Minnesota Journal of Law, Science & Technology

Volume 25 | Issue 1 Article 8

12-22-2023

Artificial Intelligence and the Administrative State: Regulating the Government Use of Decision-Making Technology

Gordon Unzen

Note

Artificial Intelligence and the Administrative State: Regulating the Government Use of Decision-Making Technology

*Gordon Unzen**

ABSTRACT

Artificial intelligence (AI) facilitates data-driven decision-making across all domains and government is no exception. AI is an imperfect technology, however, and will continue to be for the foreseeable future. Now, in the preliminary stages of AI deployment for societally consequential decision-making purposes, is the opportune moment to consider harm-mitigating regulations. This Note addresses regulating AI use in government decision-making, with a specific focus on administrative agencies. This Note recommends that Congress mandate the implementation of standardized technical and harm-based risk assessments for agency AI use. Additionally, it suggests that Congress implement increased public transparency and accountability measures and create an AI Agency with the legal authority to enforce agency compliance with best AI practices.

I. INTRODUCTION

At the dawn of the artificial intelligence (AI) era,[1] humanity is situated in an exciting yet treacherous position. With the aid

* J.D. Candidate, 2024, University of Minnesota Law School; B.A., 2021, University of New Hampshire. I would like to thank Professor Nick Smith for introducing me to the philosophy of AI, Professor Joan Howland and Chase Webber for their guidance during the writing process, and the MJLST team for their feedback and edits.

1. There are several indications that the advancement of AI is still in the beginning stages, not the end. Most tellingly, the core components of AI continue to improve without obvious signs of slowing down. AI development relies on (1) data relevant to a goal, (2) algorithms that can facilitate learning from the data, and (3) sufficient computing power to perform calculations necessary for creating the AI. Data, or information, is limitless for all practical purposes,

of data and algorithms, machines display an ever-increasing efficacy at tasks that were once exclusively within the realm of human intelligence, often surpassing human performance. AI technologies operate in the realm of predictions, recommendations, and decisions.[2] The mobilization of AI capabilities has enabled the automation of various tasks such as driving cars and drones,[3] conducting research,[4] producing art,[5] and writing essays.[6] These are just a few examples of the vast potential of AI. With the expanding competence of AI systems in a wide range of use cases, there is swelling public enthusiasm for enhancing their cognitive power, applying the technology to new tasks, and expanding their deployment in real-world applications.[7] The mounting prevalence and success of AI has led to its widespread acceptance not only by data scientists and

---

and the only restriction is collecting it. Present algorithmic practices are robust and likely have the capacity to facilitate the growth of AI for the conceivable future. Computing power, as a hardware limitation, is most likely to create roadblocks in AI development. However, Moore's law has yet to be violated and work on Quantum Computing will likely facilitate additional leaps and bounds in the potential speed of information processing. *See* David Brown, *Moore's Law vs. Quantum Computing: Is it Comparing Apples and Oranges?* ELEC. PRODS. (Oct. 18, 2021), https://www.electronicproducts.com/moores-law-vs-quantum-computing-is-it-comparing-apples-and-oranges/.

2. Matt O'Shaughnessy, *One of the Biggest Problems in Regulating AI is Agreeing on a Definition*, CARNEGIE (Oct. 6, 2022), https://carnegieendowment.org/2022/10/06/one-of-biggest-problems-in-regulating-ai-is-agreeing-on-definition-pub-88100.

3. *Autonomous Vehicles and Drones*, CETMO, https://www.cetmo.org/autonomous-vehicles-drones-transport-logistics/ (last visited Nov. 21, 2023).

4. *Is Artificial Intelligence Good or Bad for Academic Research?* ENAGO ACAD. (Oct. 10, 2023), https://www.enago.com/academy/academic-publishing-machine-learning-era/.

5. There is some contention as to whether AI produced "art" can be categorized as novel art. *See* Kevin Roose, *An A.I.-Generated Picture Won an Art Prize. Artists Aren't Happy*, N.Y. TIMES (Sept. 2, 2022), https://www.nytimes.com/2022/09/02/technology/ai-artificial-intelligence-artists.html.

6. Jonathan Vanian, *Why Tech Insiders Are So Excited About ChatGPT, a Chatbot that Answers Questions and Writes Essays*, CNBC (Dec. 13, 2022, 1:52 PM), https://www.cnbc.com/2022/12/13/chatgpt-is-a-new-ai-chatbot-that-can-answer-questions-and-write-essays.html.

7. *See* Ron Schmelzer, *The Increasing Expansion of AI in Business and Government–Insights from AI World*, FORBES (Mar. 22, 2019, 5:16 PM), https://www.forbes.com/sites/cognitiveworld/2019/03/22/the-increasing-expansion-of-ai-in-business-and-government-insights-from-ai-world/?sh=2762eda15def.

enthusiasts, but also among the general population. AI is now viewed as a legitimate and valuable tool that can support work, automate daily tasks, and even inform decision-making.[8] In light of the expected, and perhaps inevitable,[9] creep of AI power and influence over time, it is imperative for humans to learn in these early days how to coexist with this new type of intelligence.

The initial achievements of AI may lead one to assume that future advancements will only prove beneficial, or even believe that machines will become the perfect decision-makers, guiding humanity to a better tomorrow. However, AI is a tool that, like any other, has flaws inherent to its design.[10] Complications emerge whenever information processing and analysis systems, whether AI or human, utilize poor data or are tasked with functioning in the extremely complex world. AI, like its human counterparts, is not infallible and introduces its own distinctive biases and imperfections into the process of decision-making.[11]

---

8. Sage Kelly et al., *What Factors Contribute to the Acceptance of Artificial Intelligence? A Systemic Review*, 77 TELEMATICS & INFORMATICS, 2023, at 10 (compiling studies about AI acceptance among different populations).

9. Inevitable is a strong word, and perhaps too strong, but there is a defense of its use here. The argument that AI will inevitably improve and expand influence takes the following form: (1) AI technology can continue without limit foregoing unforeseen technological or theoretical limits of information processing; (2) People are incentivized to continue attempts to improve the power and use of AI due to its effectiveness; (3) If there are no technological limitations to AI development, only the intervention of a regulatory body could limit expansion of the technology; (4) AI technology can be developed within a well-regulated environment assuming individuals still have access to data and computer processors; (5) a regulatory body cannot fully limit access to data and computer processors because of the internet and cloud computing; (6) even if a particular regulatory body could fully regulate AI in their jurisdiction, there is a low probability such regulations would have global influence; (7) therefore, someone, somewhere, will always continue improving AI technology and expanding its scope of application. This renders AI development a practically inevitable process.

10. Consider the analogy of a hammer. A hammer carries little meaning until someone picks it up. Then, the hammer becomes a component of effectuating some objective. This objective may involve the act of driving nails into a surface, but it could also potentially entail the act of taking someone's life. It is also important to note that the hammer may not be suitable for certain tasks, particularly those that require precision, such as making intricate cuts on a wooden surface. Similarly, AI is an extension of human objectives, and its design is optimized for specific applications over others, rendering it ineffective in some contexts.

11. *See* Abhishek Dabas, *Algorithmic Bias in Real-World*, MEDIUM (July 27, 2020), https://adabhishekdabas.medium.com/algorithmic-bias-in-real-

The effectiveness of AI, like any other tool, is also dependent on the actions of its human wielders. Therefore, the notion of AI as a self-sufficient entity rising beyond the influence of imperfect humans is flawed. The architecture of AI models, the underlying patterns in the data,[12] and the objectives for which AI is created are all products of human design. This incorporation of human biases and imperfections into AI technology is a significant factor that influences its operations. Prudence necessitates adopting a critical approach towards both AI and its human users, particularly where the technology informs real-world decision-making.

People should be most skeptically attentive to the implementation of AI in government decision-making, especially for governing institutions purporting to operate under the principles of democracy and its associated values of transparency, accountability, and fairness.[13] Government decision-making is already detrimentally influenced by human biases like any other human endeavor, but given the greater social stakes at hand, there is understandably a heightened desire to substitute in the seemingly objective AI decision-makers. However, the use of AI in government likewise has the potential to negatively impact the public interest. Government actors might conceal biased polices in the opaque and uninterpretable mathematics of AI models. AI decision-making could worsen systemic issues, discourage human discretion, and encourage the centralization of decision-making.[14] These possible side effects of AI use may undermine fairness in government decisions and decreased transparency, ultimately

---

world-b98808e01586 (explaining AI biases in a variety of contexts, including policing, sentencing, criminality, hiring programs, advertising, and healthcare).

12. *Id.* This is particularly true for predictive applications for which historic data is used to draw conclusions about individuals. For example, using AI to predict the risk of individuals for bail assessment purposes based on racist and classist policing practices led to a particular pattern of data collection that biases the algorithm from human decision-making.

13. *See* Murat Jashari & Islam Pepaj, *The Role of the Principle of Transparency and Accountability in Public Administration*, 10 AUDA 60, 61 (2018) (indicating that lack of transparency and accountability diminishes the value of democracy and rule of law).

14. *See e.g.*, Yuval Noah Harari, *Why Technology Favors Tyranny*, ATLANTIC (Sept. 4, 2018), https://www.theatlantic.com/magazine/archive/2018/10/yuval-noah-harari-technology-tyranny/568330/ (explaining the theoretical relationship between AI and dictatorships).

limiting the potential for accountability. AI poses a plausible threat to democracy, a threat that will increase as the power and influence of AI continues to expand.

The mere possibility for AI to threaten democratic governance should be a call to action. It is time to consider how the United States should employ regulatory mechanisms to restrict the use of AI by governing institutions. Regulations should create room for the technology to enhance government efficiency, while simultaneously curbing its potentially undemocratic features. This early stage of the AI revolution presents an opportune moment to implement experimental restrictions on its use in government, assess their efficacy, and modify accordingly. The objective is to develop a versatile framework that can effectively address the challenges posed by the disruptive and unpredictable future of AI governance.[15] Successful regulatory tools implemented today may need to evolve tomorrow considering the many unknowns regarding the possible status of AI power in the coming decade. Given the rapid pace of AI development, it is imperative not to delay the development of regulatory measures until after the potential risks are fully understood.[16] By that time, it will likely be too late to effectively address the harms posed by AI uses in government.

ChatGPT, an AI system specialized in natural language processing, shared its own thoughts on whether government use of AI should be regulated, and eloquently summarized the issue:

> There are valid arguments for both regulating and not regulating the government's use of AI. Some argue that regulation is necessary to ensure that the government's use of AI is transparent, fair, and accountable to the public. Additionally, regulation could help to prevent potential abuses of power and ensure that the government's use of AI is aligned with democratic values. On the other hand, some argue that too much regulation could impede the government's ability to use AI effectively to improve public services, such as healthcare, education, and transportation. Additionally, it's also important to consider the legal and constitutional framework that may already constrain the government's use of AI and the potential implications of

---

15. *See generally* Steve Andriole, *AI: The Good, the Disruptive, and the Scary*, 31 CUTTER BUS. TECH. J. 6, 10 (2018) (advocating for further research in AI to address future issues like job displacement).

16. AI development could be outpacing Moore's Law. *See* Cliff Saran, *Stanford University Finds that AI is Outpacing Moore's Law*, COMPUTERWEEKLY (Dec. 12, 2019, 9:56 AM), https://www.computerweekly.com/news/252475371/Stanford-University-finds-that-AI-is-outpacing-Moores-Law.

> AI on the rights and freedoms of citizens. Therefore, it's important that the decision to regulate the government's use of AI is made after careful consideration and with input from a diverse group of stakeholders, including experts in AI, civil liberties and human rights, government, industry and academia.[17]

As noted by ChatGPT, determining the optimal relationship between government and AI requires a comprehensive and intricate discussion. Such a discussion implicates the involvement of private and public actors developing the technology, the vertical interplay between the layers of United States federalism, and the distinct considerations arising across government when integrating AI decision-making with the judicial, legislative, and executive branches. While these topics all warrant analysis, this Note focuses on the regulation of AI applications in federal administrative agencies. Future scholarship will expand this project to address broader questions related to algorithmic governance.[18] This could entail ensuring regulatory consistency across all governing structures within the United States or addressing these issues at the international level.

This Note aims to accomplish two primary objectives. The first is to formulate a compelling argument for regulating the use of AI by administrative agencies. The second is to analyze the ineffectiveness of existing regulatory structures and propose recommendations for Congress aimed at constraining agency AI use to safeguard democracy from the threat of AI. This Note builds most directly on the work of David F. Engstrom and his colleagues,[19] who have studied AI use cases in administrative agencies and developed a set of regulatory recommendations centered on maintaining human involvement in decision-

---

17. OpenAI, Response to "Should the Government Use of AI be Regulated?", CHATGPT (prompted Dec. 2022), https://chat.openai.com/.

18. Algorithmic governance is defined as the use of AI to support government research, decision-making, implementation, enforcement, and interaction. DAVID FREEMAN ENGSTROM ET AL., GOVERNMENT BY ALGORITHM: ARTIFICIAL INTELLIGENCE IN FEDERAL ADMINISTRATIVE AGENCIES 9 (Admin. Conf. U.S. 2020), https://www.acus.gov/sites/default/files/documents/Government%20by%20Algorithm.pdf [hereinafter GOVERNMENT BY ALGORITHM].

