accountability Act of 2022 (AAA).[80] The AAA would require companies to assess the impacts of the automated systems they use and sell, and would be regulated through FTC guidelines.[81] Just like other types of federal legislation, which are ideally applied equally in all 50 states,[82] a federal law governing AI use would provide uniformity and predictability to AI developers and end users. This bill would align with the numerous federal laws that already govern employment and hiring decisions; the U.S. Department of Labor administers and enforces more than 180 federal laws regarding employment and labor in the U.S.[83] In addition, the EEOC enforces a wide variety of laws ranging from the Americans with Disabilities Act (ADA), the Equal Pay Act (EPA),

---

79. *See* David A. Teich, *The White House Blueprint For An AI Bill Of Rights – Time For More*, FORBES (Oct. 6, 2022, 11:56 AM), https://www.forbes.com/sites/davidteich/2022/10/06/the-white-house-blueprint-for-an-ai-bill-of-rights--time-for-more/?sh=e15ccc8739a2.

80. *See* Press Release, Jon Wyden, Senator, State of Oregon, Wyden, Booker and Clarke Introduce Algorithmic Accountability Act of 2022 To Require New Transparency and Accountability For Automated Decision Systems (Feb. 3, 2022), https://www.wyden.senate.gov/news/press-releases/wyden-booker-and-clarke-introduce-algorithmic-accountability-act-of-2022-to-require-new-transparency-and-accountability-for-automated-decision-systems; Proponents of the algorithmic accountability act include U.S. Senator Ron Wyden, D-Ore., with Senator Cory Booker, D-N.J., and Representative Yvette Clarke, D-N.Y. *Id.* This act also has 39 congressional co-sponsors. Algorithmic Accountability Act, H.R. 6580, 117th Cong. (2022).

81. *See Algorithmic Accountability Act of 2022*, OFFICE OF SENATOR RON WYDEN (Feb. 3, 2022), https://www.wyden.senate.gov/download/one-pager-bill-summary-of-the-algorithmic-accountability-act-of-2022.

82. *See generally How Laws Are Made and How to Research Them,* USA.GOV, https://www.usa.gov/how-laws-are-made#:~:text=Federal%20laws%20apply%20to%20people,they%20agree%20with%20the%20Constitution (last visited March 1, 2023).

83. *Summary of the Major Laws of the Department of Labor*, U.S. DEP'T OF LABOR, https://www.dol.gov/general/aboutdol/majorlaws (last visited Nov. 24, 2022).

and Title VII of the Civil Rights Act,[84] all of which are geared towards reducing or eliminating prohibited forms of discrimination against job applicants.

Nevertheless, the EEOC's guidance on AI in employment remains voluntary with no real enforcement abilities baked into the guidance. Similarly, the AAA—which is still just a bill—also would simply require companies to self-assess, self-document, and self-report their AI systems according to FTC's non-existing guidelines for assessment and reporting.[85] The AAA's enforcement section remains vague, stating that "the same jurisdiction, powers, and duties as . . . the Federal Trade Commission Act" would apply to the AAA.[86] Thus, both the EEOC guidance and the AAA suffer from their lack of enforceability.

B. The AAA's Flaws Show that Congress is Not Equipped to Regulate AI

The AAA is intended to cover all types of algorithms that are used to make a "critical decision" which includes employment.[87] Ironically though, the AAA does not define the term "algorithm" and instead defines "automated decision system" as "any system, software, or process . . . that uses computation, the result of which serves as a basis for a decision or judgment."[88] This definition includes "machine learning, statistics, or . . . artificial intelligence techniques" and excludes "passive computing infrastructure."[89]

The AAA's broad approach in lumping all algorithms together under one federal law, regardless of function, model, industry, or any other variable, is an overreach.

First, many non-AI systems could fall under this broad definition of an automated decision system. For example, a system made by a human that uses computation and is used to make a critical decision could easily fall into this category.

Second, researchers are already finding gaps and issues arising within the AAA. For example, some argue that because certain entities in the healthcare sector are exempt from FTC jurisdiction, it could affect how algorithms are developed and marketed, further distorting and creating an imbalance between which AI systems

---

84. *Statutes Enforced by EEOC*, EEOC, https://www.eeoc.gov/statutes/laws-enforced-eeoc (last visited March 11, 2023).

85. *Algorithmic Accountability Act of 2022, supra* note 80.

86. Algorithmic Accountability Act, H.R. 6580, 117th Cong. § 9(a)(2)(A) (2022).

87. *See Press Release*, *supra* note 80.

88. Algorithmic Accountability Act, H.R. 6580, 117th Cong. § 2(2) (2022).

89. *Id.*

are "regulated" and which ones are not.[90] Over 75% of hospitals in the United States are either not-for-profit or government owned, and are thus exempt from FTC jurisdiction.[91] If the AAA is passed, none of these entities would have to conduct AAA-required assessments of the automated decision tools they develop or use.[92] In addition, if only for-profit hospitals are conducting evaluations of automated decision tools in healthcare, these assessments would not accurately represent the overall performance of these systems.[93]

A sweeping federal bill like the AAA—which is supposed to cover all kinds of AI systems including those used in employment decisions—would be imprudent and unwise. This broad-stroke proposal signals that Congress members currently lack the expertise and depth of knowledge to regulate this industry. In a press release by one of the proposers of the AAA, Rep. Yvette D. Clarke stated that "[a]lgorithms shouldn't have an exemption from our anti-discrimination laws."[94] But ironically, given that the AAA would be powered through the FTC, numerous employers and entities would remain exempt from FTC jurisdiction. What about the not-for-profit or government hospitals that employ thousands of individuals? Since the AAA would only apply to entities that make over $50 million—what about employer entities that make less than that?[95]

The passage of the AAA could also chill the AI industry and innovation. For example: the EU's General Data Protection Regulation (GDPR) resulted in a severe drop in the number of venture deals.[96] In effect in the EU since 2018, the GDPR imposes obligations onto organizations anywhere, so long as they target or collect

---

90. Mithal et al., *The Algorithmic Accountability Act: Potential Coverage Gaps in the Healthcare Sector*, ANTITRUST MAG. ONLINE, Aug. 2022, at 5.

