The Constitutionality of Algorithmic Decision Systems Output in Light of the Probable Cause Requirement of the Fourth Amendment

Lacey Strahm

Columbia College, Columbia University

American Studies, Senior Thesis

Professor Roosevelt Montás, Program Advisor

Professor Steven Bellovin, Subject Matter Advisor

Spring 2020

TABLE OF CONTENTS

1. INTRODUCTION

As we enter the Fourth Industrial Revolution, ushered in by ubiquitous cyber-physical systems, the Internet of Things, and the Internet of Systems, our legal precedents are lagging behind our technical innovations.[1] In his 2017 book, *The Fourth Industrial Revolution*, Klaus Schwab describes how this revolution is different from all the previous epochs. The First Industrial Revolution began in the latter half of the 18th century and focused on the power of the steam engine, which lead to the mechanization of many manufacturing industries. The Second Industrial Revolution began late in the 19th century and was characterized by scientific advancements such as electricity that enabled mass production. The Third Industrial Revolution began in the late 1950s and marked the shift from analog to digital through computing. The Fourth Industrial Revolution, Schwab explains, is happening now and does not resemble anything we have seen before. It is disrupting nearly every academic discipline, professional industry, and the global economy in a non-linear way. The Fourth Industrial Revolution is characterized by the way in which the physical, digital, and biological spheres of an individual's life are connecting at a speed previously unseen. The speed, quantity, and efficacy with which the Fourth Industrial Revolution is raging throughout the globe have strained the legal frameworks tasked with protecting citizens from the unintended consequences of such massive innovation.

Advances in Artificial Intelligence (AI), Machine Learning (ML), Algorithmic Decision Systems (ADS), and other approaches to training computers in decision-making are at the forefront of a new set of challenges to our legal frameworks. These systems are becoming so sophisticated that institutions are replacing human decision-makers with automated systems to boost efficiency and consistency. We see their dissemination in policing decisions, such as

[1] Schwab, Klaus. *The Fourth Industrial Revolution*. Portfolio Penguin, 2017.

in so-called 'predictive policing,' where police apply statistical or machine learning algorithms to data from police records to look for potential patterns that predict when, where or what crime may be committed.[2] We see their dissemination in courtroom decisions where judges use computerized systems risk assessments to inform who can be set free at every stage of the criminal justice process.[3] We see their dissemination in insurance decisions where state government agencies use automated systems to determine if a claimant has committed insurance fraud.[4] In short, we see their dissemination across a wide range of industries that utilize complex decision-making in their operation.

One important question that arises from these advances and their dissemination throughout society is whether and how these systems output can be used in legal proceedings. Under the Fourth Amendment of the United States Constitution, "no Warrants shall [be] issue[d], but upon probable cause." This stipulation means citizens are secure from searches and seizures of their persons, houses, papers, and effects unless probable cause can be established. In United States criminal law, probable cause is the standard of reasonable grounds by which police have a suspicion, supported by circumstances sufficiently strong to justify a prudent and cautious person's belief that specific facts are probably true.[5] An Algorithmic Decision System (ADS) could be built to address applications for search warrants, trained to look at evidence to establish whether probable cause exists and, produce an output of probable cause with given statistical likelihoods. Under this definition, could a

---

[2] Lacambra, S. (2017, October 30). Predictive Policing one pager. Retrieved from https://www.eff.org/document/predictive-policing-one-pager
[3] Angwin, J., Larson, J., Kirchner, L., & Mattu, S. (2019, March 9). Machine Bias. Retrieved from https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing
[4] Felton, R. (2016, December 18). Michigan unemployment agency made 20,000 false fraud accusations – report. Retrieved from https://www.theguardian.com/us-news/2016/dec/18/michigan-unemployment-agency-fraud-accusations
[5] Handler, J. G. (1994) Ballentine's Law Dictionary (Legal Assistant ed.). Albany: Delmar. p. 431. ISBN 0827348746.

judge consider the output of the aforementioned ADS to establish probable cause for a search warrant?

Legal scholarship has not yet definitively answered this question, although Orin Kerr, a professor of law at the UC Berkeley School of Law, wrote a compelling defense for keeping quantification out of the court room. Kerr suggests quantification specifically of "probable cause would eclipse intuitions and instead facilitate distortions of probability resulting from cognitive biases."[6] He reveals how quantification would lead to less accurate probable cause determinations because humans determine varying degrees of reasonableness through intuition, which is a feature of cognition machines lack. Quantification, Kerr explains, "would override the critical intuitions of judges about missing information in the affidavit that are critical to assessing probable cause accurately." Kerr's analysis hints at a human desire to quantify immaterial things that are not quantifiable, forcing such quantifications to use poor input or proxy variables that lead to inadequate measures of the outcome. This potent desire for reification forces immaterial things that no one has a good way of quantifying to be reduced to an outcome that does not always make sense in a particular context.