19. GOVERNMENT BY ALGORITHM *supra* note 18; David Freeman Engstrom & Daniel E. Ho, *Algorithmic Accountability in the Administrative State*, 37 YALE J. ON REGUL. 800 (2020).

making processes.[20] However, the analysis here adopts a different perspective by highlighting the significance of AI regulation for minimizing the potential threats to democracy and human rights. This approach aims to ensure that administrative agencies use AI in a responsible and ethical manner by arguing for the implementation of stronger oversight mechanisms both inside and outside of government.

Part II establishes the technical foundation for the analysis by introducing AI as a powerful decision-making technology. Section A presents this Note's operational definition of AI given its scope. Section B explores how machines make decisions. Section C highlights the limitations of AI and how they cause data bias and opaqueness issues that negatively impact AI decision-making.

Part III considers the implementation of AI decision-making in the operations of the administrative state. Section A presents the argument for the integration of AI in government decision-making processes, emphasizing the ability of AI models to exceed the physical and cognitive limitations of humans and enable more effective decision-making. Section B considers the potential dangers of integrating AI in government, with a particular focus on how AI could potentially undermine democratic values and interests. Section C introduces a risk-based methodology for categorizing AI implementations and employs this perspective to identify the risks associated with the current use of AI in federal administrative agencies.

Part IV analyzes the regulatory mechanisms best equipped to ensure that agency use of AI is transparent, accountable, and fair. Section A provides a framework for the establishment of effective AI regulations, emphasizing key principles such as the importance of considering AI substance in regulatory procedure, the potential risks associated with relying solely on the executive for regulation, and the need to prioritize caution over the rapid expansion of AI implementation. Section B considers the current government actions and regulatory mechanisms that could constrain irresponsible or dangerous uses of AI and examines their inadequacies. Section C argues that Congress should introduce new regulatory mechanisms targeting

---

20. A human is "in the loop" when they contribute or provide direct oversight to the AI decision-making process. *See* Arne Wolfewicz, *Human-in-the-Loop in Machine Learning: What is it and How Does it Work?*, LEVITY (Nov. 16, 2022), https://levity.ai/blog/human-in-the-loop.

administrative agency AI use. These new regulations must minimally require technical and harm-based AI risk assessments of AI uses, public disclosures and opportunities to hold agencies accountable for AI use, and legally enforceable interagency oversight with the creation of an independent AI Agency.

This Note concludes that regulating agency AI use is crucial for guaranteeing effective and safe algorithmic governance. The efficacy of regulatory mechanisms, however, no matter how comprehensive, will be compromised in the absence of a societal commitment to their enforcement. It is nevertheless crucial for the United States to act promptly by restricting the use of AI by agencies and mitigate the potential risks the technology poses to democratic principles and human rights.

II. ARTIFICIAL INTELLIGENCE: A POWERFUL DECISION-MAKING TECHNOLOGY

The notion of a "thinking machine" or artificial intelligence (AI) is not a recent concept. In fact, it can be traced back to at least the 1872 satirical novel *Erewhon*.[21] When the mathematician Alan Turing published his 1950 paper titled *Computing Machinery and Intelligence*, which established a logical framework for constructing and evaluating intelligent machines, the idea of AI had already occupied the minds of a generation of scientists and philosophers.[22] By the 1990s, AI had progressed significantly from a theoretical concept to a thriving technology. The advent of expert systems, which are programs that could solve complex problems by adhering to conditional logic decision paths derived from human expert knowledge, drove this progression.[23] Expert systems could detect cancer in patients, analyze molecular structures,[24] and achieve

---

21. Jeremy Norman, *In His Novel "Erewhon" Samuel Butler Describes Artificial Consciousness*, HIST. INFO., https://historyofinformation.com/detail.php?entryid=3850 (last visited Nov. 14, 2022) (noting *Erewhon* contemplates conscious self-replicating machines).

22. Rockwell Anyoha, *The History of Artificial Intelligence*, HARV. UNIV. GRADUATE SCH. ARTS & SCIS. (Aug. 28, 2017), https://sitn.hms.harvard.edu/flash/2017/history-artificial-intelligence/.

23. Michael Haenlein & Andreas Kaplan, *A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence*, 61 CAL. MGMT. REV., July 2019, at 4.

24. Aastha Aneja, *Expert Systems*, GEEKSFORGEEKS (June 16, 2023), https://www.geeksforgeeks.org/expert-systems/.

remarkable feats of intelligence such as defeating the then-world chess champion, Gary Kasparov, in 1997.[25] The present state of AI, and the reason people both love and fear the technology, however, stems from advancements to another approach to developing AI: machine learning (ML). ML requires substantial processing power,[26] extensive amounts of data,[27] and efficient "learning" algorithms.[28] Developments in information processing, data collection methods, and algorithm invention since the 1990s allowed ML to become a viable method for creating powerful AI systems.

Relying on the methods of ML, AI now exhibits superior performance to human intelligence in various tasks such as playing chess,[29] researching alternative physics,[30] and detecting

---

25. Haenlein & Kaplan, *supra* note 23 at 4.

26. The exploding popularity of video games encouraged the advancement of increasingly powerful graphical processing units (GPUs) that could facilitate the complex calculations needed for ML. Huw James, *How Gaming Has Aided GPU Rendering for Volume Visualization*, OFFSHORE ENG'R (Feb. 2, 2011), https://www.oedigital.com/news/445085-how-gaming-has-aided-gpu-rendering-for-volume-visualization ("The consumer-driven appetite for very large video game environments has been behind the push for performance improvements and price reductions for leading 3D graphics cards."); Andrew Brust, *NVIDIA Morphs from Graphics and Gaming to AI and Deep Learning*, ZDNET (Sept. 8, 2017, 3:50 PM), https://www.zdnet.com/article/nvidia-morphs-from-graphics-and-gaming-to-ai-and-deep-learning/ ("As it turns out, the kind of mathematical capabilities required to render high-resolution, high frame-rate graphics are also directly applicable to AI.").

27. The rise of the internet and the collection of information from its users led to "big data," huge sets of data that could be used to train ML algorithms. *Big Data and Artificial Intelligence: How They Work Together*, MARYVILLE UNIV. (July 21, 2017), https://online.maryville.edu/blog/big-data-is-too-big-without-ai/ (discussing how big data drives better AI while also encouraging the development of new AI techniques to analyze the massive data sets).

28. Advancements in mathematics and computer science brought about new "learning" algorithms as well as "boosting" algorithms that reduced bias and improved the efficacy of ML. Keith D. Foote, *A Brief History of Machine Learning*, DATAVERSITY (Dec. 3, 2021), https://www.dataversity.net/a-brief-history-of-machine-learning/.

29. The deep reinforcement learning ML algorithm AlphaZero learned to play chess in four hours and went on to uniformly beat Stockfish 8, an open-source chess engine designed through trial and error of the best human chess strategies. James Somers, *How the Artificial-Intelligence Program AlphaZero Mastered Its Games*, NEW YORKER (Dec. 28, 2018), https://www.newyorker.com/science/elements/how-the-artificial-intelligence-program-alphazero-mastered-its-games.

30. The AI was fed raw footage of physics phenomena and told to find the minimal set of fundamental variables describing the dynamics. The AI used different variables and different numbers of fundamental variables than

cancer.[31] AI also complements human intelligence by automating manufacturing,[32] driving cars,[33] and assisting in the development of new AI.[34] The power of AI is remarkable, disruptive, and potentially dangerous in a society structured to operate on the basis of human decision-making. This section introduces AI, including how it works and why it is an imperfect technology. Part A discusses the narrow definition of AI used in this Note: AI via ML. Part B provides an overview of how ML algorithms use mathematical operations to make probabilistic decisions. Part C examines the limits of statistical reasoning to explain why AI technology is imperfect.

A. THE OPERATIONAL DEFINITION FOR ARTIFICIAL INTELLIGENCE

The term AI refers to a broad range of concepts, encompassing a diverse array of definitions put forth by experts from various fields of study. The definition and conception of AI lacks a universally accepted standard for several reasons.[35] First, definitions of significant intelligence depend on reference to human intelligence given there is no higher intelligence form with which to compare. Second, human intelligence is often defined by its performance on distinct tasks such as learning, remembering, reasoning, abstracting, and adapting, which are difficult to reduce to a single definition. Third, AI is currently a

---

current physical laws. Columbia Univ. Sch. of Eng'g and Applied Sci., *Artificial Intelligence Discovers Alternative Physics*, SCITECHDAILY (July 27, 2022), https://scitechdaily.com/artificial-intelligence-discovers-alternative-physics/.

31. Nadia Jaber, *Can Artificial Intelligence Help See Cancer in New, and Better, Ways?* NAT'L CANCER INST. (Mar. 22, 2022), https://www.cancer.gov/news-events/cancer-currents-blog/2022/artificial-intelligence-cancer-imaging.

32. Madan Mohan Mewari & Gurudatta Kamath, *17 Remarkable Use Cases of AI in the Manufacturing Industry*, BIRLASOFT (July 1, 2021), https://www.birlasoft.com/articles/17-use-cases-of-ai-in-manufacturing.

33. Edwin Lisowski, *Artificial Intelligence in Self-Driving Cars*, ADDEPTO (July 16, 2021), https://addepto.com/blog/artificial-intelligence-in-self-driving-cars/ (noting that self-driving cars lead to a decrease in the number of accidents, at least in instances where a crash is caused by human error).

34. Anil Ananthaswamy, *Researchers Build AI that Builds AI*, QUANTA MAG. (Jan. 25, 2022), https://www.quantamagazine.org/researchers-build-ai-that-builds-ai-20220125/.

35. The following list is non-exhaustive. For further discussion regarding the difficulties of defining AI, *see* Matthew U. Scherer, *Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies*, 29 HARV. J.L. & TECH. 353, 359–62 (2016).

tool reflecting and supplementing human knowledge, which makes it dependent on the existence of other intelligences. Fourth, AI research involves both replicating human intelligence and striving to surpass it with new intelligence capabilities. Finally, there is no clear consensus on how to evaluate AI's intelligence, which can be approached through considering its capacity to think and engage in cognitive reasoning processes or behave through its ability to interact with an environment and exhibit appropriate behavior.

Defining AI poses a potential obstacle in discourse about its utility and the degree to which it should be subject to regulation. A vague or insufficient definition of AI may constrain the effectiveness of regulations aimed at governing its use. Additionally, an individual's perception of AI can impact one's perceived regulation priorities. Those trained in the philosophy of AI or futurism may take a more stringent regulatory stance on AI to address the potential risks associated with a technological singularity—a scenario where AI surpasses human intelligence at an exponential rate.[36] Economists may express greater concern regarding the displacement of human discretion, autonomy, and participation within the labor market.[37] Mathematicians, statisticians, and computer scientists may favor regulations ensuring sound data science practice but not restricting the use of AI.[38] Policymakers and ethicists may prioritize the regulation of specific AI applications over others, particularly those that have the potential to introduce biases affecting socially significant decision-making

---

36. One well-known public communicator of singularity concerns is the futurist philosopher Nick Bostrom. He frames the issue around the emergence of a super intelligent AI. *See* Nick Bostrom, *Ethical Issues in Advanced Artificial Intelligence*, NICK BOSTROM, https://nickbostrom.com/ethics/ai (last visited Nov. 21, 2023).

37. *See generally* Matthew Urwin, *Robots and AI Taking Over Jobs: What to Know About the Future of Jobs*, BUILT IN (Sept. 12, 2023), https://builtin.com/artificial-intelligence/ai-replacing-jobs-creating-jobs (explaining jobs that are more and less likely to be replaced by AI).

38. Machine Learning texts frame issues around data science limitations such as nonrepresentative training data, poor quality data, irrelevant features, and over/underfitting training data. *See* AIURELIEN GERON, HANDS-ON MACHINE LEARNING WITH SCIKIT-LEARN, KERAS, AND TENSORFLOW: CONCEPTS, TOOLS, AND TECHNIQUES TO BUILD INTELLIGENT SYSTEMS (Rachel Roumeliotis & Nicole Tache eds., 2nd ed. 2019).

processes.[39] Granted, these concerns are certainly not so segregated by field as displayed here. Nevertheless, the central problem remains: one's priorities for regulation are influenced by one's perception, definition, and understanding of AI. This underscores the importance of having a clear and comprehensive definition for AI before discussing regulation.