91. According to the American Hospital Association's Annual Survey from 1999 – 2020: 18.5% are State/Local Government; 57.6% are Non-Profit. *Hospitals by Ownership Type*, KFF, https://www.kff.org/other/state-indicator/hospitals-by-ownership/ (last visited Nov. 24, 2022); Mithal et al., *supra* note 90, at 5.

92. Mithal et al., *supra* note 90, at 5.

93. *Id.* at 6.

94. *Booker, Wyden, Clarke Introduce Bill Requiring Companies To Target Bias In Corporate Algorithms*, OFFICE OF SENATOR CORY BOOKER (April 10, 2019) https://www.booker.senate.gov/news/press/booker-wyden-clarke-introduce-bill-requiring-companies-to-target-bias-in-corporate-algorithms.

95. Algorithmic Accountability Act, H.R.6580, 117th Cong. (2022).

96. *See* Jia et al., *The Short-Run Effects of GDPR on Technology Venture Investment*, NATIONAL BUREAU OF ECONOMIC RESEARCH (Nov. 1, 2018), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3278912 ("We found evidence suggesting negative and pronounced effects following the rollout of GDPR on the number of venture deals, particularly in the period immediately after GDPR's rollout, and particularly for newer, data-related, and consumer facing ventures.").

data related to people in the EU.[97] Noncompliance with the GDPR leads to harsh fines against those who violate its privacy and security standards, with penalties reaching tens of millions of euros.[98] Many provisions of the GDPR impact AI developers in the EU, since many of them focus on limiting data collection and data use.[99]

In the AI industry, the GDPR rollout led AI developing companies to reallocate their limited funding to handle GDPR compliance.[100] In the same study, most of the companies indicated that, because of the GDPR, their training datasets were deficient, and this problem has led to unrepresentative or inaccurate models.[101]

The limited dataset issue is something AI developers are all too familiar with. Having access to "enough" data is important, since large amounts of data are at the heart of any AI system or model.[102] Not having enough data can lead to mediocre AI systems that do not provide any useful tool or outcome, resulting in a meaningless model. Numerous AI developers claim that "your model is only as good as your data", which refers to both the quality and quantity of the data.[103] Thus, Congress must proceed with caution if it chooses to regulate this developing industry since there are tradeoffs between the quality and innovation of AI systems in limiting access to useful data.

---

97. *What is GDPR, the EU's new data protection law?*, GDPR.EU, https://gdpr.eu/what-is-gdpr/ (last visited Nov. 24, 2022).

98. *Id.*

99. *See* Avantika Bhandari, *The Impact of the GDPR on Artificial Intelligence*, MONTREAL AI ETHICS INSTITUTE (Feb. 20, 2022), https://montrealethics.ai/the-impact-of-the-gdpr-on-artificial-intelligence/.

100. Bessen et al., *GDPR and the Importance of Data to AI Startups*, SCHOLARLY COMMONS AT BOSTON UNIVERSITY SCHOOL OF LAW (April 1, 2022), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3576714.

101. *Id.* ("Data is important to the success of AI startups. They need access to training data to run sophisticated algorithms, that in some cases are at the cutting edge of technology, such as neural networks and ensemble learning. [Majority of firms in this study] indicate[d] that their current training data are deficient . . . and that they would benefit from refreshing their models more often. Though GDPR drives changes that are important to safeguarding personally identifiable information, there are also costs associated with this increased regulation").

102. *See generally* Alexandre Gonfalonieri, *5 Ways to Deal with the Lack of Data in Machine Learning*, KDNUGGETS https://www.kdnuggets.com/2019/06/5-ways-lack-data-machine-learning.html (last visited Nov. 24, 2022).

103. *The Size and Quality of a Data Set*, GOOGLE MACHINE LEARNING (July 18, 2022), https://developers.google.com/machine-learning/data-prep/construct/collect/data-size-quality; Thomas Redman, *If Your Data Is Bad, Your Machine Learning Tools Are Useless*, HARVARD BUS. REV. (April 2, 2018), https://hbr.org/2018/04/if-your-data-is-bad-your-machine-learning-tools-are-useless.

V. AI SHOULD NOT BE HELD TO A HIGHER STANDARD THAN HUMANS

The movement for XAI and "ethical AI" stems from a fear of creating and using decisions that are not justifiable, legitimate, or simply do not allow for detailed explanations.[104] General users and government entities have expressed concerns surrounding discrimination by algorithms.[105] It has become a common criticism of AI assessments of any kind that there may be bias hardwired into the AI system that could severely hurt people.[106] Organizations such as the Algorithmic Justice League exist to "illuminate the social implications and harms of artificial intelligence."[107] There does not appear to be a parallel "Human Fallibility League" even if the number of errors is vastly greater.

Articles among numerous news agencies and academics alike warn readers about the horrors of algorithmic decision-making, and if algorithmic decision-making tools are used "at every turn", then it would "add up to a discriminatory society."[108] This type of fear-mongering has become increasingly commonplace and it plagues many discussions surrounding AI.[109] To illustrate some

---

104. *See* Arrieta et al., *Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI*, ARXIV (Dec. 26, 2019), https://arxiv.org/abs/1910.10045; *See* Patrick Thibodeau, *A backlash emerges over automated interviewing*, TECHTARGET (Oct. 17, 2019) https://www.techtarget.com/searchhrsoftware/news/252472497/A-backlash-emerges-over-automated-interviewing ("'I think one of the biggest concerns people have is the cultural differences in society,' Andrade said. In some cultures, people don't make direct eye contact or do not smile as much. 'Does that now knock them out of any customer service position?' he said").