Think of doctors asking patients to rate the scale of their pain 1-10. Everyone's tolerance for pain is different depending on what injuries, ailments, or diseases a patient has been exposed to before. Yet the scale remains the same 1-10. An experienced athlete who had dislocated their knee before can give a new dislocation a grade of six when previously they gave the same injury a grade of eight. A fifth grader who caught the flu for the first time can also rate their pain as a six, whereas the mother who catches that same strain of the flu from her child can rate her pain a two. The flu and a dislocated knee are very different kinds of

---

[6] Kerr, Orin S., Why Courts Should Not Quantify Probable Cause (March 28, 2011). The Political Heart of Criminal Procedure: Essays on Themes of William J. Stuntz (Michael Klarman, David Skeel, and Carol Steiker, eds), pages 131-43 (Cambridge 2012). . Available at SSRN: https://ssrn.com/abstract=1797824

pain, yet they both can receive the same pain grade, revealing a shortcoming in the pain quantification system. It is the doctor’s responsibility to determine the reasonableness of the pain grade by utilizing their critical intuition contrived from their knowledge of the patient and the ailment to assess the pain accurately, and ultimately successfully do their job. If pain quantification were to eclipse doctor's critical intuition in diagnosing patients and guiding them to appropriate healing options, what would our health care system look like? This error-ridden vision of quantified-centric institutions is precisely the reason why Kerr's defense is so tenable.

Many immaterial things are multi-dimensional, so quantification becomes tricky when you try to compare multiple points, on multiple axis, in multiple planes. The question of what becomes your combining rule for these various combinations of points, axes, and planes would require a complete mathematical theory for how to go about solving this problem. However, the mathematician’s personal values would be baked into the framework of the theoretical structure because their choice in picking which variables to compute would be a direct reflection on their values. This speaks to the differing values humans can ascribe to a situation that complicates the problem even further, as it is highly unlikely two humans have the same value base. Humans can have similar values but tend to differ on nuanced contexts or situations that distinguish their value set from others.

I agree with Kerr’s disposition towards opposing quantification of probable cause. However, in the ten years since Kerr’s article advances in technology have made progress toward the opposite, making quantification of probable cause seem not only plausible but highly attractive in some spaces. I will take Kerr’s argument to keep quantification out of the courtroom a step further and contend that ADS output cannot meet the explainability standard required to establish probable cause in a legal proceedings, and thus should not be a viable option in the courtroom.[7] ADS provide black and white answers to color-filled questions with no explanation as to how they arrived at their conclusion. In the event the system does produce some form of an explanation, the explanations are often unintelligible to humans because the machine's infrastructure is not designed to justify its work. ADS have no mechanism to interpret and understand the cause of a situation nor explain how it arrived at its decision output. Therefore, a legal claim established from ADS does not meet the standard necessary for establishing probable cause under the Fourth Amendment. ADS cannot provide enough reliable evidence to justify the output and thus could not meet the explainability requirement needed for a search warrant.

2. TECHNICAL APPROACHES

The Fourth Industrial Revolution is profound because of the vast and far-reaching progress made in computing machinery and intelligence research. The fuel that has excited the modern explosion of innovation has been in place for years, finding its genesis largely in the work of Alan Turing, widely considered to be the father of theoretical computer science and Artificial Intelligence (AI).[8] Turing's seminal paper, *Computing Machinery and Intelligence*, centered around the question "Can machines think" where he explores this concept in detail and answers what he presumes may be common objections towards the idea.[9] Two years later in 1952, Claude Shannon, a Bell Labs researcher, shared one of the first examples of machine

[7] The explainability standard will be defined *infra* as we survey United States legal history. Reference explainability section.

[8] Beavers, Anthony (2013). "Alan Turing: Mathematical Mechanist". In Cooper, S. Barry; van Leeuwen, Jan (eds.). Alan Turing: His Work and Impact. Waltham: Elsevier. pp. 481–485. ISBN 978-0-12-386980-7.

[9] Turing, Alan (October 1950), "Computing Machinery and Intelligence" (PDF), Mind, LIX (236): 433–460, doi:10.1093/mind/LIX.236.433

learning with the world[10]. Theseus, a robotic maze-solving mouse, could 'remember' its path through telephone relay switches.

With trailblazing electromechanical devices such as Theseus emerging out of research labs, John McCarthy, an assistant Professor of Mathematics at Dartmouth college, recognized there was an opening for some development in the research area of thinking machines. In the summer of 1956 McCarthy organized the Dartmouth Summer Research Project on AI, where mathematicians, scientists, and people interested in the subject were invited to study features of intelligence that a machine can be made to simulate.[11] The U.S. government was particularly supportive of machine translation research at this time because Cold War politics fueled their heightened interest in automatically translating documents, particularly of Russian origin. In 1958, the U.S. Navy funded Frank Rosenblatt, the head of the cognitive systems section at Cornell Aeronautical Laboratory, in his perceptron project. The perceptron algorithm learned by trial and error using a specific kind of neural network that simulated the human thought process. Rosenblatt built the perceptron algorithm with the hopes of gaining insights into the human brain by organizing computer systems in a way that he believed mimicked the organization of the human brain.

