



Legal Studies Research Paper Series

Automation Rights: How to Rationally Design Humans-Out-Of-The-Loop Law

Research Paper 24-013

Orly Lobel

This paper can be downloaded without charge from the
Social Science Research Network Electronic Paper Collection

AUTOMATION RIGHTS:
HOW TO RATIONALLY DESIGN HUMANS-OUT-OF-THE-LOOP LAW

*Orly Lobel**

This essay begins with the following puzzle: in sharp contrast to significant evidence demonstrating the effectiveness of AI-based automation in high stakes spheres – health care, transportation, national security, finance, workplace safety, public administration, and more – the contemporary impulse is to legally require a human-into-the-loop is heightened the higher the stakes of the activity or decision. Indeed, in the legislation emerging in both the EU and the United States, ironically showcases the assumption that when it comes to AI, high stakes equal high-risk of tackling the stakes through the most advanced technology. Moreover, while there are hundreds of bills, reports, and executive orders that seek to prohibit or restrain certain uses or applications of AI, there are virtually no equivalent frameworks, or even language, that would mandate automation when such a shift has been empirically shown to be the safest, or most consistent in achieving agreed upon goals or courses of action. This essay, written for the University of Chicago Symposium on *AI How AI Will Change the Law*, argues for the development of more robust – and balanced – law that focuses not only on the risks but also the potential that AI brings. In turn, it argues that there is a need to develop a framework for laws and policies that incentivize, and at times, mandate transitions to AI-based automation. Automation rights – the right to demand and the duty to deploy AI-based technology when it outperforms human-based action – should become part of the legal landscape. A rational analysis of the costs and benefits of AI deployment would suggest that certain high-stakes circumstances compel automation because of the high costs and risks of not adopting the best available technologies. Inevitably, the rapid advancements in machine learning will mean that law soon must embrace AI, accelerate deployment, and under certain circumstances prohibit human intervention, as a matter of fairness, welfare, and justice.

The essay suggests that the thinness of legal thinking on automation rights in both policy and scholarship can be connected to several flaws in AI debates. First and foremost is the dearth of [comparative analysis](#) of automation versus a human decision-maker, a comparison that must be made along at least six related yet distinct axis:

1. *Machine v. human current performance in relation to the desired outcomes – accuracy, consistency, safety, speed.*
2. *Machine v. human scalability/access.*
3. *Machine v. human blackbox opacity/explainability.*
4. *Machine v. human traceability/detection of failures.*
5. *Machine v. human learning/improvement prospects.*

* Warren Distinguished Professor of Law and Director of the Center for Employment and Labor Policy (CELP), University of San Diego.

6. *Machine v. human liability schemes.*

Instead of developing such a comparative matrix, debates about AI tend to focus on the risks and fails of the technology in absolute terms, resulting in a double standard and a strong bias toward human action. Further propelling the failure to adopt a comparative advantage analysis when it comes to AI are broadly documented behavioral biases, including the status quo bias – the human tendency to favor what is currently in place and to fear change; the related *loss aversion effect* (Kahneman and Tversky’s famous dictum that *losses loom larger than gains*, in which people impute greater value to a given endowment when they give it up); the human tendency to [*distrust what is perceived as artificial*](#) as opposed to human or natural; and *the holier than thou bias*, the tendency to overestimate one’s own performance, despite statistical evidence of high human error. This essay argues that the responsible way for policymakers to adapt the law to the era of AI is to consider the costs and harms of staying static just as we consider the risks of an AI shift itself. Moreover, policymakers should assume the role of aiding the public to adopt a more rational relationship with AI applications, adopting laws designed to mitigate both algorithmic adoration and algorithmic aversion.

The lack of robust AI/human comparative analysis is exacerbated by two additional flaws in contemporary law and policy. First is the [*privileging of privacy*](#), including deontological or theoretical privacy rights as trumping other individual rights and social goals. The outsized and often misleading fear of the loss of privacy has contributed to [*a dearth of law pertaining*](#) to mandates on fuller data collection. Second, a conflation between technological readiness and the effects of technological deployment. For example, questions about the safety of autonomous trucks have been muddled with the separate, albeit important, questions about the inevitable job losses that will result from their legal deployment on the roads.