105. *Blueprint for an AI Bill of Rights*, *supra* note 5.

106. *See generally* David Moschella, *AI Bias Is Correctable. Human Bias? Not So Much*, INFORMATION TECHNOLOGY & INNOVATION FOUNDATION (April 25, 2022), https://itif.org/publications/2022/04/25/ai-bias-correctable-human-bias-not-so-much/ ("Fears of AI bias also reflect the familiar hi-tech double-standard. A single potent example of an AI failure—such as the racist rantings of Microsoft's chatbot, Tay, or a deadly accident involving a self-driving Tesla—will get vastly more media coverage than everyday car crashes or human misbehavior. Likewise, organizations such as the Algorithmic Justice League will continue to effectively publicize AI's risks, failures, and injustices, sometimes in ways that are misleading").

107. *About Page*, ALGORITHMIC JUSTICE LEAGUE, https://www.ajl.org/about (last visited Nov. 24, 2022) ("The deeper we dig, the more remnants of prejudice we will find in our technology. We cannot afford to look away this time because the stakes are simply too high. We risk losing the gains made with the civil rights movement and other movements for equality under the false assumption of machine neutrality").

108. *See* Khoo et al., *D.C. might ban 'algorithmic discrimination.' That's good for everyone.*, WASHINGTON POST (Oct. 7, 2022, 10:00 AM), https://www.washingtonpost.com/opinions/2022/10/07/dc-should-ban-algorithmic-discrimination/.

109. Marr, *supra* note 1; Philip Ball, *How dangerous is AI?*, PROSPECT MAG. (Dec. 22, 2021), https://www.prospectmagazine.co.uk/science-and-technology/how-dangerous-is-ai-artificial-intelligence-review-bbc-reith-lectures; *See* Kelsey Piper, *The case for taking AI*

absurd commentary around AI, the following quote is from Vox, a news website with a readership of nearly 20 million visitors just in 2021:[110]

"Here's one scenario that keeps experts up at night: We develop a sophisticated AI system with the goal of, say, estimating some number with high confidence. The AI realizes it can achieve more confidence in its calculation if it uses all the world's computing hardware, and it realizes that releasing a biological superweapon to wipe out humanity would allow it free use of all the hardware. Having exterminated humanity, it then calculates the number with higher confidence."

While not all articles contain such outlandish perspectives, there still exists a significant amount that tend not to be solution-orientated and do not add productive dialogue to the societal implications of AI.

Perhaps this "Responsible AI" movement is demanding something greater than what we require of humans.[111] For example, HireVue acknowledges that their "[AI] systems might be less capable of giving high scores to atypical . . . candidates whose answers . . . are radically different."[112] But even in traditional hiring settings, HireVue argues that "it cannot be guaranteed that those candidates would have been selected by humans", given that their answers were unusual.[113] We routinely accept human conclusions and decisions without fully understanding their origin, like when we trust a doctor's prescription despite us not having extensive (and for some, basic) medical knowledge. Why must AI decisions be held to a greater standard?[114]

If the end goal is to reduce bias and provide equal opportunity for all applicants, using AI hiring tools (rather than trying to fundamentally and idealistically fix human nature) might move us forward toward that goal. In designing AI-based assessments,

---

seriously as a threat to humanity, VOX (Oct. 15, 2020, 1:30 PM), https://www.vox.com/future-perfect/2018/12/21/18126576/ai-artificial-intelligence-machine-learning-safety-alignment.

110. Piper, *supra* note 109.

111. *See generally* Robert Atkinson, *'It's Going to Kill Us!' And Other Myths About the Future of Artificial Intelligence,* INFORMATION TECHNOLOGY & INNOVATION FOUNDATION, (June 6, 2016), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3066182 ("In summary, no technology is an unalloyed good, but virtually all technologies that emerge in the marketplace produce benefits vastly in excess of their costs. AI is and will continue to be no different").

112. *HireVue AI Explainability Statement*, *supra* note 8, at 19.

113. *Id.*

114. *See* Elizabeth Holm, *In defense of the black box*, SCIENCE (April 5, 2019), https://par.nsf.gov/servlets/purl/10101197.

developers like HireVue can minimize the impact of certain data points which could lead to discriminatory decisions. In traditional job assessments, questions that are found to cause adverse impact can be removed by the employer, but it can lead to "significant accuracy reductions" since each pre-selected question is already part of the pre-made assessment question bank.[115] Many human biases are also unconscious in nature and can include stereotypes, preconceptions, self-interests, etc.[116]

By contrast, AI systems can assign "weights" to each word, word type, statement, and contextual phrase which predict the competency being rated or measured.[117] In an example provided by HireVue: If the word "aggressive" predicts the competency 'Drive for Results', and more men than women use the word, it might add bias against women in the competency score. By lowering the weight assigned to the word "aggressive", the bias against women will be decreased and the prediction of the competency still remains.[118] In other AI hiring tools, the removal or omission of characteristics that identify "gender, national origin, skin color, and age" also allow hiring managers to focus purely on candidates' qualifications and experience.[119] Unlike many human decisions, "decisions made by AI can in principle (and increasingly in practice) be opened up, examined, and interrogated."[120] For example, HireVue's video AI interviewing system provides a "Candidate Insights Report" to hiring managers and recruiters for each candidate's communication competency.[121] Key behaviors from the video interview—such as use of proper grammar and delivery of clear and concise messages in a "logical sequence"—allow for such hiring managers to understand precisely why the AI system scored a candidate the way it did.[122]

Whether decisions made by AI are less biased than decisions made by humans is something that remains to be explored. However, as we continue to scrutinize AI, we must remain cognizant of the flaws of human decision-making as well. Just in 2021, the

---

115. *HireVue AI Explainability Statement*, *supra* note 8, at 3.

116. Moschella, *supra* note 106.

117. *HireVue AI Explainability Statement*, *supra* note 8, at 3.