In 1964 the United States government desired to evaluate the progress of computational linguistics and machine translations, so it set up a committee of seven scientists called the Automatic Language Processing Advisory Committee (ALPAC). The report subsequently produced from ALPAC's inquiry encouraged a more basic approach to computational linguistics and machine translation compelling the research discipline to take

[10] Klein, D. (2018, December 19). Mighty mouse. Retrieved from https://www.technologyreview.com/s/612529/mighty-mouse/
[11] McCarthy, J., Minsky, M., Rochester, N., Shannon, C.E., A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence., http://raysolomonoff.com/dartmouth/boxa/dart564props.pdf August, 1955

a step back and revaluate where it was focusing its energy.[12] Following ALPAC's findings in 1969, Marvin Minsky, a cognitive scientist best known at the time for co-founding the Massachusetts Institute of Technology's AI laboratory, published, *Perceptrons: An Introduction to Computational Geometry*, a book harshly criticizing the work of Frank Rosenblatt on his perceptron algorithm. This book introduced major controversy within the AI community centered around the fundamental divide between believing Minsky's pessimistic predictions on the limitations of the perceptron or hoping that Rosenblatt's work would usher in more research that could lead to ground-breaking innovation. Minsky's pessimism was able to win the debate for the time being, and U.S. support for AI research was drastically decreased.

Across the pond, the British government harbored similar reservations about the progress of AI research. James Lighthill's report compiled for the British Science Research Council in 1973 entitled, *Artificial Intelligence: A General Survey*, was a cynical prognosis of academic research in the AI field that confirmed the British crown's suspicions.[13] Now, both the U.S. government and the British government had ended general support for further academic research into AI, leading to a period of reduced funding known as an 'AI winter'.

Significant AI research would not resurrect until the 1980s when Japan's Ministry of International Trade and Industry (MITI), embarked on the Fifth Generation Computer Systems (FGCS) initiative to build supercomputers that could yield a platform for future advances in AI. In response the British Government used their substantial war chest to fund the Alvey Programme, which supported research in knowledge engineering and opened AI research in England again.[14] The U.S. government also reacted to the changing sentiments by

---

[12] ALPAC Report, Language and Machines — Computers in Translation and Linguistics. A Report by the Automatic Language Processing Advisory Committee, Washington, DC, 1966
[13] James Lighthill (1973): "Artificial Intelligence: A General Survey" in Artificial Intelligence: a paper symposium, Science Research Council
[14] Aleksander, Igor (2013). Decision and Intelligence, Volume 6. London: Kogan Page. p. 185. ISBN 9781850914075.

founding the Strategic Computing Initiative under the Defense Advanced Research Projects Agency (DARPA)[15] which tripled funding for AI research between 1984 and 1988. This period in AI research was centered around a rule-based approach where computer scientists would give the system a set of rules and constraints to follow and observe how the program operated under such conditions. A popular programming language associated with AI at the time was Prolog. Programs coded in Prolog expressed logic in terms of relations, represented as facts and were useful for particular tasks that benefited from rule-based logical queries such as searching databases.[16]

Some early AI systems required specialized hardware for its processing power. However, advances in hardware technology from companies such as Apple and IBM collapsed the market for such specialized hardware. As predicted by Moore's Law in 1965, the speed and memory capacity of computers doubles every two years so the fundamental problem of 'raw computer power' would gradually be overcome.[17] IBM used this shift to their advantage and became a leader in the AI industry, most notably for Deep Blue, their chess-playing computer that was the first computing system to defeat Garry Kasparov, the reigning world chess champion in 1996. Games became the new playground for researchers to test the capabilities of their AI-powered computing systems. DeepMind followed suit in 2015 with their creation of the first deep learning AI model to "successfully learn control policies directly from high-dimensional sensory input."[18] The model played seven Atari 2600 games and outperformed previous approaches to six of the games while surpassing human experts

---

[15] McCorduck 2004, pp. 426–432, NRC 1999 under "Shift to Applied Research Increases Investment"
[16] Kowalski, R. A. (1988). "The early years of logic programming" (PDF). Communications of the ACM. 31: 38. doi:10.1145/35043.35046
[17] Moore, G. E. (1965). Cramming more components onto integrated circuits. Electronics, 38(8). Retrieved from https://drive.google.com/file/d/0By83v5TWkGjvQkpBcXJKT1I1TTA/view
[18] Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Riedmiller, M., & Wierstra, D. (n.d.). Playing Atari with Deep Reinforcement Learning. DeepMind Technologies. Retrieved from https://arxiv.org/pdf/1312.5602v1.pdf

on three of the games, achieving a superhuman level of play.[19] The gaming programs and advances in computer vision that came out of this time period were early attempts at AI focused computing. However, they were limited in functionality because their logic was hard-coded, only being capable of executing specific tasks to accomplish specific goals such as playing chess.[20] Thus evolutionarily, these advances were a technological dead-end.

At the turn of the century, the concept and use of big data was popularized, and it revolutionized old conceptions about how researchers could approach computing. Big data typically refers to data sets with a size so great that they are beyond the ability of regular software tools to process the data within a reasonable amount of time. With so much data now available, researchers discovered they could get dramatically better performance out of neural networks with many layers as opposed to the few layers they were previously using. This breakthrough in deep neural network learning transformed research in AI as deep learning applications could be used across various industries. In the automotive industry, deep learning is pioneering the automated driver charge, helping vehicles automatically detect objects such as pedestrians, traffic lights, and street signs.