Giving Up the Wheel

A striking number of recent laws have sought to ban or slow down welfare enhancing, and even lifesaving, AI-based technology in transportation, medicine, law, criminal justice administration, employment, finance and education. With the rapid advancements in machine learning, the legal landscape is witnessing [*calls for bans, moratoriums, and limits*](#) on the deployment of the technology. Legal scholars call for an overarching [*precautionary rule*](#) when it comes to AI and have even [*proposed*](#) “a system of ‘unlawfulness by default’ for AI systems.” In a forthcoming article, *The AI Regulatory Toolbox*, I show that prohibitory AI laws have been skewed to the top of the regulatory pyramid, focused on bans and command-and-control prohibition with little attention to other forms and methods of regulation and governance, including standardization, public investment, assessment, and incentives to learn and develop best-practices. In 2023, the United States saw countless proposals to ban AI technologies ranging from biometrics and monitoring technologies to weapon systems, to bans on autonomous vehicles as well as the use of AI in decision-making processes, like criminal justice administration, hiring, or loan approvals. [*The Federal Algorithmic Accountability Act*](#) would require an assessment of the need for “guard rails for or limitation on certain uses of applications of the automated decision system.”

The newly enacted [EU AI Act](#) includes bans on certain “high-risk” AI practices: “Banned applications . . . biometric categorization systems that use . . . political, religious, philosophical beliefs, sexual orientation, race; untargeted scraping facial images . . . emotion recognition . . . social scoring based on social behaviour or personal characteristics; AI systems that manipulate human behaviour to circumvent their free will; AI used to exploit the vulnerabilities of people (due to their age, disability, social or economic situation).” There is no upshot path to automate when, for example, AI is used to protect against the exploitation of the vulnerabilities and fallibilities of humans. The EU AI Act differentiates between high and low risk AI systems, providing that “human oversight shall aim at preventing or [minimizing] the risks to health, safety or fundamental rights that may emerge when a high-risk AI system is used in accordance with its intended purpose or under conditions of reasonably foreseeable misuse.” The draft regulation requires high-risk AI systems to be “designed and developed in such a way, including with appropriate human-machine interface tools, that they can be effectively overseen by natural persons.” It also bans certain uses of AI that create “unacceptable risks,” although it does not specify what those risks are. It further divides the world of AI risk to high and low, subjecting high-risk AI systems to elaborate risk regulation, then funneling lower-risk AI systems into the abstract requirement of an acceptable level. There are no principles in the Act on what are acceptable levels of risks or what requirements are necessary to compare to risks emanating from the status quo of human processing.

One important example of attempts to ban life-saving AI comes from the debates over deployment of AI in transportation. Tesla, Waymo, and Uber are at the forefront of developing autonomous vehicles (AVs), with AI-enabled navigation, obstacle detection, and decision-making in real-time traffic scenarios. The primary advantage of AVs is their potential to significantly reduce accidents caused by human error. According to the National Highway Traffic Safety Administration, [94% of serious crashes are due to human error](#). AVs, through their precise and consistent operation, could lower this statistic dramatically. Automation on the roads is the antidote to driver fatigue and risky maneuvers, potentially saving lives while boosting efficiency and lowering shipping costs. The trajectory towards mandating AI applications in such an area where they are set to outperform humans is both inevitable and responsible. Yet, even the legality of their deployment is currently in question.

In 2023, California [Assembly Bill 316](#) aimed to put the brakes on self-driving trucks by mandating the presence of human safety drivers in these trucks. The proponents of the bill championed it as both a shield against job losses for the state’s 60,000 truckers and a safety measure. Governor Newsom [vetoed the bill](#), rightfully in my opinion, pointing to the [evidence](#) that self-driving trucks actually improve safety and that a human trucker over-ride could reduce rather than augment road safety. Deployment of other AI-based transportation advancements has also been slowed down by laws with questionable rationality. The [Federal Aviation Administration rules](#) requiring drone operation to occur within a human operator’s view have slowed down important advancements in delivery and other services. The California law prohibiting opt-in AI-based driving safety monitoring technologies has [prevented lifesaving technology](#) even as a voluntary private insurance-based measure.

The job loss question should not muddle the evidence about safety. Much of the fear and resistance to AI has to do with our fear of being replaced by machines. These are valid fears. We are facing undeniable waves of seismic changes in the labor market. We should anticipate and address these

changes that will occur in every industry. For example, like transportation, healthcare is facing seismic change. Already, medical AI outperforms the work of healthcare professionals in a wide variety of diagnostics and patient care, including [AI-assisted surgery](#), [AI-nursing assistants](#), and [telemedicine](#). Parallel changes are happening in the legal field, in art and entertainment, in science and technology. It is critical that we keep the questions distinct – Is AI safe and ready? Does it outperform human decision-making? Is it less prone to accidents and inconsistencies? And separately: What must the new social contract look like in the face of rapid job loss and changes in employment patterns? For everyone to enjoy the benefits of AI, the labor market effects need to be tackled with investment in reskilling, education, taxation, social welfare, public options, public procurement, and access to these revolutionary technological capabilities.