118. *HireVue AI Explainability Statement*, *supra* note 8, at 4.

119. Glenn Gow, *How To Use AI To Eliminate Bias*, FORBES (July 17, 2022, 6:00 AM), https://www.forbes.com/sites/glenngow/2022/07/17/how-to-use-ai-to-eliminate-bias/?sh=4964ccc91f1f.

120. *See* Silberg et al., *Tackling bias in artificial intelligence (and in humans)*, MCKINSEY GLOBAL INSTITUTE (June 6, 2019), https://www.mckinsey.com/featured-insights/artificial-intelligence/tackling-bias-in-artificial-intelligence-and-in-humans.

121. *HireVue AI Explainability Statement*, *supra* note 8, at 27.

122. *HireVue AI Explainability Statement*, *supra* note 8, at 13, 27.

EEOC received over 22,000 disability discrimination charges,[123] over 20,000 race-based discrimination charges,[124] and over 18,000 sex-based discrimination charges.[125] Bear in mind that these numbers are simply the ones that were reported and do not indicate the true extent of employment discrimination that may occur in the U.S.

Lastly, some studies even show that job applicants have a direct and positive reaction to AI-enabled job application processes, indicating "intrinsic satisfaction" and an increase in "feelings of being innovative."[126] While no conclusion can be made as to whether job applicants prefer traditional, human-based job interviews over AI-enabled video interviews, these studies do demonstrate that at least some population of applicants might accept or prefer AI-enabled hiring processes in the future.[127]

---

123. *Americans with Disabilities Act of 1990 (ADA) Charges (Charges filed with EEOC) (includes concurrent charges with Title VII, ADEA, EPA, and GINA) FY 1997- FY 2021*, EEOC, https://www.eeoc.gov/data/americans-disabilities-act-1990-ada-charges (last visited Nov. 24, 2022).

124. *Id.*

125. *Sex-Based Charges (Charges filed with EEOC) FY 1997 - FY 2021*, EEOC, https://www.eeoc.gov/data/sex-based-charges-charges-filed-eeoc-fy-1997-fy-2021_(last visited Nov. 24, 2022).

126. Esch et al., *Job candidates' reactions to AI-Enabled job application processes*, 1 AI AND ETHICS 119, 121, 124 (2021), https://link.springer.com/article/10.1007/s43681-020-00025-0#Sec12.

127. *See generally* Cynthia Rudin & Joanna Radin, *Why Are We Using Black Box Models in AI When We Don't Need To? A Lesson From an Explainable AI Competition*, HARV. DATA SCIENCE REV. (Nov. 22, 2019), https://hdsr.mitpress.mit.edu/pub/f9kuryi8/release/8 (Although this is not a hiring decision: "[T]he audience—consisting of power players in the realms of finance, robotics, and machine learning—were asked to engage in a thought experiment where they had cancer and needed surgery to remove a tumor. Two images were displayed on the screen. One image depicted a human surgeon, who could explain anything about the surgery, but had a 15% chance of causing death during the surgery. The other image showed a robotic arm that could perform the surgery with only a 2% chance of failure. . . In this scenario, total trust in [AI] was required; . . . and no specific understanding of how it came to its decisions would be provided. The audience was then asked to raise a hand to vote for which of the two they would prefer to perform life-saving surgery. All but one hand voted for the robot.").

VI. A SOLUTION TO REGULATING AI TOOLS IN HIRING: A MODEL ACT

A. Why Have a Model Act

Model Acts are suggested examples of laws and are drafted to be disseminated and enacted across state legislatures.[128] Traditionally, model acts have provided uniformity across jurisdictions in a specific body of law.[129] State legislatures can reject them, enact them in their entirety, or enact them with modifications.[130] Some widely known examples include:

- ABA's Model Business Corporation Act – Adopted (with modifications) by 30 states[131]
- ALI's Model Penal Code – Adopted (with modifications) by 37 states[132]
- NNA's Model Notary Act – Adopted (with modifications) by 40 states[133]
- Uniform Commercial Code – Adopted (with modifications) by all 50 states[134]

The groups who write and promote Model Acts are made up of practicing lawyers, judges, legislators, and law professors—essentially those who are experts in the field.[135] The groups also tend to be made up of representatives from different states and can be appointed by state governments, which reflect the diverse experience of states.[136]

Model Act drafters like the Uniform Law Commission provide that if "there is substantial reason to anticipate enactment in a

---

128. *What is a Model Act?*, UNIFORM LAW COMMISSION, https://www.uniformlaws.org/acts/overview/modelacts (last visited Nov. 25, 2022).

129. *Id.*

130. Deanna Barmakian, *Uniform Laws and Model Acts*, HARVARD LAW SCHOOL LIBRARY (Jan. 20, 2023), https://guides.library.harvard.edu/law/unifmodelacts ("Model Acts are generally used as a basis for designing state laws. They are rarely enacted in entirety.").

131. *Model Business Corporation Act Annotated,Fifth Edition*, AMERICAN BAR ASS'N, https://www.americanbar.org/products/inv/book/402676454/ (last visited Nov. 25, 2022).

132. Model Penal Code (MPC) - Criminal Law Basics, UPCOUNSEL, https://www.upcounsel.com/lectl-model-penal-code-mpc-criminal-law-basics (last visited March 1, 2023); *see generally* MARKUS DUBBER, AN INTRODUCTION TO THE MODEL PENAL CODE 5–6 (2d ed. 2015).

133. NATIONAL NOTARY ASSOCIATION, MODEL NOTARY ACT ADOPTIONS (2010), https://www.nationalnotary.org/file%20library/nna/reference-library/us_jurisdictions_-adopting_model_notary_act.pdf.

134. *Uniform Commercial Code*, UNIFORM LAW COMMISSION, https://www.uniformlaws.org/acts/ucc (last visited Nov. 25, 2022).