Given the broad range of definitions attributed to AI, this Note adopts a functional approach. The regulatory discussions that follow do not pertain to expert systems or potential future approaches to developing AI technologies. The AI at issue for this Note stems from the most effective set of techniques currently used to perform tasks that are typically associated with presumptively intelligent agents. This describes ML: the method by which computers learn without explicit programming.[40] While ML is a subcategory of AI in the sense that there exist multiple approaches to achieving AI and ML is only one of them, for applied discussions it is advisable to think of ML and AI as interchangeable.[41] ML is simply the current method to create the most "intelligent" AI systems. To illustrate, ML is to AI as a rocket ship is to possible methods of space travel. The rocket ship is used because it is the best available technology for space travel, not because it is the only theoretically possible method for traversing the cosmos. Likewise, ML is not the only way to achieve AI, nor is it necessarily the best possible way to achieve AI, but it is the best technology currently in use for making machines perform intelligent tasks. By focusing on methods, this Note will prioritize regulation for ensuring smart data science practices and ethical AI uses. The discussion of technological singularities will be left to the domain of philosophy, at least for the time

---

39. Olga Akselrod, *How Artificial Intelligence Can Deepen Racial and Economic Inequalities*, ACLU (July 13, 2021), https://www.aclu.org/news/privacy-technology/how-artificial-intelligence-can-deepen-racial-and-economic-inequities (noting that data bias can lead to discriminatory harm to people of color, women, and other marginalized groups).

40. *See* GERON, *supra* note 38, at 2.

41. Of course, future advancements in AI may eventually use non-ML methods. *See, e.g.*, Ron Schmelzer, *Going Beyond Machine Learning to Machine Reasoning*, FORBES (Jan. 9, 2020) https://www.forbes.com/sites/cognitiveworld/2020/01/09/going-beyond-machine-learning-to-machine-reasoning/?sh=27ffbcaf426b (arguing for a new machine reasoning design philosophy to further improve AI).

being.[42] The term AI will hereinafter be reserved for theoretical discussions of the concept of AI, although this is not a significant concern here. The term ML will be exclusively employed in the subsequent sections of Part II to convey that the techniques and limitations of current AI technologies are a direct result of ML methods. For discussions concerning ML-based AI technology in Parts III and IV, AI/ML will be used, serving as the general term for applied AI. AI/ML is the focus for regulation.

Like AI, ML carries several definitions, yet they all convey the same fundamental concept. Generally, ML is the field of study concerned with giving computers the ability to learn, and therefore acquire some notion of intelligence, without explicit programming.[43] The more technical definition states "[a] computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E."[44] Most simply, "Machine learning is the science (and art) of programming computers so they can *learn from data*."[45] The following section discusses the process by which computers learn from data and the potential applications of this learning.

B. HOW MACHINES MAKE DECISIONS

Machines learn from data using algorithms and statistical methods. This section presents a general overview of the types of ML models, systems, and methodologies. ML is often depicted as an enigmatic and complex field, and perhaps correctly so, but some understanding of the inner workings of the technology is necessary to demonstrate the capabilities and constraints of AI/ML decision-making.

---

42. Scenarios such as super intelligent AI and the complete supremacy of AI decision-making over society are plausible in the future and important problems to consider, but experts do not agree whether such scenarios are possible with existing technology, or any technology. *Compare* Bostrom, *supra* note 36 *with* Luke Dormehl, *Why AI Will Never Rule the World*, DIGITALTRENDS (Sept. 25, 2022), https://www.digitaltrends.com/computing/why-ai-will-never-rule-the-world/.

43. Without explicit programming is the crucial feature that separates ML from other AI methods such as the expert systems discussed previously. *See* GERON, *supra* note 38, at 2.

44. *Id.* at 2.

45. *Id.*

1. Introduction to Machine Learning

ML is perhaps best understood by contrast to traditional computer programming methods. Traditional programming is often analogized to following a baking recipe.[46] To make a computer perform a task, the programmer creates a series of instructions, like a recipe, that dictate the precise steps the computer should follow—what ingredients to include, in what order and proportions, and how long they should be baked.[47] The quality of the computer's output is dependent on the accuracy and comprehensiveness of the recipe. ML, on the other hand, is a computational approach that involves training the computer with data and using statistical analyses to generate a value representing the answer to a given problem.[48] The problem may call for a descriptive, predictive, or prescriptive response.[49] A given task may require a specific approach to training the computer, but the fundamental objective remains consistent: to enable the computer to create its own recipe through experiential learning.[50]

The technical term for this machine-constructed recipe is a ML model, defined as "a program that can find patterns or make decisions from a previously unseen dataset."[51] To build a model, the programmer must first have access to thousands and even millions of relevant data instances, whether in the form of numbers, photos, sounds, or text.[52] The significance of data in ML cannot be overstated. Even basic problems can require hundreds to millions of samples for effective training.[53] The programmer must collect the gathered data in a dataset and prepare it for analysis, which includes exploring the structures and content of the data, validating the data, and formatting the

---

46. Sarah Brown, *Machine Learning, Explained*, MIT SLOAN SCH. MGMT. (Apr. 21, 2021), https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained.

47. *Id.*

48. Lisa Tagliaferri, *An Introduction to Machine Learning*, DIGITAL OCEAN (May 31, 2022), https://www.digitalocean.com/community/tutorials/an-introduction-to-machine-learning.

49. Brown, *supra* note 46.

50. *Id.*

51. *Machine Learning Models,* DATABRICKS, https://www.databricks.com/glossary/machine-learning-models (last visited Nov. 21, 2023).

52. Brown, *supra* note 46.

53. GERON, *supra* note 38, at 23.

data.[54] They usually split the dataset into two sets: a set for training and a set for evaluation.[55] The programmer then chooses the ML algorithm for training. Selecting an appropriate algorithm depends on various factors, including the intended objective of the model, the desired output type, and the problem's scope.[56] The process may also require trial-and-error. Training methods can vary, but most ML algorithms use a blend of statistics, calculus, probability, and linear algebra to minimize some measurement of error and iteratively enhance accuracy.[57] After training the ML model, the programmer will assess its accuracy using new data from the evaluation set. If necessary, they may modify the model hyperparameters, variables, or the training algorithm to improve results.[58] The ML model can then be implemented in a ML system that incorporates the ongoing process of data collection, additional training, and updating the application that employs the model.[59]

2. Identifying Categories of Machine Learning Models

The classification of ML models and their corresponding training algorithms into categories is primarily based on the presence or absence of human supervision during the training process. Human supervision plays a role in ML when the training data contains the intended solution for a given problem. Frequently, however, ML models are used to address problems for which there is no pre-existing solution, thereby altering the

---

54. *See* Larysa Visengeriyeva et al., *Three Levels of ML Software*, MLOPS, https://ml-ops.org/content/three-levels-of-ml-software (last visited Nov. 21, 2023).

55. *Id.*

56. The field of study, time and resource limitations, and feature preferences are also important considerations. *See* Iryna Sydorenko, *How to Choose the Right Machine Learning Algorithm: A Pragmatic Approach*, LABEL YOUR DATA (May 3, 2021), https://labelyourdata.com/articles/how-to-choose-a-machine-learning-algorithm.

57. Ananya Chakaborty, *How to Learn Mathematics for Machine Learning? What Concepts Do You Need to Master in Data Science?,* ANALYTICS VIDHYA (Mar. 31, 2023), https://www.analyticsvidhya.com/blog/2021/06/how-to-learn-mathematics-for-machine-learning-what-concepts-do-you-need-to-master-in-data-science/.

58. Brown, *supra* note 46.

59. Alexander Reshytko, *Machine Learning Systems Versus Machine Learning Models*, MEDIUM (Aug. 29, 2022), https://towardsdatascience.com/machine-learning-systems-versus-machine-learning-models-3955d038ea1f.

learning approaches available for the ML model. The ML model learning categories include supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning. Models can also be categorized based on their data input method, either through discrete batches or continuously via an online connection.

In ***supervised learning***, the training data set includes desired solutions, called labels.[60] The term supervised is used because human intervention is required to ensure that the data is cleaned,[61] randomized, structured, and annotated with the appropriate labels.[62] During the training process of a supervised learning model, such as one designed to identify a picture of a dog, the model provides an answer, verifies the accuracy of the answer against the label, and subsequently adjusts its parameters in response to inaccuracies. Different supervised ML algorithms are appropriate for specific tasks, such as *regression* for predicting a target value[63] or *classification* for determining the group to which a data instance belongs.[64] This highlights the importance of selecting the appropriate algorithm to achieve accurate results.

Unsupervised learning can be considered the antithesis of supervised learning because it works with unlabeled, unstructured, and unprocessed data, thereby eliminating the possibility of comparing the results with correct answers.[65] This can limit the tasks that unsupervised models can perform. For example, a *clustering algorithm* can detect different object groups based on their distinctive characteristics, such as differentiating between an apple and a car.[66] However, it lacks the ability to recognize the specific identity of each group and

---

60. Labels are the answers as depicted in the real world. Supervised learning works to shape themselves to predicting the provided answers. *See* GERON, *supra* note 38, at 7–8.

61. Cleaning refers to the process of preparing data for analysis. This can involve combining variables into a single measure, changing the data type, or removing poor data, for example.

62. Sydorenko, *supra* note 56.

63. GERON, *supra* note 38, at 8.

64. *Id.* Geron uses the example of a spam filter in an email inbox to describe classification. By providing a model with examples of spam and non-spam emails, the model will learn to identify a new email it has not been trained on as either spam or non-spam.

65. Sydorenko, *supra* note 56.

66. *Id.*

can only acknowledge their dissimilarity. The absence of supervision, despite its potential drawbacks, does permit unique capabilities. For example, a *visualization algorithm* can take in complex and unstructured data to return a two-dimensional or three-dimensional representation that facilitates the identification of patterns that may not have been initially anticipated by a human.[67]

Semi-supervised models can be advantageous where the available data is partially labeled.[68] The typical semi-supervised learning process is to train the model as if it were unsupervised and then fine-tune the functionality with supervised learning.[69] This process leverages both labeled and unlabeled data to improve the model's performance. For example, photo-hosting services often employ unsupervised learning to identify images featuring identical faces and cluster them together without relying on explicit labels.[70] If the user opts to add labels, in this example the name of the individual to whom the face belongs, then the model can classify new images based on both the facial features and the associated label, effectively simulating a supervised learning approach.[71]

Reinforcement learning is a distinctive variant of ML model because it involves an agent that assumes the role of an observer in an environment.[72] The agent selects actions to perform and receives rewards or penalties from the environment based on the chosen action. The agent's goal is to formulate a policy or strategy that maximizes rewards while minimizing penalties.[73] For example, the reinforcement learning model AlphaZero created a chess policy that surpassed human capabilities in the game.[74] The model analyzed chess games to

---

67. Unsupervised learning can do a variety of tasks beyond the scope of this Note. *See* GERON, *supra* note 38, at 11–13 (describing other types of unsupervised learning algorithms).

68. *Id.* at 13.

69. *Id.* at 13 (using the example of Google Photos as a semi-supervised model).

70. *Id.*

71. *Id.*

72. *Id.* at 14.

73. *Id.*

74. Maxim Khovanskiy, *AlphaZero Chess: How it Works, What Sets it Apart, and What it Can Tell Us*, MEDIUM (May 5, 2022), https://towardsdatascience.com/alphazero-chess-how-it-works-*what*-sets-it-apart-and-what-it-can-tell-us-4ab3d2d08867.

learn the game's rules, then through trial-and-error it optimized its ability to win. It is noteworthy that AlphaZero not only outperforms human players in the game of chess, but it also uses a strategy that diverges from the approach developed by humans over the past thousand years.[75]

ML systems can also be categorized by how they receive data for further learning. The two categories are batch and online learning. In ***batch learning***, the model for the system is trained offline using all available data before deployment.[76] To incorporate new data into a batch learning system, the model must be retrained with a dataset including both the old and new data.[77] Batch learning models exhibit stability but are nonadaptive to changes in data without additional training. In ***online learning***, the system continues the learning process after deployment by feeding the model data in batches.[78] Online learning systems can be more adaptable and use fewer resources, but they also tend to degrade over time from the accumulation of erroneous data.[79]

3. Deep Learning

The final topic for this overview pertains to neural networks and deep learning. An ***artificial neural network*** (ANN) is a special type of ML model that was designed as an analogy to biological neuron networks in the brain.[80] The components of an ANN include the simulated "neurons," the synapses that connect them, and a series of layers that organize the neurons.[81] The simplest ANN model consists of three layers: the input layer composed of input neurons, the output layer using output neurons to represent the final result, and the hidden layer, composed of more neurons, in between.[82] The input neurons

---

75. *Id.* ("[S]ince AlphaZero does not make use of any human knowledge, unlike traditional engines (which use not only human-built heuristics, but also opening books and sometimes endgame tablebases), we can expect it to come up with brand-new ideas previously unknown to mankind.").

76. GERON, *supra* note 38, at 15.

77. *Id.*

78. *Id.* at 15.

79. *Id.* at 17.

80. *Neural Networks and Deep Learning Explained*, WGU (Mar. 10, 2020), https://www.wgu.edu/blog/neural-networks-deep-learning-explained2003.html#close.