Big data is the foundation for today's machine learning methods, which are ubiquitous within individual's public and private lives. People not only interact with AI almost every day whether they are aware of it or not, but people also contribute to machine learning algorithms with their personal data. The Fourth Industrial Revolution was made possible through this growing frequency of human-computer interactions, but the road to getting here had been in place for decades.

---

[19] Ibid.
[20] The concept of hard-coded programs will be discussed *infra* in the Machine Learning section. Reference Machine Learning section.

A. Algorithmic Decision Systems

One of the technical approaches ubiquitous within our public and private institutions is Algorithmic Decision Systems (ADS). As a generic term, ADS encompass any deterministic rules-based algorithmic system. Algorithmic, in this context, refers to an explicit process or set of rules to be followed in calculations or other problem-solving operations. Algorithms, themselves, are finite and well-defined instructions that can be executed routinely. This broad definition does not explicitly require machine computing, although in society and in terms of this paper, algorithms usually refer to complex machine computing operations. This more conventional brand of machine-centric algorithms can be statically hand-coded by programmers or automatically generated from data input. In ADS, algorithms define the rules for which the system can then analyze high quantities of data to find correlations and parse relevant information out for decision-making.

When evaluating ADS, individuals must consider every piece of information that goes into an algorithm and contextualize it within the programming goal. ADS typically have training data, which is the initial set of data used to understand how to apply its algorithm to given data points. Similar to humans, algorithms learn from exposure and experience, so training data functions as algorithms' exposure and experience. If an ADS is built to estimate if there is probable cause for a police officer to search a vehicle, the training data could include the number of stops that turned into arrests, the number of stops that did not turn into arrests, and any other recorded data points that the coder feels necessary to include. They also have parameters that are values passed into the function, which set the bounds for what data the argument will evaluate. That same ADS built to estimate if there is probable cause for a police officer to search a vehicle would be passed a parameter such as 'arrest' that corresponds to the training data. The algorithm would have learned how to handle similar parameters and would instruct the system to evaluate the arrest data point based on its education. ADS also can have classifiers that relate given input to particular categories of output.[21] Sticking with our previous ADS example, classifiers in that system could be 'arrest' and 'no arrest,' which assign what the system evaluates into these two categories to then inform the output prediction. In considering how an ADS is constructed, observers can understand with more clarity how the system functions, for what purpose, and what recourse needs to be taken to rectify an error.

Humans are involved in the construction process of building the algorithms out by selecting appropriate training data for the system to model and correctly labeling data for supervised algorithmic systems. Coders who create the algorithm make judgment calls on what is or is not possibly relevant to ADS, constructing the algorithm's logic base on values they wish to codify. These values maybe personal, corporate, philanthropic or driven by some other factor that helps achieve the desired engineering goal. After this creation stage, human intervention in ADS output is not necessary and often entirely removed from the equation.[22] Once the algorithm is built, trained, and tested, it is ready to be deployed in the real world, subject to all the chaos that comes with reality.

Public discourse around ADS has increased through recent news regarding Roger Stone, a prominent Republican political consultant and lobbyist, who was convicted of seven felonies and sentenced to around four years in federal prison. Attorney General William Barr and the Department of Justice (DOJ) decided that Stone's convictions should remain in place,

[21] The concept of classifiers will be discussed *infra* in the Bad Classifier section. Reference Bad Classifier section.
[22] Castelluccia, C., & Le Métayer, D. Understanding algorithmic decision-making: Opportunities and challenges, Understanding algorithmic decision-making: Opportunities and challenges (2019). Retrieved from https://www.europarl.europa.eu/RegData/etudes/STUD/2019/624261/EPRS_STU(2019)624261_EN.pdf

but the sentence did not have to reflect the prosecution's recommendation of nine years.[23] In every federal criminal case involving a defendant who has been convicted, U.S. Sentencing Guidelines are used to prescribe sentences or punishment in broad terms with the goal of ensuring equity.[24] The guidelines offer a calculation of what the sentence should be within a range of months based on the offense and the characteristics of the offender. This range is then plotted on the U.S. Sentencing Table, which is a grid where the offense and offender characteristics are given numerical values. The offense numerical value is plotted on the vertical axis and can be increased or decreased by various factors such as threat of violence or pleading guilty. The offender's characteristics numerical value is plotted on the horizontal axis and can be increased or decreased by evaluating the defendant's criminal history with input such as how many prior convictions the offender has and how many crimes the offender has committed while on probation, to name a few. The point at which the offense level and criminal history category meet generates the recommended sentence.