135. *About Us*, UNIFORM LAW COMMISSION, https://my.uniformlaws.org/aboutulc/overview (last visited Nov. 25, 2022).

136. *Id.*

large number of jurisdictions", then "uniformity . . . among the various jurisdictions is a principal objective."[137] In other words, Model Acts can provide useful frameworks and enable states to remain consistent while also giving them the ability to prioritize or include certain features of a law that are important to a state's values.

Studies report that the global AI market size will likely grow to $422 billion by 2028.[138] Because of this projected growth, states should regulate this industry as they see fit. In Justice Brandeis's famous words, states "serve as a laborator[ies]" and allow for "novel social and economic experiments" to exist without risk to the rest of the country[139]—and here—without the risk of a grand sweeping bill like the AAA, which is already riddled with potential defects.

But simultaneously, AI developers and employers using AI tools in hiring should be given some form of consistency across jurisdictions.

The work-from-home movement spurred by COVID-19 has led to an even larger increase in workers living in states different from their employers.[140] Consistency across state laws governing AI used in employment and hiring will allow AI developers and employers to tailor their products and policies on a broad level. If states are provided with a model act, states will be able to adopt such a model act—and at the very least, states would share the same basic framework for regulation on AI tools used in employment. This way, AI developers and employers can be compliant across state lines without having to dedicate extra time and resources.[141] It would be a boon to the industry as a whole and would likely not hamper innovation and would dissuade AI developers from building models focusing only on or limited by specific

---

137. *What is a Uniform Act?*, UNIFORM LAW COMMISSION, https://www.uniformlaws.org/acts/overview/uniformacts (last visited Nov. 25, 2022).

138. Zion Market Research, *$422.37+ Billion Global Artificial Intelligence (AI) Market Size Likely to Grow at 39.4% CAGR During 2022-2028*, BLOOMBERG (June 27, 2022, 10:01 AM), https://www.bloomberg.com/press-releases/2022-06-27/-422-37-billion-global-artificial-intelligence-ai-market-size-likely-to-grow-at-39-4-cagr-during-2022-2028-industry.

139. New State Ice Co. v. Liebmann, 285 U.S. 262, 311 (1932).

140. *See generally* Parker et al., *COVID-19 Pandemic Continues To Reshape Work in America*, PEW RESEARCH CENTER (Feb. 16, 2022), https://www.pewresearch.org/social-trends/2022/02/16/covid-19-pandemic-continues-to-reshape-work-in-america/.

141. *See* Luca Bertuzzi, *Inside the EU's rocky path to regulate artificial intelligence*, INT'L ASS'N OF PRIVACY PROS. (Feb. 9, 2022), https://iapp.org/news/a/inside-the-eus-rocky-path-to-regulate-artificial-intelligence/ ("Adopting industry standards creates a presumption of conformity, minimizing the risk and costs for compliance. These incentives are so strong that harmonized standards tend to become universally adopted by industry practitioners, as the cost for departing from them become prohibitive.").

geographic regions. Thus, a Model Act—written by practicing lawyers, judges, and AI technologists—could provide states an avenue for AI regulation without the potential harm of a federal, overly-broad law drafted by Congress members who are not experts in the field.

Furthermore, AI software departs from traditional software in important ways. For example, a Software as a Service (SaaS) company's journey ends with sale and use of its software products, with newer versions, integrations, and updates along the way.[142] However, an AI software product's journey does not necessarily end at sale or use because the value derived from AI software is in its ability to provide useful predictions.[143] AI software also tends to have "rapid iteration" with algorithms and data sets constantly being tweaked or updated.[144] To illustrate: traditional software will likely have the same functionality once deployed unless humans update it. But for example, an AI system in a vehicle may stumble upon new data (such as a roundabout) and might discover a new approach to navigating roads.[145] AI systems are uniquely situated because they are essentially feedback loops, which are cycles without an end that are constantly changing.[146] An AI system continues to refine its recommendations based on the latest data that the system is fed.[147] Because of how rapidly evolving AI systems are, legislation in this field must remain flexible and open to innovation, and a model act would provide this malleability.

Representative Jamie Andrade, the drafter of Illinois' AIVIA, when asked about AI use stated that "we really don't quite understand [it] as legislators and just regular citizens",[148] which further confirms the need for a Model Act created by experts in the field. Since the Model Act would be drafted by qualified individuals, state legislators not familiar with the technology will feel more

---

142. *See generally What is SaaS customer journey mapping?*, TWILIO SEGMENT, https://segment.com/growth-center/customer-journey/customer-journey-saas/ (last visited Nov. 25, 2022).

143. Eric Lofgren, *How AI/ML margins differ from traditional software*, ACQUISITION TALK (Sept. 13, 2022), https://acquisitiontalk.com/2022/09/how-ai-ml-margins-differ-from-traditional-software/.

144. *See generally* Thanh (Bruce) Pham, *The Difference Between AI Software And Traditional Software Business*, SAIGON TECHNOLOGY, https://saigontechnology.com/blog/the-difference-between-ai-software-and-traditional-software-business (last visited Nov. 25, 2022).

145. *See* Dmitry Baraishuk, *What is Artificial Intelligence? AI vs Traditional Software*, BELITSOFT SOFTWARE (Feb. 12, 2021), https://belitsoft.com/ai-development/artificial-intelligence-vs-conventional-software.

146. *Id.*

147. *Id.*

148. Thibodeau, *supra* note 104.

confident and compelled in adopting such an act. As states would adopt this Model Act, different policies surrounding AI in hiring would be tested and applied without impacting each AI system in the country. This route may require AI developers to make multiple versions of their software. However, needless variations in policy would likely be limited since the base Model Act would have already been implemented as the foundation. In other words, even if AI developers are required to make multiple versions of their software to comply with different state laws, the differences between such state laws likely would not be considerable since they have the same Model Act foundation.