81. *Id.*

82. *Id.*

receive the data and then pass them onto the hidden layer neurons via the synapses.[83] The hidden layer performs calculations, and then passes the result onto the output layer which will activate a neuron corresponding to the result.[84] For an ANN trained to recognize a number in an image, there is an input neuron for each pixel of the image and an output neuron for each number the image could represent. Importantly, like the real brain, ANN neurons only fire at the behest of an activation function, which determines whether a given neuron will send information to the next layer.[85] The learning occurs through backpropagation, which is the process of adjusting the calculations applied at each neuron to improve accuracy.[86] ANNs can be used for many of the ML tasks discussed above such as classification, clustering, and prediction.[87]

Deep learning models follow the ANN structure but contain up to hundreds of hidden layers, thus deep.[88] These additional layers give the model increased power and efficiency, but they come with considerable tradeoffs. The primary challenge is that the complexity of the hidden layers poses a significant obstacle for human comprehension of the model's decision-making process.[89] One can theoretically calculate the decision output step-by-step, but this approach is not practically feasible and provides limited insights into the critical questions about a decision-making process. The basis for a decision lives in the complex mathematics of the behavior of thousands of simulated neurons.[90] Deep ANNs are an opaque black box that perform extraordinary, yet unexplainable, feats of intelligence.[91]

Although this overview only provides a cursory understanding of the complexities of ML, the information presented herein should provide the requisite foundation for the remainder of the Note. The following section explores how the

---

83. *Id.*
84. *Id.*
85. *Id.*
86. *Id.*
87. *Id.*
88. *Id.*
89. Valeryia Shchutskaya, *Deep Learning: Strengths and Challenges*, INDATA LABS (July 27, 2021), https://indatalabs.com/blog/deep-learning-strengths-challenges.
90. *Id.*
91. *Id.*

mechanics of ML models can result in consequential flaws that negatively impact their decision-making capabilities.

C. WHERE ALGORITHMIC DECISION-MAKING GOES WRONG

ML models possess significant computational capabilities, yet they are accompanied by a multitude of problems. First, ML is an exercise in data science rooted in mathematics. Both data science and mathematical limitations inhibit the efficacy of ML decision-making, and at worst, incorporate human biases into the decisions. Second, the most intelligent ML models, deep ANNs, exhibit the least transparency. The trustworthiness of their decisions is questionable when human beings cannot comprehend them. Third, ML is not suitable for every task.[92] The excitement surrounding the technology may result in ML decision-making applications that are unsuitable or unethical for the given problem. Finally, the ML model may perform well in a controlled environment, but the effectiveness may be difficult to replicate when implemented in real-world scenarios. Without sufficient trial-testing, ML models may cause large-scale damage before people notice the errors.

ML issues occur because of the training data, the chosen learning algorithm, or the validation technique. ML requires a large quantity of data and without sufficient information, a ML model will not yield accurate results.[93] For a model to generalize to new data, the training data must also be representative of the future.[94] The presence of sampling bias, historic data bias, and underrepresented data can result in the development of inaccurate models, which may perpetuate discriminatory practices based on race or gender.[95] The effectiveness of a ML

---

92. Sakshi Gupta, *When Should You Not Use Machine Learning?*, SPRINGBOARD (Sept. 25, 2020), https://www.springboard.com/blog/data-science/when-not-to-use-ml/ (explaining scenarios where ML is not appropriate for a problem, such as when data are lacking or simple rules are sufficient).

93. This problem is one of the most fundamental in ML. Several researchers have shown that with sufficient data, very different learning algorithms performed quite similarly. Lack of data not only diminishes accuracy, but it also requires the careful selection of learning algorithms. GERON, *supra* note 38, at 24.

94. *Id.* at 25.

95. *Id.* at 26 (noting that data collection methods can make the resulting dataset unrepresentative of reality). Historical bias refers to the incorporation of bias in the real world into ML model decision-making because the data reflects the historic biases. Some examples are the gender wage gap and over-policing of minority communities. For example, facial recognition algorithms

model also depends on the quality of the data. If the dataset is contaminated with errors, outliers, random noise, and irrelevant features, the resulting model will reflect these issues.[96] The data can also be underfitted or overfitted by poorly chosen learning algorithms.[97] Finally, when evaluating the model, using an inappropriate performance measure can lead to misleading conclusions about the accuracy.[98]

ML models can make untrustworthy or unjustifiable decisions. Deep learning ANNs are the most effective ML models at finding patterns, making predictions, and surpassing human intelligence.[99] However, these amazing capabilities cause an opaqueness limitation. Deep ANN models find patterns and correlations overlooked by human experts, arrive at conclusions that even ML engineers cannot understand, and therefore, can make consequential decisions that lack explanation.[100] If a financial institution denies a loan application based on the output of a deep ANN, the institution would be unable to clarify the decision beyond "the AI said so."

---

worked better for white faces because that was the primary source of the data. Representation matters. *See* Mary Reagan, *Understanding Bias and Fairness in AI Systems*, TOWARDS DATA SCI. (Mar. 24, 2021), https://towardsdatascience.com/understanding-bias-and-fairness-in-ai-systems-6f7fbfe267f3.

96. As the adage goes, "garbage in, garbage out." Poor quality data can be rectified with effective data cleaning and the most relevant features can be found with feature engineering. GERON, *supra* note 38, at 27; Visengeriyeva et al., *supra* note 54.

97. Overfitting occurs when an overly complex model is used with too many degrees of freedom so that the model fits the training data very well but does not generalize to new data. The solution is to use a training algorithm that simplifies the model, such as turning a polynomial model into a linear model. Underfitting is the opposite issue where the model is too simple to adequately learn the underlying structures of the data. A more powerful model with more parameters and fewer hyperparameter constraints will resolve this issue. GERON, *supra* note 38, at 27–29.

98. *Id.* at 89, 90.

99. Ben Dickson, *The Limits and Challenges of Deep Learning*, TECHTALKS (Feb. 27, 2018), https://bdtechtalks.com/2018/02/27/limits-challenges-deep-learning-gary-marcus/.

100. *Id.* All current ML models are inherently narrow: they can perform a given task very well but cannot generalize their intelligence as expected by general intelligence systems like humans. Deep learning is also shallow, meaning the models do not understand the context of the data they process. *See generally* Schmelzer, *supra* note 41. These attributes of ML will limit the ability to ingrain abstract ethical principles in the models. Deep learning has a way to go before it becomes a fully mature technology.

ML can be overused. ML is most effective for solving problems when existing solutions require lengthy rule-lists, lack traditional solutions, occur in dynamic environments,[101] and involve vast amounts of data.[102] Overenthusiasm about AI/ML systems can lead to model deployment for problems lacking sufficient data or for which traditional programming methods present more effective solutions, causing inefficient or inaccurate decision-making. This issue also arises when someone opts to use a deep ANN rather than a simpler ML model such as a decision tree for decision-making that requires clear explanation.[103]

Finally, the real world is a difficult environment for evaluating certain types of ML models. In the context of statistical hypothesis testing, there are two types of errors: type I and type II.[104] Type I errors are false positives: the inaccurate finding of statistical significance.[105] A type II error is a false negative: the inaccurate finding of a lack of statistical significance.[106] ML models, as an exercise in statistics, need to engage with both types of errors. However, real-world deployment can limit the ability to account for them. For example, consider a ML model trained to forecast the risk of recidivism for criminal offenders and determine whether the offender should be detained pre-trial.[107] If the ML model predicts that an offender presents a threat to the public or a recidivism risk, the offender will likely be jailed. It is not feasible to verify the accuracy of this prediction because the offender would not be

---

101. Note, however, that a fluctuating environment can be a detriment. Data can grow stale and batch learning systems may have difficulties keeping up with changes over time. *See* Michael Segner, *Stale Data Explained: Why it Kills Data-Driven Organizations*, MONTE CARLO (Mar. 28, 2023), https://www.montecarlodata.com/blog-stale-data/.

102. GERON, *supra* note 38, at 5.

103. Decision trees work like a flow chart, so each step and fork through the decision-making process is traceable. *Decision Tree*, GEEKSFORGEEKS (last updated Aug. 20, 2023), https://www.geeksforgeeks.org/decision-tree/.

104. Saul McLeod, *Type 1 and Type 2 Errors in Statistics*, SIMPLYPSYCHOLOGY (Oct. 5, 2023), https://www.simplypsychology.org/type_I_and_type_II_errors.html.

105. *Id.*

106. *Id.*

107. Noel L. Hillman, *The Use of Artificial Intelligence in Gauging the Risk of Recidivism*, AM. BAR ASS'N: JUDGES' J. (Jan. 1, 2019), https://www.americanbar.org/groups/judicial/publications/judges_journal/2019/winter/the-use-artificial-intelligence-gauging-risk-recidivism/.

afforded the chance to reoffend. Thus, the ML model cannot be evaluated for type I errors. It would only exhibit type II errors when someone is released and recidivates. This creates a significant need for robust experimentation before deploying a consequential ML model, as real-world scenarios may not offer opportunities for evaluation.[108]

These problems, however, should not completely overshadow the benefits of ML in supplementing and improving upon human intelligence capabilities. This is why regulation is important: to mitigate the problems and ensure responsible use. Responsible use means complying with the best data science practices, trial-running before deployment to detect errors before they become consequential, and developing transparency and accountability measures for addressing errors when they do occur. Part III discusses AI/ML in the context of administrative agency decision-making, the regulatory environment that is the focus for the Note.

III. ALGORITHMIC GOVERNANCE: A NEW WAY TO MAKE GOVERNING DECISIONS

Algorithmic governance refers to the use of AI/ML to support government decision-making and action. Responsible algorithmic governance must reconcile two conflicting issues: (1) the need for AI/ML to improve government by limiting the influence of human biases in decision-making and (2) the need to mitigate the extent AI/ML will make government actions less transparent, accountable, and fair. This Part presents the advantages and disadvantages of implementing AI/ML in government decision-making in Sections A and B respectively, and, in Section C, offers a risk assessment survey of current uses of AI/ML in the administrative state. The following establishes the context for the discussion in Part IV, which pertains to regulations addressing concerns for AI/ML agency use.

---

108. There is also an important lesson to be learned about the scope of ML models here. Models may be very accurate at predicting group behavior, but not a given individual's behavior. A person will not necessarily recidivate because they exhibit factors that strongly correlate with or predict recidivism. *See generally id.* The most a ML model can conclude is simply that the person is more likely to recidivate. Probabilities should not be mistaken for definite conclusions.

A. THE NEED FOR AI IN GOVERNMENT

Despite the issues implicit to AI/ML technologies, there is a substantial need for data-driven governance in the increasingly complex world. Humans are themselves algorithmic agents, and sometimes their algorithms are not appropriate for finding optimal solutions to a problem. Humans often take cognitive shortcuts when making decisions because the brain cannot possibly process all available and potentially relevant information. These cognitive shortcuts were advantageous for survival in human's evolutionary history, and can still be useful today, but shortcuts also create biases.[109] For example, an individual who owns a thing tends to value it more than someone who does not, a bias known as the endowment effect.[110] Humans also dislike losses far more than gains, overestimate the predictability of past events, assume that the most readily available examples are the most important or prevalent, search for favorable information confirming existing beliefs while ignoring information that contradicts them, are susceptible to framing effects that change the evaluation of risk, make decisions based on anchored values, and use implicit racial and gender biases to make decisions.[111] Not all of the biases are necessarily irrational at the individual level because rapid decision-making is often a value in itself. However, when making decisions with national consequences impacting millions of people, the biases become dangerous errors. AI/ML, especially reinforcement models,[112] is not beholden to evolutionarily bound modes of reasoning and can thus avoid these biases.

Humans, as biological entities, also have physical limitations that AI/ML does not experience. Human working memory can handle approximately four variables at once, while AI/ML can work with hundreds of thousands of datapoints.[113] Human error greatly increases with fatigue, while AI/ML does

---

109. Alexander S. Gillis, *Cognitive Bias*, TECHTARGET (Apr. 2023), https://www.techtarget.com/searchenterpriseai/definition/cognitive-bias.

110. CARY COGLIANESE, ADMIN. CONF. U.S., A FRAMEWORK FOR GOVERNMENTAL USE OF MACHINE LEARNING 15 (Dec. 8, 2020), https://www.acus.gov/sites/default/files/documents/Coglianese%20ACUS%20Final%20Report.pdf.

111. *Id.* at 20 (discussing loss aversion, hindsight bias, availability bias, confirmation bias, framing, anchoring, and implicit racial and gender biases).

112. *See* discussion *supra* Part II.B.2.

113. COGLIANESE, *supra* note 110, at 10.

not tire.[114] Machines are always available for work because they do not sleep, eat, or take breaks. AI/ML does not age,[115] it is not impulsive, and does not exhibit the same types of perceptual inaccuracies as humans.[116]

Simply by avoiding human errors, AI/ML can outperform humans in problem-solving and decision-making capacities. Machines are not beholden to biology. AI/ML is not biased by the way information is articulated, or whether a group of authoritative people endorse an opinion.[117] AI/ML models can be trained to solve a simple problem and scaled to have a much broader impact on society. Additionally, AI/ML allows for novel ways to communicate to the public, determine resource priorities, and manage bureaucracy.[118] AI/ML decision-making can be bounded by restrictive parameters in the code, ensuring the AI/ML stays consistent with law.[119] This is particularly important for improving the operations of administrative agencies, where it is difficult to keep the thousands of humans in an agency within the mandates of the agency's authorizing statute. Effective governance will require political states to take advantage of these AI/ML benefits, if only to stay in competition with other governments that do.