Although humans currently administer the U.S. Sentencing Guidelines, this type of nonautomated algorithmic-based decision-making process lends itself as one of the more basic examples of what ADS can do and where ADS are deployed in society. However, dissimilar to many other ADS discussed below, the U.S. Sentencing Guidelines have explainability. Humans operate the guidelines, and fashion all judgments remaining at the heart of every decision. Hence the decisions produced from these guidelines are explainable to humans, empowering the U.S. Sentencing Guidelines to be a more acceptable form of ADS; despite their controversial record with overriding sentence guidelines. There are many other

[23] McCarthy, A. C. (2020, February 14). Roger Stone's sentence: How the judge will really decide to 'lock him up'. Retrieved from https://thehill.com/opinion/judiciary/483090-roger-stones-sentence-how-the-judge-will-really-decide-to-lock-him-up
[24] Ibid.

types of ADS percolating throughout individuals' everyday lives, but the most powerful and criticized models are Artificial Intelligence (AI) and Machine Learning (ML).

B. Artificial Intelligence

AI as defined by MIT Technology Review, is 'the quest to build machines that can reason, learn, and act intelligently.'[25] It is a quest that started with Turning and has been around for many decades but has barely scratched the surface of its potential. AI, as an umbrella term, is used to describe much of the innovation seen in machine intelligence today, although many technical approaches fall under this field of study. The theory behind AI is to develop computer systems that can perform cognitive functions conventionally associated with the human mind. Some examples of these functions include visual perception, learning, language translation, problem-solving, speech recognition, and, most important for this paper, decision-making. AI research seeks to display in machines an intelligence comparable to the natural intelligence possessed by human beings. Engineers model AI devices to be intelligent agents that can perceive their environment and take actions that maximize their chance of successfully achieving their goals.[26] AI as a broad field of academic study can be broken down into subfields based on technical applications. Since AI draws on so many different academic disciplines, the space for this division is vast and continually growing.

C. Machine Learning

Machine Learning (ML) is a practical subfield of AI that has garnered much attention recently for its advanced computing capabilities. On a technical level, ML is a method of data

---

[25] Artificial Intelligence. (n.d.). Retrieved from https://www.technologyreview.com/artificial-intelligence/

[26] Poole, David; Mackworth, Alan; Goebel, Randy (1998). Computational Intelligence: A Logical Approach. New York: Oxford University Press. ISBN 978-0-19-510270-3.

analysis that automates analytical model building.[27] It does this by iterating through data autonomously and adapting to new data by relying on previous computations. The algorithm is trained to learn what some unknown variable looks like. Then that association is assigned a numeric value through various linear algebra calculations and told to produce repeatable results within a given probability. Matrix operations can be used directly to solve key computations or provide the foundation to use more complex operations in the description of a machine learning method.[28]

To further explain how ML functions, consider Privee, a software architecture that analyzes website privacy policies. Privee uses ML to perform automatic classifications on inputted privacy policies by checking if the inputted privacy policy matches privacy policies in its repository.[29] If so, the policy is labeled with an overall letter grade that is based on the classification metrics it was trained on and displays the label to the user. If the inputted privacy policy does not match privacy policies in its repository, the policy is evaluated by either the rule classifier, the ML classifier, or both to determine the policy's classification.[30] Once the classification is determined the algorithm trains itself on the new policy and stores that classification with the other training policies it has learned, completing the nonlinear feedback loop. Then the policy is labeled and displayed to the user.

The foundational idea for ML arose when AI based systems were trying to solve the problem of how to address hard-coded programs.[31] Hard-coded programs are a software

---

[27] Machine Learning: What it is and why it matters. (n.d.). Retrieved from https://www.sas.com/en_us/insights/analytics/machine-learning.html

[28] Brownlee, J. (2019, August 9). A Gentle Introduction to Matrix Operations for Machine Learning. Retrieved from https://machinelearningmastery.com/matrix-operations-for-machine-learning/

[29] Sebastian Zimmeck and Steven M. Bellovin. Privee: An architecture for automatically analyzing web privacy policies. In 23rd USENIX Security Symposium (USENIX Security 14), pages 1--16, San Diego, CA, August 2014. USENIX Association.

[30] The concept of classifiers will be discussed infra in the Bad Classifier section. Reference Bad Classifier section.

[31] Khan, M., Jan, B., & Farman, H. (2019). Deep Learning: Convergence to Big Data Analytics. Springer Singapore. doi: 10.1007/978-981-13-3459-7

development practice of embedding data directly into the source code of a program as opposed to getting the data from external sources or generating it at run-time. This practice can be problematic because the static nature of hard-coded programs means they cannot adapt to innovation. An inability to adjust to change can result in becoming obsolete at the hands of such change, and hard-coded programs are experiencing such a fate. In modern computer science circles, hard-coded programs are seen as a failure of technology to predict and address problem test cases adequately. Thus, this failure led to a train of thought that envisioned placing machines in the driver's seat of extracting patterns from the data by themselves, freeing them from their previous static programming nature.

D. Supervised & Unsupervised Learning

Like AI in general, ML has a broad array of methods it can deploy in programs, two of them being supervised learning and unsupervised learning. Supervised learning uses data sets containing training examples with associated correct labels as prior knowledge to anticipate what the output will be.[32] The program learns the relationship between the training example and the associated correct label by identifying patterns in the data and forming heuristics. Heuristics are techniques designed to expedite the problem-solving process and find approximate answers. They are short-cut rules of thumb that guide decision-makers to a satisfactory solution, not an optimal or perfect one. In ML, heuristics are derived from the compilation of previous experiences that make up the algorithm's general information base.