Arguably, since many concerns surrounding AI tools in hiring are—at their core—fears of employment discrimination, it may be appealing to have the EEOC handle arising issues of AI tools. The EEOC already has provided a guidance as detailed above, so it seems that this approach would be obvious and fitting.[149]

However, the EEOC, similar to the Congress members who proposed the AAA, may not be well equipped to handle these discrimination cases. Even though the U.S. workforce has grown by 50 percent since 1980,[150] Congress has kept the EEOC's funding nearly flat, resulting in more cases without the resources to handle them.[151] Just in 2021, the EEOC took in about 61,000 complaints, with about half the staff since 1980.[152] In addition, through a government survey in 2018, almost half of EEOC staff stated that they did not have the resources to do their jobs, higher than average for federal agencies.[153] While the October 2022 AI Training Act could help to remedy this problem, the AI Training Act itself does not indicate that any training provided to federal workers will be industry-specific, nor does it provide additional funding for the

149. EEOC, *supra* note 30.

150. Anna Brown, *Key findings about the American workforce and the changing job market*, Vox (Oct. 6, 2016), https://www.pewresearch.org/fact-tank/2016/10/06/key-findings-about-the-american-workforce-and-the-changing-job-market/.

151. *Charge Statistics (Charges filed with EEOC) FY 1997 Through FY 2021*, EEOC, https://www.eeoc.gov/data/charge-statistics-charges-filed-eeoc-fy-1997-through-fy-2021 (last visited Nov. 25, 2022); *U.S. Equal Employment Opportunity Commission Budget And Staffing History 1980 To Present*, EEOC, https://web.archive.org/web/20091209025619/http://archive.eeoc.gov/abouteeoc/plan/budgetandstaffing.html (last visited Nov. 25, 2022); *See generally* Maryam Jameel, *More and more workplace discrimination cases are being closed before they're even investigated*, VOX (June 14, 2019, 9:30 AM), https://www.vox.com/identities/2019/6/14/18663296/congress-eeoc-workplace-discrimination; Brown, *supra* note 148.

152. *Charge Statistics, supra* note 151.

153. *Federal Employee Viewpoint Survey 2018*, U.S. OFFICE OF PERSONNEL MGMT., https://www.opm.gov/fevs/reports/data-reports/data-reports/report-by-agency/2018/2018-agency-report-part-1.pdf (last visited Nov. 25, 2022); *See generally* Jameel, *supra* note 151

EEOC.[154] Rather, the act entails that the training will have "introductory concepts."[155] Given the complexity of employment discrimination through an AI system, the EEOC's large backlog of cases, and the EEOC's current lack of expertise, a Model Act drafted by lawyers and AI experts could serve as a better solution.

B. What Should be in it: The Model Act Should Move Away from an Explainability Requirement

Illinois' AIVIA, while not comprehensive, does contain aspects worth using and adjusting for in future regulations or a potential Model Act. For example, AIVIA's requirement for notice and disclosure to applicants in advance may provide job applicants the ability to choose whether they want to participate in a video interview AI system.[156] However, it does not specify how far in advance the notice must be provided or its format.[157] Similarly, AIVIA's requirement of "explaining how the artificial intelligence works and what general types of characteristics it uses to evaluate applicants" is a baseline starting point.[158] But it does not provide what the explanation should contain or what level of detail such an explanation should provide.[159]

Using the AIVIA as a framework for a potential Model Act, the explainability requirement for employers should stem from a local explainability standard. To provide a meaningful explanation to job applicants, employers must be able to justify decisions made by the AI system on an individual level.[160] Employers providing a global explanation may be evading the true intent behind regulations like the AIVIA, which as Representative Andrade describes as the desire for "simple transparency."[161]

But a better revision could include entail shifting focus away from an explainability requirement entirely. At the end of the day, how valuable can an explainability statement be for an end user? For example, if employers or companies like HireVue provide 30-page in-depth documents to job applicants before an interview,

---

154. AI Training Act, S.2551, 117th Cong. (2022).

155. AI Training Act, S.2551(b)(3)(B), 117th Cong. (2022).

156. Illinois Artificial Intelligence Video Interview Act, 820 ILL. COMP. STAT. 42/1 (2020).

157. *Id.*

158. *Id.*

159. *Id.*

160. Dhinakaran, *supra* note 62 ("Companies in regulated industries need to be able to justify decisions made by their model and prove that the decision was not made using features or derivatives of protected class information.").

161. Thibodeau, *supra* note 104.

would these documents provide something meaningful for a layperson?[162] To demonstrate, HireVue's explainability statement provides:[163]

> "The scores provided by our ridge regression models can be explained by looking at the input variables (known as 'features') and assessing their relative importance to generating the output score . . . [I]n addition to individual input features for each word used (Bag of Words feature sets) in the answer, our CAKE model calculates an embedding (i.e., list of numbers, or vector) to each answer. This embedding reflects the weightings that the neural network has attributed to the transcript. The CAKE embedding reflects 768 dimensions or features of the input sentence."

Would a layperson understand the complexity of AI models which AI experts themselves struggle to explain?[164]

Legislators and AI developers' emphasis on providing "transparency" for end users through explainability requirements parallels something every human being in modern-day society is plagued by: Terms and conditions of service agreements. Surveys over the years indicate that nearly 90% of consumers accept legal terms and conditions without reading them. For younger people (ages 18-34), the rate is even higher with 97% agreeing to conditions before reading.[165] Similarly, as companies continue to

---

162. Dwork et al., *supra* note 2 ("But disclosure of learning algorithms themselves has limited usefulness in the absence of data features with comprehensible meanings and explanations of weight determining the contribution of each feature to outcomes. Machine learning algorithms use mathematical operations to generate data features that almost always are not humanly understandable, even if disclosed, and whose learned combinations would do nothing to explain outcomes, even to expert auditors.").

163. *HireVue AI Explainability Statement*, *supra* note 8, at 13.