B. HOW AI CAN BE UNDEMOCRATIC

Combining AI/ML and government is also a dangerous proposition that could cause undemocratic outcomes. Democratic decision-making requires transparency. Without access to sufficient relevant information, the public cannot make informed voting decisions, protest, or otherwise engage with the government to ensure accountability. A healthy democracy also requires plurality in decision-making. The more concentrated the exercise of power, the greater the risk that decisions unfairly

---

114. *Id.* at 11–12.

115. AI/ML models can, however, experience cognitive decline. *See* discussion *supra* Part II.B, C.

116. Human perceptual errors are based on information collection tradeoffs as evolution optimized humans to attend to certain informational sources rather than others. Such evolutionary prioritization of information is not exhibited in AI/ML systems, although poor data could lead to similar issues. COGLIANESE, *supra* note 110, at 13–14.

117. *Id.* at 22.

118. *See* discussion *infra* Part IV.C.

119. Cary Coglianese, *Administrative Law in the Automated State*, 150 DAEDALUS 104, 111 (July 1, 2021).

target minority groups, are erroneous, and are untouchable by counterbalancing tools of political power. AI decision-making also encourages power centralization as a means to increase the efficacy of models. Finally, AI/ML can be anti-democratic because it is opaque. These problems are of great consequence to regulating administrative agencies.

AI/ML's is opaque because its decisions are unexplainable, particularly when employing deep ANN models.[120] There is no transparency as to what happens under the hood, hiding potential decision errors or biases. This issue becomes more pronounced when a government uses the technology. When AI/ML creates rules, makes adjudicatory decisions, or enforces the law, the decisions implicit to these actions cannot be understood by the model engineers, much less the United States public or courts. If agencies use AI/ML for research or in other early stages of the decision-making process, the influence of AI/ML may be completely imperceptible. Bad-faith actors in government could rely on this opaqueness to develop models that appear effective on the surface but encode biased and unjust attitudes in their application. Once (and if) these issues are discovered, millions of people may already be disparately impacted by the algorithm's decision-making. Finally, even without the involvement of bad faith, if AI/ML decision-making appears effective on the surface, government decision-makers may continue to rely on the outputs and simply sign-off on the work product models produce.[121] This greatly exacerbates the previous problems because the removal of human discretion will further limit AI/ML oversight and the possibility of accountability.

AI/ML decision-making also encourages centralization. The futurist philosopher Yuval Harari argues that the difference between democracy and dictatorship boils down to a conflict between different data-processing systems—democracies represent a pluralistic spread of information processing and decision-making responsibilities, while dictatorships concentrate information and decision-making.[122] While perhaps reductionist, Harari's description of political systems in terms of data raises concerns when viewed in the context of AI/ML. As a

120. For more on deep ANNs, s*ee* discussion *supra* Part II.B.3.
121. GOVERNMENT BY ALGORITHM, *supra* note 18, at 11.
122. Harari, *supra* note 14.

tool for data-processing, AI may incentivize certain types of government structures, specifically dictatorships, because concentrating data will permit more information for training models and subsequently better algorithmic decision-making.[123] Whether this concern is relevant to the United States is unclear, given the institutional divides created by federalism and the three branches of government. However, within the administrative state, where information-sharing practices are more common and incentivized,[124] the process of power concentration could have a more pronounced effect.

These problems are not a sufficient reason to denounce algorithmic governance in its entirety. As discussed in the previous section, there are many reasons why AI/ML use in administrative agency actions can improve the administrative and greater political state. However, these risks do require attention and response, especially because the integration of AI/ML and government is no longer a hypothetical scenario. The next section surveys the current landscape of algorithmic governance and the risks it already imposes.

C. SURVEY OF RISK IN ADMINISTRATIVE AGENCY AI USE CASES

The era of algorithmic governance, marked by the use of AI/ML to support government research, decision-making, implementation, enforcement, and interaction[125] is underway.[126] According to a report from the 2020 Administrative Conference of the United States (ACUS), 64 of the 142 federal departments, agencies, and subagencies surveyed were at least experimenting with AI/ML technology.[127] AI/ML models primarily performed classification and regression tasks using structured and textual

---

123. *Id.*

124. For example, the Office of Information and Regulatory Affairs (OIRA) facilitates interagency review of regulatory actions. *OMB Approval Process*, US DEP'T. DEFENSE, https://open.defense.gov/Regulatory-Program/Process/OMBApproval/ (last visited Nov. 26, 2023).

125. GOVERNMENT BY ALGORITHM, *supra* note 18, at 9.

126. *Id.* at 10. Attempts at using AI in government are not new. The current landscape of algorithmic governance comes from decades of government experiments with data mining, efforts to "reinvent government through data-based performance management and oversight" in the 1990s, and the creation of "expert systems" that relied on domain experts to craft logical rules to automate decision-making in the 1960s and 1970s.

127. *Id.* at 16.

data.[128] Agencies used these capabilities to enforce regulatory mandates, adjudicate government benefits and privileges, conduct research, monitor and analyze public health and safety risks, extract information from big government data, communicate with the public, and support internal management.[129]

This section surveys the various agency uses of AI/ML discussed in the ACUS report to emphasize the immediate need for regulation. One way to show this need is by categorizing current uses in risk categories. There are many frameworks for categorizing AI/ML risk,[130] but this Note will employ the system recommended by the European Union in their Artificial Intelligence Act proposal (AI Act) because its metrics generally correspond to the extent the AI/ML implementation poses a threat to democracy.[131] The EU's system is highly relevant for evaluating algorithmic governance in the administrative state.

The AI Act creates four categories of risk: minimal or no-risk, limited-risk, high-risk, and unacceptable risk.[132] A minimal/no-risk AI/ML use would be a spam filter or search tool.[133] A limited-risk AI/ML system would implicate transparency issues such as a chatbot, which a user may falsely

---

128. *Id.* at 19. For more on classification and regression, *see* discussion *supra* Part II.B.2.

129. GOVERNMENT BY ALGORITHM *supra* note 18, at 17.

130. One way of categorizing risk would be by the extent human discretion remains in the decision-making process. *See generally id.*

131. *Proposal for a Regulation of the European Parliament and the of Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts*, COM (2021) 206, final (Apr. 4, 2021). The European Union reached a provisional agreement on the AI Act on December 8, 2023, so ratification is likely. Clara Hainsdorf et al., *Dawn of the EU's AI Act: Political Agreement Reached on World's First Comprehensive Horizontal AI Regulation*, WHITE & CASE (Dec. 14, 2023), https://www.whitecase.com/insight-alert/dawn-eus-ai-act-political-agreement-reached-worlds-first-comprehensive-horizontal-ai. Given that negotiations have softened the categorization of AI systems as high-risk, this Note uses the standards as articulated in the original proposed AI Act. *See* Emilia David, *The EU AI Act Passed—Now Comes the Waiting*, THE VERGE (Dec. 14, 2023), https://www.theverge.com/2023/12/14/24001919/eu-ai-act-foundation-models-regulation-data.

132. *Regulatory Framework Proposal on Artificial Intelligence*, EUR. COMM'N, https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai (Nov. 15, 2023).

133. *Id.*

assume is a human.[134] A high-risk system would be an implementation pertaining to critical infrastructure, educational training, essential public services like loan or benefit services, law enforcement, or the administration of justice.[135] Finally, unacceptable risks are those that present a clear threat to safety and rights with explicitly anti-democratic uses like government-assigned social scores.[136] The following discussion categorizes the AI/ML uses uncovered in the ACUS study using the AI Act framework to convey the current level of public harm.

The good news is that many of the current AI/ML use cases in government agencies fall within the no-risk and limited-risk categories. Under no-risk are the uses that facilitate or speed up typical bureaucratic agency tasks. The Consumer Financial Protection Bureau uses AI/ML to analyze customer complaints, identify trends, and predict the consumer harm.[137] The Bureau of Labor Statistics uses similar systems to identify specific characteristics in worker injury narratives.[138] The Social Security Administration uses clustering models[139] in formal adjudication case processing to send substantively similar cases to specialized adjudicators.[140] The U.S. Patent and Trademark Office uses classification systems for informal adjudication to process applications for informal adjudication.[141] The office also has experimented with ANNs to facilitate patent prior art searches, automate the classification of trademarks, and retrieve trademarks in more accurate searches.[142] The U.S. Postal Service employs handwriting recognition tools to decipher illegible writing.[143]

Use cases are limited-risk, the second category from the EU framework, when AI/ML shifts agency priorities or facilitates novel action that would be impossible without the technology. The Social Security Administration uses a program to identify

---

134. *Id.*
135. *Id.*
136. *Id.*
137. GOVERNMENT BY ALGORITHM, *supra* note 18 at 10.
138. *Id.* at 10.
139. *See* discussion *supra* Part II.B.2.
140. GOVERNMENT BY ALGORITHM, *supra* note 18, at 39.
141. *Id.* at 48.
142. *Id.*
143. *Id.* at 10.

and accelerate appeals that are predicted to be successful.[144] They also use natural language processing systems to assist with writing hearing and Appeals Council decisions.[145] The Department of Health and Human Services used AI/ML to compare vendor products and services and develop acquisition strategies for spending procurement dollars.[146] The Department of Homeland Security deploys AI/ML to counter cyberattacks.[147] Limited-risk uses also occur when agencies use AI to interface with the public, who may not understand they are talking to an AI/ML system. The Department of Housing and Urban Development uses a chatbot for inquiries about rental assistance, agency programs, and civil rights complaint procedures.[148] The U.S. Citizenship and Immigration Services uses a chatbot for immigration questions.[149]

Concerningly, some current agency AI/ML use cases are high-risk because they could impact public rights and safety. This level of risk arises in rulemaking, enforcement, and infrastructural contexts. For example, the Food and Drug Administration uses AI/ML for preapproval studies to conduct post-market surveillance and risk assessment of adverse events and medication error reports.[150] The AI/ML outputs inform rulemaking and may prompt reevaluation of approval decisions.[151] The General Services Administration employs a system to ensure federal solicitors are legally compliant.[152] The Securities and Exchange Commission, Centers for Medicare and Medicaid Services, and Internal Revenue Service use AI/ML to predict potential violators of laws and regulations.[153] The Customs and Border Protection and Transportation Security Administration use facial recognition and risk detection systems to identify security threats like bombs, known criminals, and

144. *Id.* at 39.
145. *Id.* at 40.
146. GCN Staff, *How HHS Used AI to Become a Smarter Buyer*, GCN (Jan. 30, 2019), https://gcn.com/cloud-infrastructure/2019/01/how-hhs-used-ai-to-become-a-smarter-buyer/298313/.
147. GOVERNMENT BY ALGORITHM, *supra* note 18, at 10.
148. *Id.* at 16.
149. *Id.* at 17.
150. *Id.* at 53.
151. *Id.*
152. *Id.* at 10.
153. *Id.*

people at risk of becoming criminals or victims.[154] The Food Safety and Inspection Service uses prediction systems for deciding which sites to test for food safety.[155] The U.S. Postal Service is piloting the use of autonomous delivery vehicles and long-haul trucks, and exploring the use of unmanned aerial vehicles.[156]

Two final points are important to keep in mind: the varied sophistication of AI/ML and the limited public knowledge about agency AI/ML use.[157] Of the known agency AI/ML use cases, computer scientists identified approximately equal numbers of low, medium, and high sophistication implementations.[158] The high-risk uses discussed here could employ unsophisticated AI/ML that make more comprehendible decisions, thus lowering the risk. However, the ACUS Report authors note that the largest category in their sophistication level survey is "insufficient detail."[159] This highlights the public knowledge problem. While there are few high-risk and no reported unacceptable risk AI/ML use cases described in the ACUS report, this information is incomplete. Further, as the power of AI/ML increases, high-risk applications will only become more prevalent. AI/ML advocates imagine AI/ML systems replacing brick-and-mortar government centers, predicting social problems like housing and food insecurity to facilitate proactive government responses, leading military operations, engaging in pandemic and environmental control, monitoring and maintaining city infrastructure, managing prisons, and testing policy with simulations.[160]

---

154. *Id.* at 30.

155. *Id.* at 10.

156. *Id.* at 66.

157. *Id.* at 19–20; *see* DELOITTE AI INSTITUTE, THE AI DOSSIER 40 https://www2.deloitte.com/content/dam/Deloitte/us/Documents/deloitte-analytics/us-ai-institute-ai-dossier-full-report.pdf (last visited Dec. 2, 2022) (noting that AI use differs among government agencies).