The late Daniel Kahneman [predicted in an interview](#) that “[b]eing a lot safer than people is not going to be enough. The factor by which they [AVs] have to be more safe than humans is really very high.” That is an alarming prediction. Yet in another high-stakes transportation context, the international community has already agreed that automation is much safer in high-stakes travel, such as air travel. The entirety of the international aviation industry operates with the gold standard of autopilot when weather conditions are harsh. In riskiest conditions of commercial aviation, the international community have long established rules about humans-out-of-the-loop. Reduced Vertical Separation Minimums (RVSM) allows flights to have a small vertical separation (1000’), if the instruments meet accuracy requirements, and the airplanes are operated on automatic pilot. Pilots undergo required training but give up controls in riskier proximities. If regulators, pilots, and passengers are comfortable with this standard, there is no reason to believe that we cannot learn to love “a lot safer” autonomous cars—as well as fully autonomous commercial planes. An acceptance of an “a lot safer” standard instead of a “really very high” factor of better performance requires law and policy, research, design, and education. Although the willingness of people to give up control of the wheel may be different across generations, once our individual rights and public goals can be clearly better realized with technology, a right and duty to automate become not only possible but morally correct.

A Right to an AI Decision-Maker

Where AI is proving to be more accurate and effective than human decision-making, law requires new directions for analyzing the necessity of mandating AI applications in certain fields. For example, in the medical field, particularly in diagnostic procedures, recent advancements in AI have enabled systems to diagnose certain conditions, such as skin and breast cancer or retinal diseases, more accurately than human doctors. In the financial sector, algorithms are now able to analyze market data and consumer behavior with a precision that far surpasses human capabilities. For example, JPMorgan Chase’s AI program, COIN, can interpret commercial loan agreements in seconds, a task that takes legal professionals 360,000 hours annually. By automating such tasks, AI not only brings efficiency but also minimizes human errors and bias that could have significant consequences. Failing to develop laws prohibiting human decision-making under certain circumstances is a normative failure with serious costs. Yet again, with regard to recommendation algorithms and sorting applications, we witness a slew of laws that are aimed to prohibit such

technologies while no equivalent laws mandate automation currently exist. In 2023, for example, [a New York state bill](#) was introduced to prohibit the use of algorithmic decision-making in hiring and job screening decisions, without the involvement of a human final decision-maker. Article 22 of [EU General Data Protection Regulation \(GDPR\)](#) states that “the data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her.” The exceptions are narrow and require a showing of necessity, special authorization, or “the data subject’s explicit consent.” Again, no equivalent duties to adopt AI have been contemplated in our recent laws. In particular, when it comes to civil rights, access to justice, and public administration, AI regulation should include mandates to automate—to deploy AI systems that have been proven safer, fairer, or more accurate than human decision-making and existing systems.

- **Good Governance:** Cary Coglianese and I argue in a forthcoming article *Algorithmic Administration as Constitutional Governance* that government agencies, as part of their constitutional duty to ensure good government, should shift to digitization and automation when those systems are more efficient and better at achieving our public goals.
- **Clean Slate:** Colleen Chien has [offered](#) a [compelling argument](#) that to make laws that increase the eligibility of people with criminal background to clear their records and regain the right to vote, a reality, we need to “shift administrative burdens from the defendant and onto the state and algorithms through automation, standardization, and ruthless iteration.”
- **Criminal Justice:** In criminal justice, AI systems like COMPAS have been used for sentencing and assessing the likelihood of a defendant reoffending. Their application has sparked debates around the biases inherent in AI algorithms and their impact on fairness and justice, but [the empirical evidence demonstrates](#) that shifting to such algorithmic application reduces overall racial bias and the harms of arbitrary policing.
- **Child Welfare:** In a recent empirical study of algorithmic tools that assist caseworkers in investigating child abuse or neglect, Amit Haim finds that the tools improved the abilities of case workers and lowered the risks of invasive investigations.
- **Pay Equity:** In [recent research](#) on gender and racial pay gaps, I show that to tackle longstanding, stagnating pay inequities, AI-based software should not only be voluntary but should become an industry gold standard, and be required by law.
- **Patenting:** Using AI can automate patent application drafting, error-checking, distinguishing filed patent claims from what came before, flagging mismatches between the specification and the claimed invention, providing first draft specification language, and placing a patent application in a better condition for being granted. [Expanded and equitable access](#) to technology can increase patent access, quality, efficiency, and equity.
- **Freedom of Information:** If agencies adopt clear rules requiring automation in scanning, digitizing, making searchable and findable government documents, the mandate of freedom of information will be more of a reality. In a recent [ACUS report](#), my co-authors and I call for the accelerated development of digital tools to realize the mandate of affirmative disclosure of agency materials.
- **Public Safety:** The Illinois bill Biometric Privacy and Security Act considered a near-total ban on Facial Recognition Technology (FRT) use by law enforcement. In August of 2023, however, Governor JB Pritzker vetoed the bill, citing concerns about its impact on public