164. *See* Will Knight, *The Dark Secret at the Heart of AI*, MIT TECH. REV. (April 11, 2017) https://www.technologyreview.com/2017/04/11/5113/the-dark-secret-at-the-heart-of-ai/ ("The workings of any machine-learning technology are inherently more opaque, even to computer scientists, than a hand-coded system. This is not to say that all future AI techniques will be equally unknowable. But by its nature, deep learning is a particularly dark black box . . . You can't just look inside a deep neural network to see how it works. A network's reasoning is embedded in the behavior of thousands of simulated neurons, arranged into dozens or even hundreds of intricately interconnected layers.").

165. *See* Caroline Cakebread, *You're not alone, no one reads terms of service agreements*, BUSINESS INSIDER (Nov. 15, 207, 6:30 AM) https://www.businessinsider.com/deloitte-study-91-percent-agree-terms-of-service-without-reading-2017-11; *Click to agree with what? No one reads terms of service, studies confirm,* THE GUARDIAN, https://www.theguardian.com/technology/2017/mar/03/terms-of-service-online-contracts-fine-print (last visited Nov. 25, 2022) ("Hundreds of college students tapped the big green 'Join' button to become members of NameDrop, a new social network. But according to paragraph 2.3.1 of the terms of service, they'd agreed to give NameDrop their future first-born children . . . They were confirming, in the lab, what other scholars have found by painstakingly combing data on actual user behavior: nobody reads online contracts, license agreements, terms of service, privacy policies and other agreements.").

incorporate AI tools in their hiring protocols, there is likely a high chance that any explainability statements provided to job applicants will become another "click-to-agree" phenomena—where job applicants gloss over any documents provided (if they even open them at all).

Lastly, there are even studies that indicate that providing model explanations may lead to a false sense of trust, especially in faulty models. For example, in experiments led by Microsoft Research, individuals were asked to predict the price of a New York City apartment given eight features, ranging from the number of bedrooms and bathrooms to the distance to subways or schools.[166] Before making their predictions, participants were shown the rental price predicted by a model.[167] For some participants, the model's calculations and explanation were displayed. For other participants, they were not provided with any explanation of the model's predictions.[168]

When the model was applied to an unusual apartment (e.g., one with more bathrooms than bedrooms), the explained model participants were more likely to defer to the model's predictions than the no-explanation participants, despite the model's "sizeable mistakes" when evaluating unusual apartments.[169] This study indicated that people were more apt to accept and believe the output of an explained model, despite its glaring output mistakes, while the no-explanation participants remained more cautious.[170] Providing an explained model doesn't necessarily benefit end users in the long run if they end up falsely relying on the model due to an "explanation."[171]

Thus, while explainability may continue to be desired by legislators or advocacy groups, we must consider how truly helpful any explanation would be for job applicants.[172] It might still be worthwhile as a high-level requirement, but whether an AI system is deemed "good" or "bad" should not be rooted in its ability to be

---

166. *See* Poursabzi-Sangdeh at al., *Manipulating and Measuring Model Interpretability*, ARXIV (Aug. 15, 2021) https://arxiv.org/abs/1802.07810.

167. *Id.*

168. *Id.*

169. *Id.*

170. Katharine Miller, *Should AI Models Be Explainable? That depends.*, STANFORD INSTITUTE FOR HUMAN-CENTERED ARTIFICIAL INTELLIGENCE (March 16, 2021) https://hai.stanford.edu/news/should-ai-models-be-explainable-depends.

171. *Id.*

172. *Id.* ("If an AI model yields accurate predictions that help clinicians better treat their patients, then it may be useful even without a detailed explanation of how or why it works.").

explained to end users like job applicants, but perhaps rather in the quality of its output.[173]

C. What Should be in it: The Model Act Should be Limited in Scope

As highlighted before, a shortcoming of the AAA is that it is intended to apply to any type of algorithm or AI system used for "critical decisions"—a broad definition—regardless of industry. AI systems have numerous applications in countless realms, ranging from employment to automotive industries.[174] The ways in which AI tools aid users in such fields vary drastically— shouldn't the laws or regulations surrounding these tools be specific to those applications within those fields? With the AAA's framework, does it make sense to have the same standards for AI systems used by electricity providers and as for AI systems used by doctors or health practitioners? As it stands, the AAA would hold each of those AI systems in both scenarios to the same level of scrutiny or transparency, despite the obvious contrast in application.

One solution would be to develop individual Model Acts tailored to each type of "critical decision" category provided in the AAA, but this approach could be time-intensive and inefficient. An alternative solution could be to adopt the Risk Categories approach from the EU's draft law known as the Artificial Intelligence Act (AIA).[175]

The EU's AIA distinguishes three categories of AI uses: Prohibited AI uses, High-Risk AI uses, and Systems with Limited Risk.[176]

- The Prohibited AI use category refers to manipulation of persons through subliminal techniques beyond their consciousness or exploitation of specific vulnerable groups such as children or persons with disabilities in

---

173. Moschella, *supra* note 106.

174. *See* Louis Lehot, *United States: Artificial Intelligence Comparative Guide*, MONDAQ (Jan. 6, 2022) https://www.mondaq.com/unitedstates/technology/1059776/artificial-intelligence-comparative-guide.

175. *See* Benjamin Mueller, *The Artificial Intelligence Act: A Quick Explainer*, CENTER FOR DATA INNOVATION (May 4, 2021) https://datainnovation.org/2021/05/the-artificial-intelligence-act-a-quick-explainer/.