158. GOVERNMENT BY ALGORITHM, *supra* note 18, at 20.

159. *Id.*

160. THE DELOITTE AI INSTITUTE, *supra* note 157, at 40–51 (discussing the many use cases for AI in government). *See* The Future of Artificial Intelligence (AI) in Government, INTEL, https://www.intel.com/content/www/us/en/government/artificial-intelligence.html (last visited Nov. 26, 2023) (outlining Intel technology's compatibility with government AI models); *see* Stephan Zheng et al., *The AI Economist: Taxation Policy Design Via Two-Level Deep Multiagent Reinforcement Learning*, SCI. ADVANCES, May 4, 2022, at 1, 1 (describing the AI

The possibility that AI/ML will make administrative agency actions less fair, transparent, and accountable raises a significant need for regulation. Regulation is most critical for responding to the present and future high or unacceptable risk applications. However, it is important that the regulatory framework also illuminates the use of AI/ML at all risk levels for public and intragovernmental oversight purposes. While the use of AI/ML to search a database presents considerably fewer risks than using it to predict future criminals, it nonetheless imposes non-human agency in a decision-making process.[161] Part IV addresses these issues by examining current and hypothetical regulatory tools that could reasonably constrain agency use of AI.

IV. REGULATING THE REGULATORS: TRANSPARENCY, ACCOUNTABILITY, AND FAIRNESS

AI/ML is a powerful technology that already influences the operations of administrative agencies. The need for governments to circumvent human psychological and physical limitations, automate processes, and make informed data-backed decisions drives the implementation of AI/ML. However, the unity between AI/ML power and the administrative state could be treacherous. AI/ML decision-making is far from perfect—data biases replace human biases, the outputs of the most powerful AI/ML models are unexplainable, and AI/ML is difficult to evaluate for accuracy once deployed. The risk level analysis of current AI/ML agency uses in Part III identified several high-risk implementations that could impact public rights and safety. The amount of high-risk AI/ML uses will likely only increase with time as the power of AI/ML continues to grow. In light of these circumstances, the rise of the AI administrative state must be slowed down and monitored carefully to ensure the technology is used responsibly and safely. Now is the time to impose regulation upon administrative agencies constraining

---

Economist, which is a deep reinforcement learning model trained to optimize economic policy by simulating a simple economy environment).

161. To highlight this point, consider how an AI/ML model changes the research process. The model is trained to identify relevant content and deliver it to the user. By deciding what content is relevant and what is not, the AI/ML is shaping and guiding the research. Certain content will be shown, other content will not. AI/ML discretion constrains the universe of information that goes into the research process, for better or worse.

their use of AI/ML. Future capabilities of AI/ML are impossible to predict, so timely action is necessary to prevent future damage to democratic institutions, public safety, and human rights.

The United States government has taken initial steps to provide guidance on agency use of AI/ML and many regulatory mechanisms presently exist that could constrain the implementation of the technology; however, these measures are not sufficient. AI/ML regulation must translate transparency, accountability, and fairness concerns into legal mechanisms holding agencies to democratically appropriate procedures that ensure substantively sound AI/ML uses. Current checks to AI/ML deployment do not sufficiently permit the enforcement of the regulatory goals and are too permissive towards the expansion of the algorithmic administrative state. The dangers presented by current regulatory practices regarding agency AI/ML uses justify new mechanisms limiting agencies and guiding them towards permissible uses.

Part IV analyzes the regulatory mechanisms best equipped to ensure these requirements are met. Section A lays out the guiding principles for AI/ML regulation. The section considers the need to regulate the substance of agency AI/ML actions, the dangers of leaving AI/ML regulation entirely to the discretion of the executive branch, and the importance of sacrificing government efficiency for public safety. Section B outlines the most significant regulatory actions already taken to constrain agency use of AI/ML and discusses the existing regulatory tools that could facilitate greater transparency, accountability, and oversight. The section concludes that this regulatory structure spanning the executive, legislative, and judicial branches is insufficient to manage the undemocratic features of AI/ML and ensure safe use. Section C argues for the passage of an AI/ML Act that minimally requires technical and harm-based risk assessments of agency AI/ML uses, amends existing mechanisms to support transparency and public avenues for accountability, and establishes an AI Agency to legally enforce regulation from within the executive branch.

A. GUIDING PRINCIPLES FOR REGULATION

The power of AI/ML presents novel challenges for regulation that will conflict with the preexisting framework for administrative agency oversight. There is a significant need to rethink agency regulation specifically in response to the capacity

for AI/ML to undermine government transparency, accountability, and fairness. Three principles will guide the analysis of the existing regulatory framework and the proposed best practices for future AI/ML regulation: (1) procedure must contend with the technological and applied substance of AI/ML agency actions, (2) regulation must be a whole-government effort and not insulated within the executive branch, and (3) regulation must err on the side of incurring government inefficiency to protect from dangerous AI/ML uses. This section defends these principles.

1. The Importance of Substance

There is a longstanding debate in the administrative law context about the extent agency actions should be evaluated by oversight bodies, primarily courts, to ensure adherence to substantively-sound decision-making practices. In other words, the disagreement centers on whether agency procedural constraints can and should embody empirical and normative substance considerations.[162] The argument against substance review points out that agencies act with congressionally delegated authority and are most competent as experts to decide how to best interpret and act with respect to the law.[163] This perspective has generally won out in administrative law, limiting substance considerations to ensuring agencies act within the confines of statutory purpose and explain their reasoning for a given decision.[164] This doctrine leaves, however, little room for evaluating whether agencies made the right choice.

Limiting the misuse of AI/ML power, however, requires oversight that can contend with both empirical and normative substance considerations. Addressing AI/ML under the existing agency regulatory philosophy will likely prove ineffective. AI/ML decision-making must be carefully constrained because of its power, scope of decision-making, potential for biases, need for trial-testing, independence from human discretion, and potential for substantial and transformative harm to democratic

---

162. David Dyzenhaus, *Process and Substance as Aspects of the Public Law Form*, 74 CAMBRIDGE L.J. 286, 286 (2015); Maria Ponomarenko, *Substance and Procedure in Local Administrative Law*, 170 U. PA. L. REV. 1527, 1532 (2022).

163. Shannon Roesler, *Agency Reasons at the Intersection of Expertise and Presidential Preferences*, 71 ADMIN. L. REV. 491, 497 (2019).

164. *Id.* at 499.

institutions and human rights. Agencies must make sound empirical and normative judgments to avoid these issues. AI/ML raises technical questions concerning whether the model embodies legal constraints, uses the correct information for decision-making, achieves a fair outcome not grounded in biased data, and has undergone sufficient testing to ensure accuracy in real-world scenarios. AI/ML also raises normative questions about whether the technology is appropriate for the proposed agency action compared to alternative solutions, and whether the utility of deployment outweighs the possible risks. Agency use of AI/ML in decision-making could follow all current procedures and nevertheless erode democratic values and human rights. Considering the extent of possible harm, regulation must contend with these two layers of substance considerations and not simply leave the technical and normative judgments to administrative agency discretion.

2. The Danger of Executive-Only Oversight

Agency deployment of AI could be governed entirely within the executive branch, mitigating the need for Congress, courts, and the public to engage in regulatory oversight. Commentators have endorsed this possibility on grounds that administrative law does not generally demand extensive transparency or the involvement of the public,[165] and the executive branch's expertise and capacity to act quickly renders the institution more appropriate for regulation than Congress or the courts.[166] However, insulating AI/ML regulation in the executive will be an unstable scheme. The President can provide for and remove oversight by means of executive orders. There is a significant risk that future presidents will endorse the irresponsible expansion of AI/ML uses in the administrative state, and regulations must constrain this possible action. The regulatory analysis here will thus presume the need for judicial intervention and public action are important components of agency AI/ML accountability, and that oversight internal to the executive branch must be legislatively bound to ensure that presidents cannot choose to remove constrictive mechanisms.

---

165. Coglianese, *supra* note 119, at 108.

166. Aram A. Gavoor & Raffi Teperdjian, *A Structural Solution to Mitigating Artificial Intelligence Bias in Administrative Agencies*, 89 GEO. WASH. L. REV. ARGUENDO 71, 87 (2021).

3. The Need to Trade Efficiency for Safety

As previously discussed, AI/ML holds great potential to improve the operations of the administrative state, and there is good reason to use the technology to facilitate the project of governance. AI/ML, however, is a far from perfect technology that will only grow in power. Many high-risk AI/ML uses could greatly improve the functioning of government, but they could also present a significant danger to democratic values and human rights, both through technical failures and poorly considered implementation. The unpredictability of present and future AI/ML capabilities should be a call for prudence. A regulatory scheme at this early age of algorithmic governance must favor safety over inefficiency. Arguments for improving government with unchecked AI/ML deployment should not triumph over the great need to tread cautiously and carefully towards the AI/ML future. Regulations can be burdensome towards effective governance, but this is a valuable tradeoff—they will ensure the long-term project of algorithmic governance is accomplished responsibly. Inefficient government in the short-term is a worthy sacrifice for effective algorithmic governance in the long-term. This consideration, along with previously described principles, guide the critique of existing AI/ML regulatory mechanisms and the recommendations that follow.

B. CURRENT AI/ML REGULATION CONCERNING THE ADMINISTRATIVE STATE

All three branches of government have acted, or have the capacity to act, in response to agency use of AI/ML. This section surveys the possible avenues for agency AI/ML use regulation to highlight the main problems with the existing framework. The problems broadly concern the inability for existing mechanisms to handle AI/ML threats to transparency, accountability, and fairness, or their failure to respect the AI/ML regulatory principles calling for substance considerations, presidential control, and safety prioritization. The executive branch is not always transparent about AI/ML uses, lacks internal accountability measures, and broadly encourages the expansive use of AI/ML. The legislative branch does not sufficiently require agency transparency nor provide avenues to demand accountability. The judicial branch is limited in its capacity to review AI/ML uses in agency decision-making. These limitations

raise a significant need for reconsidering the regulatory approach to AI/ML in agencies.

1. Executive Oversight

The executive branch has taken limited action to restrain agency deployment of AI/ML. The first significant activity came from the Obama Administration in 2016, which released the *Preparing for the Future of Artificial Intelligence* report[167] and the *National Artificial Intelligence Research and Development Strategic Plan*.[168] These actions indicated concern over the use of AI/ML in government, but only provided unenforceable guidance to agencies. The next steps were taken by the Trump Administration through Executive Orders 13,859[169] and 13,960.[170] Executive Order 13,960 laid out nine principles for agencies to adhere to when designing and using AI/ML, generally conveying that the uses should be lawful, purposeful, performance-driven, accurate, reliable, effective, safe, secure, understandable, regularly monitored, transparent, and accountable.[171] The order also ordered the Office of Management and Budget (OMB) to develop policy guidance to support AI/ML use by agencies, and required agencies to prepare an inventory of use cases, review their consistency with the executive order, and share the inventory with other agencies or the public when the use is non-classified and non-sensitive.[172] While this executive order does call for greater executive oversight of AI/ML, the action is not enforceable by courts, exists at the whim of the president, leaves accountability to the discretion of the executive,[173] and encourages the expansion of algorithmic

167. NAT'L SCI. & TECH. COUNCIL, PREPARING FOR THE FUTURE OF ARTIFICIAL INTELLIGENCE (2016) (emphasizing that AI/ML challenges will arise in adopting laboratory-developed technology into the "open world," in ensuring ethical use of the technology, and the need for regulation.).

168. NAT'L SCI. & TECH. COUNCIL, THE NATIONAL ARTIFICIAL INTELLIGENCE RESEARCH AND DEVELOPMENT STRATEGY PLAN (2016) (calling for awareness of ethical, legal, and societal implications of AI/ML, as well as recommending benchmarking and AI/ML standards for evaluation purposes).

169. Exec. Order No. 13,859, 84 Fed. Reg. 3967 (Feb. 14, 2019) (requiring the Director of the Office of Management and Budget to issue a memorandum to federal agencies providing guidance on developing AI/ML regulations).

170. Exec. Order No. 13,960, 85 Fed. Reg. 78939 (Dec. 3, 2020).

171. *Id.*

172. *Id.*

173. *See* GOVERNMENT BY ALGORITHM, *supra* note 18, at 75–78 (explaining transparency and accountability concerns).

governance without clear enforcement mechanisms. The Biden administration has primarily grappled with AI/ML regulations for private actors in response to ChatGPT.[174] The administration also released a blueprint for an AI Bill of Rights which explicitly notes that the principles in the blueprint, which call for safe and effective systems, may not necessarily be appropriate as applied to agency AI/ML uses.[175]

The executive branch has existing structures that could provide greater oversight of agency AI/ML uses. The Office of Information and Regulatory Affairs (OIRA) and the Office of Science and Technology Policy (OSTP) are offices within the Executive Office of the President (EOP) and could be relevant for AI/ML regulation purposes. OIRA was established in 1980 and is statutorily tasked with facilitating executive compliance with the Regulatory Flexibility Act and Congressional Review Act.[176] Pursuant to Executive Order 12,866, OIRA's duties have expanded to encompass soft power review of any "significant regulatory action" by requiring that agencies undergo regulatory impact analysis.[177] OSTP was founded in 1976 and advises the President on the technological aspects of the economy, national security, and other subjects while coordinating science and technology actions amongst agencies.[178] Both offices could bring the President's attention to AI/ML agency action, and the President could then exert regulatory pressure on undesirable uses.[179] However, this oversight is an exercise in soft power and would not necessarily hinder an agency from deploying high risk AI/ML. OIRA review is also limited to significant regulatory actions. Significant actions include those that materially alter the budgetary impact of programs or raise novel legal or policy issues.[180] Unfortunately, discretion over this determination lies

---

174. David Shepardson & Diane Bartz, *US Begins Study of Possible Rules to Regulate AI Like ChatGPT*, REUTERS (Apr. 12, 2023, 1:28 AM), https://www.reuters.com/technology/us-begins-study-possible-rules-regulate-ai-like-chatgpt-2023-04-11/.