safety and investigations. While Illinois' attempt at a broad ban fell short, the state still has some of the country's strictest laws regarding biometric data collection and use. California has enacted laws that temporarily prohibit the use of facial recognition technology in body cameras worn by law enforcement officers. Similarly, cities like San Francisco, Oakland, and Boston have passed ordinances banning the use of facial recognition technology by city departments, including police departments. Portland, Oregon, passed one of the most stringent bans, prohibiting private entities from using facial recognition technology in public places. The similar Facial Recognition and Biometric Technology Moratorium Act, proposed in Congress, seeks to impose a moratorium on federal use of the technology until certain conditions are met. Ironically, avoiding profiling, arbitrariness, and exclusion more often than not requires more complete data collection and publicly available datasets. The specific resistance to biometrics and facial recognition and the privileging of broad notions of individual privacy come at significant costs that have not yet received enough attention in law and legal scholarship.

Operationalizing Automation Rights

The adoption of AI advancements is not just a technological inevitability, but a societal responsibility. In private law - from torts to occupational safety regulations, anti-discrimination to environmental protection - doctrines about duty of care and standards about state-of-the-art safety and compliance are already embedded in existing laws. In public law, core values of constitutional law include equal treatment, consistency, and effective governance.

Automation rights would also include a focus on how to foster rational trust in AI. Beyond the assessment of the effectiveness and accuracy of the AI models themselves, law and policy should pay more [attention](#) to how individuals – users, passengers, patients, citizens – assess AI. The nascent field of behavioral human-machine interactions [indicates](#) the push and pull of the resistance to AI: driven by both the perception that AI is a mysterious “blackbox” and the illusion (perhaps delusion) that human decision-making is easy to understand. Both in lay and expert settings, people routinely overestimate their ability to perform and decide accurately, without bias. In experimental research, [patients](#) are willing to forego better health care to have a human, rather than an AI, decision-maker. Our current laws have likely contributed to such irrationalities. As I argued in [The Law of AI for Good](#), “requiring humans to be the final decision-makers in high stakes processes is not only a flawed solution in contexts where AI has clearly reached comparative advantages, but it also risks perpetuating irrational fears about AI instead of helping debias citizens about the comparative risks of technology.” In a new article, [Do We Need to Know about Artificiality: Unpacking Disclosure and Generating Trust in an Era of Algorithmic Action](#), I argue that there should not be a default overarching right to know that a decision or content was generated by AI. The right to know that you are interacting with a bot, or that you are subject to automated decision-making, is a centerpiece of EU/US legislative proposals. Both the EU's General Data Protection Regulation (GDPR) and the California Consumer Privacy Act (CCPA) already include the right to know if AI is making decisions about them and to request explanations for those decisions. Under the new [EU AI Act](#), consumers will have a right to see disclosures that they are

chatting with or seeing images produced by AI. Title IV, art. 52 states: “Providers shall ensure that AI systems intended to interact with natural persons are designed and developed in such a way that natural persons are informed that they are interacting with an AI system, unless this is obvious from the circumstances and the context of use.” Quebec similarly passed a [law](#) that requires individuals to be informed when automated decision-making tools are being used. Other [private](#) and public declarations about ethical AI similarly emphasize such disclosures about artificiality as a keystone of AI governance. In recent surveys, most Americans want to know when they are interacting with an [AI](#). And yet, I argue, the reasons for disclosing artificiality are complex and often in tension with other goals of generating rational decision-making, trust, and safety. AI law and policy should be based on analysis, not epithets. Automation rights are inevitable, but we have not even begun to conceptualize, develop terminology, and consider the complex design of the approaching legal landscape.