176. *See* Mauritz Kop, *EU Artificial Intelligence Act: The European Approach to AI*, STANFORD LAW SCHOOL (Oct. 1, 2021) https://law.stanford.edu/publications/eu-artificial-intelligence-act-the-european-approach-to-ai/.

order "to materially distort their behaviour" that can cause "psychological or physical harm."[177]

- The High-Risk AI category requires those systems to be carefully assessed by the European Artificial Intelligence Board (EAIB) before being put on the market and requires monitoring throughout the system's lifecycle.[178] Examples of this category as provided in the AIA include critical infrastructure that impacts citizen lives or health, law enforcement, employment or hiring, educational or exam scoring, etc.[179] This category also requires "human oversight" and is accompanied by a four-step process for market entrance.[180] Any systems that meet this category will also be required to be registered in a public EU-wide database.[181]
- The Limited Risk AI systems require "transparency obligations."[182]

While there are many critiques of this type of categorization and the EU's AIA as a whole,[183] the general framework it provides at the very least recognizes the need for different standards for different AI systems. Even the OSTP's *Blueprint for AI Bill of Rights* mentions "level[s] of risk based on context", indicating a

---

177. *Proposal For A Regulation Of The European Parliament And Of The Council Laying Down Harmonised Rules On Artificial Intelligence (Artificial Intelligence Act) And Amending Certain Union Legislative Acts*, EUR-LEX (April 21, 2021) https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206.

178. *Id.*; *see also* Kop, *supra* note 176.

179. *Proposal, supra* note 177; *see also* Kop, *supra* note 176.

180. *See* Kop, *supra* note 176 ("1. A High-Risk AI system is developed, preferably using internal ex ante AI Impact Assessments and Codes of Conduct overseen by inclusive, multidisciplinary teams. 2. The High-Risk AI system must undergo an approved conformity assessment and continuously comply with AI requirements as set forth in the EU AI Act, during its lifecycle. For certain systems an external notified body will be involved in the conformity assessment audit. This dynamic process ensures benchmarking, monitoring and validation. Moreover, in case of changes to the High-Risk AI system, step 2 has to be repeated. 3. Registration of the stand-alone Hi-Risk AI system will take place in a dedicated EU database. 4. A declaration of conformity must be signed and the Hi-Risk AI system must carry the CE marking (Conformité Européenne). Now the system is ready to enter the European markets.").

181. *Proposal, supra* note 177.

182. *Proposal, supra* note 177.

183. *See* Joanna Bryson, *Europe Is in Danger of Using the Wrong Definition of AI*, WIRED https://www.wired.com/story/artificial-intelligence-regulation-european-union/ (March 2, 2022, 9:00 AM) ("Particularly given that generative models are a special subclass of machine learning, . . . [the AIA's definition of Artificial Intelligence] makes it way too easy to argue at length in court for excluding a system from consideration by the act, and thus from the oversight and protection the AIA is intended to guarantee.").

desire for some level of risk-based categorization.[184] Thus, any Model Act should include tiers or categories by sector or by risk.

D. What Should be in it: The Model Act Should Contain Enforcement Procedures or Cause of Action

The AIVIA lacks guidance on enforcement, which has been pointed out by some scholars.[185] Without a specific cause of action, the bill does not effectively allow job applicants to sue for any damages the employer or AI developer may have caused through their AI system.[186] On the other hand, a notable feature of the AAA is that it does include a lengthy enforcement section, stating that enforcement would be led by the FTC and any violators of the act would be subject to FTC penalties.[187]

One step further is the EU's AIA, which includes forming and installing an entirely new enforcement body at the Union level: the European Artificial Intelligence Board (EAIB).[188] In a similar vein, any Model Act governing AI tools used in hiring should certainly include enforcement procedures or a cause of action and could go one step further by suggesting the creation of a dedicated enforcement body. Arguably, the issue of whether states would have enough resources or funding to support an enforcement body of this caliber would remain.

VII. CONCLUSION

AI is revolutionizing the way our world works in nearly every industry. It is natural and expected that regulations surrounding such technology are appearing. With the potential harm this type of technology can cause, it is important to create standards and oversight for this industry. However, regulation should promote

---

184. *Blueprint for an AI Bill of Rights*, *supra* note 5 ("An assessment should be done to determine the level of risk of the automated system. In settings where the consequences are high as determined by a risk assessment, or extensive oversight is expected (e.g., in criminal justice or some public sector settings), explanatory mechanisms should be built into the system design so that the system's full behavior can be explained in advance (i.e., only fully transparent models should be used), rather than as an after-the-decision interpretation. In other settings, the extent of explanation provided should be tailored to the risk level.").

185. *See* Brittany Kammerer, *Hired by a Robot: The Legal Implications of Artificial Intelligence Video Interviews and Advocating for Greater Protection of Job Applicants*, 107 IOWA L. REV. 817, 840 (2022) ("Additionally, the statute does not create a specific cause of action for applicants to bring claims against either the employer or vendor for improperly following the statute to the applicant's detriment or misusing the applicant's data.").

186. *Id.*

187. Algorithmic Accountability Act, H.R.6580, 117th Cong. (2022).

188. *See* Kop, *supra* note 176.

both innovation and conformity, and it should be done by experts and technologists in the field. Thus, a Model Act, drafted by attorneys and AI technologists, can provide an avenue for states to adopt AI regulations as they see fit, without imposing an overreaching federal law. This Model Act should combine components of draft regulations like the AAA and the EU's AIA in addition to current regulations like Illinois's AIVIA. The Model Act (1) should shift away from including an explainability standard, (2) should be limited in scope either by industry or by risk level, (3) and should contain procedures or protocols for enforcement.

The development and deployment of AI systems used in hiring are still at their beginning stages, and fear-mongering dialogue around this technology will only stifle its innovation and promote public distrust. As these systems continue to be improved, we must not hold this technology, its outcomes, and its impact to a higher standard than that we hold of human decisions. While many studies and analysts highlight the flaws within algorithms, these flaws should not be perceived in a way that dismisses the use or incorporation of this technology as a whole. AI systems certainly can be used in positive ways that promote equal protection and equal treatment for all, and there could even be a future in which these systems provide better outputs than human decisions—but only if they are given a chance to do so.

Mallika Dargan