175. WHITE HOUSE OFF. OF SCI. & TECH. POL'Y, BLUEPRINT FOR AN AI BILL OF RIGHTS: MAKING AUTOMATED SYSTEMS WORK FOR THE AMERICAN PEOPLE 2 (2022).

176. Gavoor & Teperdjian, *supra* note 166, at 76–77.

177. Exec. Order No. 12,866, 58 Fed. Reg. 51735, 51738 (Oct. 4, 1993).

178. Gavoor & Teperdjian, *supra* note 166, at 76.

179. *Id.* at 77.

180. Exec. Order No. 12,866, 58 Fed. Reg. 51735, 51738 (Oct. 4, 1993).

with the agency's director.[181] Agencies could find that high-risk AI/ML uses do not have cause for review and choose not to alert OIRA to their pending promulgation, thus limiting the regulatory capacity of OIRA.

2. Legislative Actions

Congress has attempted to impose on administrative agency AI/ML use with statutory mandates. The AI in Government Act of 2020 established the AI Center of Excellence within the General Services Administration and tasked it with facilitating the adoption of AI/ML in federal government, improving cohesion and competency in the adoption, carrying out activities benefiting the public, and enhancing the productivity and efficiency in federal government operations.[182] The statute also directs the OMB to release a memorandum facilitating AI/ML development in administrative agencies and advocating for best practices.[183] This approach to regulating agencies is insufficient for two reasons. First, the AI Center can only play a soft power regulatory role because it lacks enforcement powers. Second, the statute tasks the AI Center with encouraging the development of AI/ML in government without articulating per se limitations on its use. Many of the recommendations aim to make agency implementation of AI/ML easier. The AI in Government Act has also been ignored in part by the executive; the OMB has yet to issue the required guidance memorandum.[184] This initial agency noncompliance signals the inability for advisory measures to accomplish regulatory goals.

Congress passed two other statutes that could facilitate the regulation of agency use of AI/ML: the Administrative Procedure Act (APA) and the Freedom of Information Act (FOIA). The APA binds agencies to procedures that are enforceable by courts, while FOIA permits public access to government records.

---

181. Francesca Bignami, *Artificial Intelligence Accountability of Public Administration*, 70 AM. J. COMPAR. L. (ISSUE SUPPLEMENT 1) i312, at i337–38 (2022).

182. Consolidated Appropriations Act, 2021, Pub. L. No. 116-260, 134 Stat. 1182 (2020) (integrating the AI in Government Act).

183. *Id.*

184. Letter from Rob Portman to OMB Director Shalanda Young Regarding the Implementation Status of the AI in Government Act (Dec. 22, 2022) (found at https://www.hsgac.senate.gov/media/minority-media/portman-presses-omb-on-implementation-of-ai-in-government-act_/).

The APA binds agencies to court-enforced procedures.[185] The APA addresses three types of agency actions: (1) rulemaking, by which an agency formulates, amends, or repeals a rule interpreting or prescribing law or policy,[186] (2) adjudication, through which agencies formulate an order resolving a dispute between parties,[187] and (3) discretionary decision-making.[188] The APA provides procedures to regulate agency rulemaking and adjudicatory actions, but does not impose constraints on agency actions left to their discretion.[189] Statements of policy, interpretive rules, or agency-imposed rules of organization, procedure, or practice are also exempted as non-legislative rulemaking.[190] The distinction between legislative and non-legislative rulemaking is somewhat ambiguous.[191]

The APA requires notice to the public for all legislative rulemaking, an opportunity for comment, and publication of an explanation addressing comments with the final rule. Notice generally requires publication of the proposed rule in the Federal Register, along with: (1) a statement of the time, place, and nature of public rule making procedures, (2) reference to the legal authority under which the rule is proposed, and (3) either the terms or substance of the proposed rule or a description of the subjects and issues involved.[192] The agency then must give interested persons the opportunity to submit written data, views, or arguments as a means of participating in rulemaking.[193] After consideration of the public comments and in publication of the final rule, the agency must include a concise

185. Christopher J. Walker, *Modernizing the Administrative Procedure Act*, 69 ADMIN. L. REV. 629, 633 (2017).

186. Administrative Procedure Act, 5 U.S.C. § 551(5).

187. 5 U.S.C. § 551(7). Adjudication procedure is not covered in this Note.

188. Roni Elias, *The Legislative History of the Administrative Procedure Act*, 27 FORDHAM ENV. L. REV. 207, 214 (2015).

189. Heckler v. Chaney, 470 U.S. 821 (1985) (holding agency enforcement actions are presumptively unreviewable and committed to agency discretion under 5 U.S.C. § 701(a)(2) unless the organic statute provides guidelines to follow for exercising enforcement powers).

190. Administrative Procedure Act, 5 U.S.C. § 553(b)(3)(A).

191. *See e.g.,* Chamber of Commerce of the U.S. v. U.S. Dep't of Labor, 174 F.3d 206 (D.C. Cir. 1999) (holding that a policy incentivizing employer self-regulation of safety and health programs with fewer workplace inspections is a legislative rule).

192. 5 U.S.C. § 553(b)(1-3).

193. 5 U.S.C. § 553(c).

statement of basis and purpose[194] explaining why it made the decision given alternatives, what facts were relied on, and what factors went into a policy judgment.[195]

For the APA to affect AI/ML rulemaking, the agency use must constitute a legislative rule. This limits the extent that APA procedures can regulate AI/ML. While courts consider AI/ML used for benefits determinations to be legislative rules,[196] high-risk implementations that influence enforcement priorities and targeting may not be similarly categorized. Even when AI/ML use is clearly a guidance or interpretive statement, non-legislative rules still influence public actors by signposting future enforcement directions. AI/ML decision-making would subsequently carry unreviewable public influence. It is also unclear how much information about the AI/ML implementation is necessary to satisfy APA notice requirements. Without a specific requirement for transparency regarding the technical composition of the AI/ML and its intended use, notice and comment procedure may have limited functionality as a mode for public participation. Finally, AI/ML could influence the agency decision-making process in ways imperceivable by APA regulatory mechanisms. For example, a search engine that uses AI/ML to prioritize results could hinder robust research by showing certain sources of information and avoiding others. Even if AI/ML is not used as a component of a rule, as a rule itself, or during the adjudication process, it may still influence decision-making. APA notice and comment will allow public oversight over AI/ML uses that are deemed legislative rules, but this oversight is likely too limited to address all high-risk uses.

FOIA allows for the public to request and gain access to government records, particularly those of administrative agencies.[197] The Supreme Court has stated that FOIA "defines a structural necessity in a real democracy."[198] FOIA plays a critical role in government transparency and could be an avenue

194. 5 U.S.C. § 553(c).

195. U.S. v. Nova Scotia Food Prod. Corp., 568 F.2d 240 (2d Cir. 1977).

196. Ark. Dep't of Hum. Servs. v. Ledgerwood, 530 S.W.3d 336, 344–45 (Ark. 2017) (holding that AI/ML allocating home-care hours to disabled low-income individuals is a legislative rule).

197. David E. Pozen, *Freedom of Information Beyond the Freedom of Information Act*, 165 U. PA. L. REV. 1097, 1102 (2017); The Freedom of Information Act, 5 U.S.C. § 552.

198. *Id.* at 1098.

for increasing public awareness of the technical substance of agency AI/ML through records requests. However, several problems limit the effectiveness of FOIA for this purpose. First, statutory exemptions could preclude disclosure of AI/ML used for law enforcement or national security purposes, which likely implicate high-risk implementations.[199] Second, when AI/ML is subject to FOIA disclosure, the most relevant information may be excluded.[200] For example, a group of academics used a FOIA request to acquire information about the Department of Homeland Security's Risk Classification Assessment system but were unable to gain access to the code.[201]

3. Judicial Review

Judicial review of agency decision-making authorized under the APA could provide accountability regulations regarding the use of AI/ML, although current doctrines impede the effectiveness of this avenue. Courts may hold unlawful and set aside agency action, findings, and conclusions found to be (1) arbitrary, capricious, an abuse of discretion, or not otherwise in accordance with law, (2) in excess of statutory jurisdiction, authority, limitations, or short of statutory right, or (3) without observance of procedure required by law.[202] Court review is largely limited to ensuring agency compliance with procedure, but not with substantively sound decision-making.

Arbitrary and capricious review is a possible avenue for regulating agency AI/ML uses. Under *Motor Vehicle Mfrs. Ass'n v. State Farm Mut. Auto. Ins. Co.*, courts undertake arbitrary and capricious, or hard look review, of agency decisions by considering whether there is a rational connection between the facts found and the choice made, whether there was a consideration of the relevant factors, and whether there was a reliance on improper factors or a clear error of judgment.[203] This test can be used to overturn AI/ML implementations that are

199. Note that not all law enforcement exemptions are included. 5 U.S.C. § 552(b)(7).

200. The extent the dataset can be disclosed will be limited by the Privacy Act of 1974. 5 U.S.C. § 552(b)(7).

201. Bignami, *supra* note 181, at i336.

202. 5 U.S.C. § 706(2). Note there are additional bases for review not covered here.

203. Motor Vehicle Mfrs. Ass'n of U.S. v. State Farm Mut. Auto. Ins. Co., 463 U.S. 29, 43 (1983).

harmful to the public, either explicitly in goal or implicitly in the use of biased data. There is also a plausible argument that agencies should be entitled little deference[204] in their use of AI/ML because agencies are not specifically delegated authority or guaranteed to have expertise on the subject. However, in practice, courts give wide deference to agencies for their use of AI/ML.[205]

Two avenues exist for holding agencies accountable for potentially unfair AI/ML uses. The first is procedural due process under *Matthews v. Eldridge*.[206] The court must consider (1) the affected private interests, (2) the potential for reducing decision-making error, and (3) the government's interests in limiting fiscal and administrative burdens.[207] Unfortunately, this test will often support AI/ML uses under the third factor because of their efficiency and under the second factor (foregoing obvious error) because AI/ML models' opaqueness can hide the errors. The other possibility is challenging an AI/ML model under the Equal Protection Clause of the United States Constitution.[208] However, the difficulty of proving discriminatory intent solely on the basis that the agency deployed a biased AI/ML model will likely limit the success of this argument.[209]

Finally, as Engstrom and his colleagues note, judicial review will be ineffective for accountability purposes if judges, like most, if not all people, have difficulty critically analyzing algorithmic decision-making.[210] Without mandates and standards governing how agencies should articulate the technical substance of AI/ML models, judicial review will likely prove ineffective for regulation even if courts take a hard look at the challenged agency AI/ML action.

As noted at the outset, current government actions and regulatory tools to protect democratic institutions and human rights do not promise an effective capacity to respond to AI/ML.

---

204. *See* Skidmore v. Swift & Co., 323 U.S. 134, 140 (1944) (considered the least deferential standard for judicial review of agency action).

205. *See* discussion *supra* Part IV.A. *See also* Coglianese, *supra note* 119, at 113.

206. Mathews v. Eldridge, 424 U.S. 319 (1976).

207. *Id.* at 321.

208. U.S. CONST. amend. XIV § 1.

209. Bignami, *supra* note 181, at i344.

210. GOVERNMENT BY ALGORITHM, *supra* note 18, at 77.

Many of the actions are unenforceable or ignored, current tools are limited in addressing opaque decision-making technologies, and there are insufficient avenues means for the public or the executive to stand up to dangerous agency AI/ML use cases. The following section considers how amendments to existing regulations and the creation of novel mechanisms can fill in the gaps and allow for an effective response to the AI/ML administrative state.

C. RESPONDING TO THE DANGERS OF AI/ML WITH NEW AGENCY REGULATIONS

As the administrative state currently stands, AI/ML threatens to make agency action less transparent, accountable, and fair because of the inherent flaws of the technology. Current oversight mechanisms within the executive, from Congress, and through the courts do not appear sufficient to challenge problematic algorithmic governance. Additional regulation is needed to protect democratic values and human rights. This regulation must require substantive transparency and oversight of agency AI/ML uses. It must extend beyond the executive branch and involve Congress, the courts, and the public. Finally, regulations must be willing to sacrifice government efficiency over the deployment of high-risk, and thus potentially dangerous, implementations that may otherwise facilitate effective government.