Once a supervised learning program forms its heuristics, it can apply that understanding to new examples the machine has not seen before and emit a label for those

[32] Maini, V. (2018, May 28). Machine Learning for Humans, Part 2.1: Supervised Learning. Retrieved from https://medium.com/machine-learning-for-humans/supervised-learning-740383a2feab

new matches. The data presented in new examples can be a discrete or continuous value. If the data presented is a continuous value, meaning it can take any values, the system will parse it through a regression where the input is mapped to continuous output. Continuous, in this context, refers to the mathematical concept that between any two possible numbers there can always be another number. Between 6 and 7 is 6.5, between 6.5 and 6.6 is 6.55 and so on without exception. If the data presented is a discrete value, meaning it can take only a specific value, the system will parse it through classification where the input is mapped to output tags.[33] Discrete values have specific numeric or non-numeric values such as 6 or "book." There is no obvious way to merge, average, or combine these discrete categories because the values are independent points disconnected from each other.

Unsupervised learning starts with unlabeled data and performs learning tasks that output clusters of items that are similar to each other in some mathematical sense. It does this without trying to attach a label or particular name to any of the output clusters of items. Depending on the purpose of a project, unsupervised learning can distinguish pattern structures from input data through clustering, dimensionality reduction, and representation learning.[34] Clustering, a popular learning task in unsupervised learning, can group old and new data by similarity such that points in different clusters are dissimilar while points within a cluster are similar.[35] Dimensionality reduction, another popular learning task in unsupervised learning, tries to reduce the complexity of the data while keeping as much of the relevant structure as possible.[36]

---

[33] Ibid, footnote 27.
[34] Ibid, footnote 27.
[35] Maini, V. (2018, May 28). Machine Learning for Humans, Part 3: Unsupervised Learning. Retrieved from https://medium.com/machine-learning-for-humans/unsupervised-learning-f45587588294.
[36] Ibid.

The fundamental difference between supervised and unsupervised learning lies in the formation and understanding of the ground truth of the program. Unsupervised learning starts without a notion of ground truth and applies various mathematical techniques to draw patterns out in the data. The program bases its clusters off of a mathematical equation that it derives from inputs' similarity. Supervised learning, on the other hand, starts with a notion of ground truth that it is taught through training data and finds patterns in the data based on its training. This type of program attaches specific labels to unknown sets of input based on the clusters of output.

E. Neural Networks

A popular data modeling structure underlying many ADS frameworks is neural networks. The concept of neural networks was first proposed in 1943 by Warren McCullough and Walter Pitts, two University of Chicago researchers, in their paper, *A logical calculus of the ideas immanent in nervous activity*.[37] Research on neural networks followed the peaks and gullies that came with the AI winters of the 20th century but has enjoyed a massive resurgence recently thanks to the increased processing power of specialized graphics chips used for video on all modern computers.[38] Neural networks have neurons, inspired by biological neurons, that represent mathematical functions. Originally, neural networks were an attempt to model the brain, but today with, developed understandings of the brain, it is clear that the brain and its neurons are more complicated and work differently than scientists initially thought.

---

[37] McCulloch, W.S., Pitts, W. A logical calculus of the ideas immanent in nervous activity. Bulletin of Mathematical Biophysics 5, 115–133 (1943). https://doi.org/10.1007/BF02478259.
[38] Hardesty, Larry. "Explained: Neural Networks." MIT News, MIT News Office, 14 Apr. 2017, news.mit.edu/2017/explained-neural-networks-deep-learning-0414.

The neural network can consist of millions of simple processing nodes that are densely interconnected and organized into layers of nodes that are 'feed-forward,' meaning the data fed into them moves through in only one direction.[39] The neural network is organized into layers of these nodes with outputs from one layer connected to inputs of the next layer.[40] An individual node can be connected to several nodes in the layer beneath it and above it, receiving data from nodes below it and sending data to the above nodes.[41] When a network is active, a node can receive data, a different weighted number, over each of the connections beneath it. It then can multiply these numbers by the associated weight assigned to each of the incoming connections. After all the products of the incoming connections have been calculated, the node can add them together and determine if that calculated weight meets the threshold to send the weight up to its outgoing connections.

On a more general level, each neuron in a neural network has a mathematical function that calculates a weighted sum of its inputs which is then fed into a complex non-linear function. This function passes data through successive layers until it arrives radically transformed at the output layer. The structure learns by starting with random values set for its weights and thresholds. It takes sample input or training data and adjusts its neurons' weights based on the network's performance on this data. The weights and thresholds are adjusted continuously until the example input consistently yields homogenous output. The adjustment of its neurons' weights is a crucial component of a neural networks ability to learn, as the calculation of new weights is how the network improves itself. Thus, this component of neural networks is at the heart of modern advances in ML.

---

[39] Ibid.
[40] Lee, T. B. (2019, December 2). How neural networks work-and why they've become a big business. Retrieved from https://arstechnica.com/science/2019/12/how-neural-networks-work-and-why-theyve-become-a-big-business/.
[41] Ibid, footnote 37.