Congress should pass a statute (Proposed AI/ML Act) guiding and constraining agency use of AI/ML that embodies these considerations and principles. Congress is best situated to respond to AI/ML regulatory needs for several reasons. First, the legislature is the most democratic government institution and can most legitimately articulate a consensus view of the desired technical and normative considerations to guide agencies as they deploy AI/ML. These considerations would be important for identifying the risk level of potential AI/ML applications and evaluative metrics for deciding whether the risk is too high to permit use. Second, Congress can ensure that existing mechanisms facilitating public transparency and accountability demands of agencies are strengthened and clarified in how they interact with novel legal questions posed by AI/ML uses. The public should have ample opportunity to participate in the development of the algorithmic administrative state, judge whether the use is too harmful to be permitted, and seek

recourse through the courts. Third, Congress can create enforceable interagency oversight within the executive branch to maintain compliance with procedural requirements. The oversight would not be removable or amendable by presidential discretion. This section explains in greater detail how the Proposed AI/ML Act would operate by standardizing agency technical and risk assessment of current and future AI/ML uses, clarifying existing statutes like the APA and FOIA for public access to AI/ML information and legal avenues to challenge the use of the technology, and developing an AI/ML Agency within the executive branch for interagency oversight.

1. Requirement of Standardized Technical and Risk Assessment

The Proposed AI/ML Act should require agencies to evaluate all current and future AI/ML uses for technical substance and potential deployment harm risk. The assessment must be standardized by Congress to incorporate normative considerations regarding desirable and undesirable AI/ML designs and uses. The normative considerations must err on the side of caution with a preference for finding AI/ML uses to be poorly designed or risky when there is ambiguity. The technical assessment component should require disclosure and expert evaluation of the AI/ML model design.[211] The disclosure could include, and the evaluation consider, for example, information about the dataset the AI/ML was trained on,[212] the goals behind the training process, the type of algorithm used for training, why it was chosen, the knowable factors that drive the model outputs, the performance measures of the model, and the reason the performance measures were chosen and were determined to effectively measure performance. The risk assessment should integrate the findings from the technical assessment and further

211. Congress must decide if the experts should come from within the agency/executive branch or from neutral third parties. *See generally* Katherine Miller, *Radical Proposal: Third-Party Auditor Access for AI Accountability*, STAN. U. HAI (Oct. 20, 2021), https://hai.stanford.edu/news/radical-proposal-third-party-auditor-access-ai-accountability (advocating for third-party audits of private AI systems).

212. The extent the dataset can be disclosed will be limited by the Privacy Act of 1974, 5 U.S.C. § 552a(b), as far as disclosure would reveal personal information about individuals, and the Freedom of Information Act, 5 U.S.C. § 552(b)(7), as far as the information could interfere with law enforcement purposes.

consider the proposed implementation's potential for harm to democratic values and human rights, analogous to categorization under EU AI Act.[213] The agency would classify the technical and harm considerations under a risk category and justify the categorization. The assessments and conclusions should then be summarized in plain language in a report.

The assessment report will provide the necessary substance disclosures to greatly improve the transparency of agency AI/ML uses and thus facilitate government accountability measures. With these assessments, the public will be afforded a greater understanding of how agency AI/ML will affect individuals, groups, and society. Courts undertaking review of agency AI/ML actions will have a framework for evaluating uses on several preexisting legal grounds. Interagency oversight will have a standardized metric for tracking and evaluating AI/ML uses across the administrative state. Congress can also use the assessments for oversight, demanding agencies appear at hearings and justify high-risk uses, creating a public political issue out of perceived undesirable AI/ML implementation, and applying pressure for agencies to adhere to safe AI/ML uses when possible. Finally, requiring agencies to assess their AI/ML uses may also serve a soft regulatory purpose. The assessment process could remind agencies that certain AI/ML development practices are better than others, and certain deployments may impose greater risk than initially anticipated. Thus, agencies may be incentivized to self-govern AI/ML design and uses in anticipation of the evaluation.

2. Requirement of Public Disclosures and Accountability Measures

The Proposed AI/ML Act should require greater public access to information about governing technologies and legal avenues to hold the government accountable for their use. This can largely be achieved by amending existing statutes such as the APA and FOIA.

Modification of the APA must address specific AI/ML requirements for the notice and comment process and permit greater substantive judicial review when agencies use the

---

213. The extent the AI/ML removes human discretion may also be an important consideration for the framework, insofar as limited human discretion creates or exacerbates the risk.

technology. These steps will ensure greater transparency and accountability. The Proposed AI/ML Act should amend the APA to articulate which AI/ML uses constitute rulemaking requiring notice and comment. Rather than following current legislative rule doctrine, Congress could require that all proposed AI/ML uses above a certain risk category threshold undergo notice and comment procedure. To improve public competence for evaluating proposed AI/ML uses, the APA should require that the "concise statement of basis and purpose" disclosure under 5 U.S.C. § 553(b) occur for AI/ML rulemaking at the notice stage rather than upon rule promulgation. The concise statement should be required to contain the risk assessment report so that the public can evaluate it and give feedback prior to the AI/ML deployment.[214] Finally, APA-authorized judicial review standards such as for arbitrary and capricious agency decision-making should have AI/ML specific requirements. Courts should be directed to consider notice without the assessment report procedurally invalid, be permitted to analyze the substance of the assessment report, and granted discretion to decide whether the agency sufficiently justified deployment of the AI/ML model in consideration of the risks.

Congress should also take a few additional steps to ensure transparency of AI/ML uses. First, it should amend FOIA to clarify what AI/ML data is requestable. This change would save time and money by avoiding unnecessary litigation and signal to the public that they can and should ask for AI/ML information. While the greater volume of FOIA requests will inevitably slow the already lethargic disclosure process down further, FOIA could itself deploy AI/ML to meet the increased demand with automation. Congress should also codify the mandate of Executive Order 19,960 for agencies to publish their AI/ML use cases for public access, protecting the sound policy against a

---

214. Notice and comment may also require an update due to the unique features of AI/ML rulemaking. An AI/ML model subject to notice and comment procedures could change substantially through training during the comment period, yet still appear similar enough to pass the "logical outgrowth" test relative to the original model. Engstrom and his colleagues propose triggering notice and comment when an AI/ML model is insufficiently subject to human discretion. GOVERNMENT BY ALGORITHM, *supra* note 18, at 77. A more conservative requirement would be to require that high risk AI/ML models are trained in batches, taken offline when deployed, and then undergo additional notice and comment if the agency desires supplementary training of the model. For more on batch training, *see* discussion *supra* Part II.B.2.

president who might revoke the order. The publications should include the assessment reports so the public is made aware of the deployed AI/ML models and can evaluate the quality of the models, the proposed justification for their use, and the anticipated risks the use may cause. This will allow the public to better criticize the deployments, vote for politicians concerned with AI/ML use, or otherwise democratically engage to express discontent with agency AI/ML applications.

3. Requirement of Interagency Oversight Measures

The Proposed AI/ML Act must create legally enforceable accountability and oversight measures from within the executive, imposed on the administrative agencies. Such measures should accomplish three aims: (1) provide for an executive mechanism to further ensure agency compliance with the previously discussed regulations, (2) allow agency experts to hinder or halt dangerous AI/ML uses before they go into effect, and (3) establish a body to issue new rules constraining agencies as appropriate for future, more powerful iterations of AI/ML.

A variety of commentators have proposed oversight structures to handle interagency accountability. Gavoor & Teperdjian argue that interagency oversight should be conducted by OIRA and OSTP, with expanded powers related to AI/ML granted by the executive, because Congressional action is too tenuous to rely on for creating new oversight bodies.[215] While the concern about Congressional action is valid, relying on executive authority is not a sound long-term solution for regulating agencies. Engstrom and his colleagues propose AI oversight boards within agencies staffed with experts who could monitor, investigate, and recommend changes to how the agency uses AI/ML.[216] Boards within agencies alone, however, will likely prove insufficient in ensuring a standardized executive process for regulating AI/ML uses and could be unduly influenced by their specific agency priorities.

The Proposed AI/ML Act should thus minimally create an AI/ML Agency to conduct executive oversight.[217] The AI/ML

215. Gavoor & Teperdjian, *supra* note 166, at 87.

216. GOVERNMENT BY ALGORITHM, *supra* note 18, at 77.

217. Engstrom and his colleagues also imagine an AI agency that would serve the dual purpose of government oversight as well as operating as an "FDA-style" regulatory body for AI/ML. *Id.* at 75. Another author has floated this approach with the idea of a Federal Robotics Commission. Ryan Calo,

Agency should be delegated enforcement power over the executive to ensure compliance with the discussed assessment report, APA, and FOIA requirements. The AI/ML Agency should also have the power to consider new regulatory mechanisms as the need arises. Congress may find it prudent to insulate the AI/ML Agency as an independent agency to limit presidential removal power of its officers. However, Congress could take an alternative route by placing the AI/ML agency within the EOP or expanding the duties of existing EOP offices like OIRA or OSTP to encompass AI/ML regulation.

An AI/ML Agency would serve several key regulatory roles. First, it would centralize the flow of AI/ML related information in the executive branch. This can be done by requiring notice to the AI/ML Agency of all agency AI/ML uses and proposed AI/ML deployments through the collection of risk assessment reports. Second, the Agency would allow greater executive control over AI/ML uses by assuming final approval discretion of the risk assessment reports. Dangerous uses or underdeveloped reports can be rejected in advance of APA procedure. The AI/ML Agency would then serve as another mechanism beyond courts to ensure the assessment reports are completed and done properly. Finally, the AI/ML Agency would facilitate ex post review of deployed AI/ML models. Engstrom and Ho advocate for agencies to undergo prospective benchmarking, wherein the agency selects random cases to hold-out from AI/ML decision-making and instead subject them to human decision-making.[218] The human results would then be compared to the AI/ML results to uncover biases and evaluate the effectiveness of the technology. The AI/ML agency could require agency compliance with prospective benchmarking. Prospective benchmarking would be useful for reviewing courts,[219] and the AI/ML Agency could review them for internal accountability actions. Such a procedure would help mitigate the issues arising from AI/ML deployment in real world scenarios.[220]

While further regulation will surely be required to protect democratic values and human rights from the expanding

---

Artificial Intelligence Policy: A Primer and Roadmap, 51 U.C. DAVIS L. REV. 399, 429 (2017); *see also* Ryan Calo, *Robotics and the Lessons of Cyberlaw*, 103 CALIF. L. REV. 513, 556 (2015).

218. Engstrom & Ho, *supra* note 19, at 849.

219. *Id.* at 850.

220. *See* discussion *supra* Part II.C.

algorithmic administrative state, the recommendations here should provide a baseline. However, regulating AI/ML will only succeed with extensive public and private support of the constraints. Experts will need to develop AI/ML technical and harm risk assessment strategies that capture meaningful information that non-experts can use to understand and evaluate the technology. Congress must coordinate and find consensus on passing legislation like the Proposed AI/ML Act. The public must be willing to engage with democratic processes to demand information and hold agencies legally accountable when they use AI/ML to cause harm. The AI/ML Agency must be capable of providing competent and timely review of AI/ML uses and take seriously the dangers of AI/ML. The country must work together to ensure that these early stages of algorithmic governance are navigated with caution to achieve a desirable and responsible unity between the power of AI/ML and government.

V. CONCLUSION

Given the explosive and unpredictable advancement of AI/ML, regulating administrative agency use of the technology is of extreme importance. AI/ML can exhibit bias, its decision-making is opaque, and real-world scenarios are not conducive to accurately evaluating AI/ML models. AI/ML technologies will only grow more powerful and government mobilization of that power could threaten democracy and human rights. At the same time, AI/ML presents opportunities for effective, data-backed governance that could improve government functioning. To ensure that the administrative state uses AI/ML safely and responsibility, regulations must expand values of government transparency, accountability, and fairness. The regulations must also create opportunities for substantive evaluation of AI/ML uses, restrict the degree to which the President can promote reckless use of the technology, and prioritize safety over government efficiency.

This Note presented an overview of how machines make decisions and the various types of errors that may arise due to the limitations of data science methods and data-centric decision-making more broadly. It surveyed the arguments supporting and opposing algorithmic governance, evaluated the present status of AI/ML implementations in administrative agencies, and examined the existing regulatory mechanisms the

three branches of government could employ to regulate agency AI/ML uses. This Note argued for the implementation of additional regulations to respond to the dangerous power of AI/ML. The proposed regulations would require technical and harm risk-based assessment reports for agency AI/ML uses, amendment to statutes like the APA and FOIA to facilitate effective public oversight, and the establishment of an AI/ML Agency to accomplish regulatory goals from within the executive branch.

The regulation of agency AI/ML involves intricate nuances that are best addressed by data scientists, policy experts, and administrators within the administrative state. However, the hope is that this Note will encourage those with similar concerns to move beyond mere theoretical work on this subject. All three branches of government have a crucial role to play in the development of a regulatory framework that can effectively address and prevent the irresponsible use of AI/ML by agencies. The clock is running out when it comes to confronting the impact of AI/ML on human life, and public institutions must be prepared to meet the evolving technological landscape. The AI/ML Pandora's box has been opened, and now is the time to act.

***