3. FOURTH AMENDMENT HISTORY

The Fourth Amendment to the United States Constitution was added as a part of the Bill of Rights on December 15, 1791, but its founding sentiments arose in the 1600s when colonists began reacting to Britain's abuse of power. 17-century American Colonists were typically well-educated Englishmen aware of their rights as British citizens. These rights covered the maxim "Every man's house is his castle" as demonstrated in *Semayne's case*, argued in 1604 which established the notion of a search warrant.[42] These assumed rights also included limits to executive power with respect to searches as revealed in *Entick v Carrington (1765)*, a landmark case in UK constitutional law which decreased the scope of executive power and established civil liberties for the people.[43] So, when colonists' homes were invaded under oppressive "writs of assistance" they felt Britain was abusing its warrant power, and this motivated them to include protections from such violations in the Bill of Rights.[44]

The Fourth Amendment states:

> The right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated, and no Warrants shall issue, but upon probable cause, supported by Oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized.[45]

---

[42] Semayne's Case All ER Rep 62 5 Co Rep 91 a Cro Eliz 908 Moore KB 668 Yelv 29 77 ER 194

[43] Entick v. Carrington (1765) 19 Howell's State Trials 1029; 95 ER 807

[44] "In order to enforce the revenue laws, English authorities made use of writs of assistance, which were general warrants authorizing the bearer to enter any house or other place to search for and seize "prohibited and uncustomed" goods and commanding all subjects to assist in these endeavors. Once issued, the writs remained in force throughout the lifetime of the sovereign and six months thereafter. When, upon the death of George II in 1760, the authorities were required to obtain the issuance of new writs, opposition was led by James Otis, who attacked such writs on libertarian grounds and who asserted the invalidity of the authorizing statutes because they conflicted with English constitutionalism. Otis lost and the writs were issued and used, but his arguments were much cited in the colonies not only on the immediate subject but also with regard to judicial review." "History." Legal Information Institute, Cornell Law School, www.law.cornell.edu/constitution-conan/amendment-4/history.

[45] United States Constitution Amendment IV

This language addressed the colonist' ultimate concern of protecting citizens' privacy and safeguarding their freedom from unreasonable intrusions by the government. The scope of this amendment does not include protection from all searches and seizures, only searches and seizures that can be seen as an abuse of power by the government by not meeting the "probable cause" standard to the satisfaction of a neutral magistrate. This provision carries the original sentiments of the founders who were responding to direct government abuses of power and were not challenging the general grounds of government searches and seizures. Protected warrantless searches and seizures include instances where an officer asks for and is given consent to search, searches that are incident to a lawful arrest, probable cause, and exigent circumstances.[46]

For a citizen's Fourth Amendment right to have been violated, the citizen must prove that a justifiable expectation of privacy was arbitrarily violated by the government. Proving this arbitrary violation can be difficult when the violation arose from probable cause, a standard vague by design.

A. Probable Cause

One of the first definitions of probable cause was put forth by Chief Justice Marshall in *United States v. Aaron Burr (1807)*: "I understand probable cause to be a case made out by proof furnishing good reason to believe that the crime alleged has been committed by the person charged with having committed it."[47] This early definition laid the groundwork for the idea that probable cause is inextricably linked with "good reason" or "reasonableness," as the modern term encapsulating the iteration of this concept. However, after this early and still

[46] Kim, J. (Ed.). (2017, June). Fourth Amendment. Retrieved from https://www.law.cornell.edu/wex/fourth_amendment
[47] United States v. Burr, 25 F. Cas. 2, 1807 U.S. App. LEXIS 489 (Circuit Court, D. Virginia April 1, 1807).

vague definition, the Courts did not enforce probable cause under the Fourth amendment until about 150 years after their inception. The Bill of Rights stood as a formal declaration of federal rights and thus as held in *Barron v. Baltimore, 32 U.S. 243 (1833),* the Bill of Rights only applied to the federal government, not the states.[48] This position changed in the latter half of the 19th century when the Courts started interpreting the Fourteenth Amendment after its adoption in 1868.

The Fourteenth Amendment was born out of concerns regarding citizenship rights and equal protection under the law as they related to former slaves following the Civil War. As one of the reconstruction amendments, the Fourteenth Amendment was monumental in broadening federal enforcement within state boundaries through the Due Process Clause and the Equal Protection Clause. The Fourteenth Amendment states:

> All persons born or naturalized in the United States, and subject to the jurisdiction thereof, are citizens of the United States and of the state wherein they reside. No state shall make or enforce any law which shall abridge the privileges or immunities of citizens of the United States; nor shall any state deprive any person of life, liberty, or property, without due process of law; nor deny to any person within its jurisdiction the equal protection of the laws.[49]

The Due Process Clause prohibited states and local governments from depriving persons of life, liberty, or property without a fair procedure. The Equal Protection Clause required each state to provide equal protection under the law to all people, including all non-citizens, within

[48] Barron v. Baltimore., 32 U.S. 243, 8 L. Ed. 672, 1833 U.S. LEXIS 346 (Supreme Court of the United States February 16, 1833, Decided).
[49] United States Constitution Amendment XIV

its jurisdiction. The Supreme Court interpreted both of these clauses in conjunction with the entirety of the Fourteenth Amendment to incorporate most of the Bill of Rights as applicable legislation to be applied to the states as it is to the federal government. This new interpretation was reflected in *Chicago, Burlington and Quincy Railroad v. City of Chicago (1897)* and in *Gitlow v. New York (1925)*. Both cases applied Bill of Rights protections against the states and it was upheld.

With this new legal precedent, the temperance movement helped institute one of the most notable broad reaching federal crimes due to the Eighteenth Amendment, by banning the production, transport, and sale of intoxicating liquors. Before the Prohibition Era, crimes were mainly contained at the state level. The Courts did not see many cases reach the federal level, and thus very few federal criminal offenses had been tried, so little case law existed. Nevertheless, prohibition ushered in a new era of jurisprudence with the difficulties of nationwide enforcement taking their toll on cities. To help enforce the Eighteenth Amendment, police officers leaned on the Fourth Amendment's probable cause stipulation for help. Although prohibition formally ended in 1933, officers continued to use the probable cause stipulation to help enforce new liquor laws that arose.

One of the landmark decisions that came out of this was *Brinegar v. United States, 338 U.S. 160 (1949)*. Brinegar had a reputation for illegally transporting liquor and drove past an officer parked on the highway with a vehicle that appeared 'heavily loaded.' The officer stopped Brinegar, alleged that he saw liquor in the front seat of the car although Brinegar denied this allegation, and arrested Brinegar, seizing all the alcohol in the vehicle as it was in violation of the Liquor Enforcement Act of 1936. Brinegar challenged his arrest under the

Fourth Amendment asserting the officer did not have a search warrant for the evidence used against him.

The Court held that, "Probable cause exists where the facts and circumstances within the officers' knowledge, and of which they have reasonably trustworthy information, are sufficient in themselves to warrant a belief by a man of reasonable caution that a crime is being committed."[50] Thus the facts presented to the officers who pulled Brinegar over were sufficient to establish probable cause for the search and seizure hence they were admissible in the trial. This definition of probable cause aligns nicely with Chief Justice Marshall's 1807 definition making the substance for all explanations of probable cause a "reasonable" ground for belief of guilt, up until this point in legal history.

B. Aguilar v. Texas

Another landmark case shaping the probable cause definition came in 1964 when Nick Alford Aguilar's home was searched for narcotics on a warrant that had been issued. The warrant was based on an affidavit stating that officers had received reliable information from a credible person.[51] In *Aguilar v. Texas, 378 U.S. 108 (1964)* Justice Goldberg held:

> The point of the Fourth Amendment, which often is not grasped by zealous officers, is not that it denies law enforcement the support of the usual inferences which reasonable men draw from evidence. Its protection consists in requiring that those

[50] Brinegar v. United States, 338 U.S. 160, 69 S. Ct. 1302, 93 L. Ed. 1879, 1949 U.S. LEXIS 2084 (Supreme Court of the United States June 27, 1949, Decided).
[51] Aguilar v. Tex., 378 U.S. 108, 84 S. Ct. 1509, 12 L. Ed. 2d 723, 1964 U.S. LEXIS 994 (Supreme Court of the United States June 15, 1964, Decided).

> inferences be drawn by a neutral and detached magistrate instead of being judged by the officer engaged in the often-competitive enterprise of ferreting out crime.[52]

Justice Goldberg is keen to highlight that the process of getting a warrant approved by a judge is embedded into the Fourth Amendment as a safeguard for maintaining the integrity of the investigative process. This distinction forces the warrant process to be subject to objective review by an entity not directly involved with the criminal proceeding, which preserves the founding vision of the Fourth Amendment as a mechanism to stifle government abuse of power. However, in the case of probable cause, which stands as an exception to this safeguard, reasonableness stands at the center of determining whether the search was justified. If an officer determines they have probable cause to search a petitioner, and the petitioner later files to suppress based on this officer's probable cause determination, the judge presiding over the case will assess the reasonableness of the officer's probable cause claim, bearing in mind an officer's ability to abuse the governmental power inherent within their job.

Justice Goldberg understood this reasonableness requirement to establish probable cause and expanded upon this definition by detailing what information can provide the reliability needed to affirm reasonableness. He explains:

> Although an affidavit supporting a search warrant may be based on hearsay information and need not reflect the direct personal observations of the affiant, the magistrate must be informed of some of the underlying circumstances from which the informant concludes that contraband, such as narcotics, is where he claims it is, and some of the underlying circumstances from which the officer concludes that the

[52] Ibid.

informant, whose identity need not be disclosed, is credible or his information reliable.[53]

Thus, officers must produce some sort of evidence or documentation that the information an informant is providing is reliable. The form of evidence may vary dependent on the case's specific factual background, but the essence of the evidence must be substantial enough to convince a reasonable person that a crime is being committed where their information suggests.

This holding established a legal guideline for evaluating the validity of probable cause: the magistrate must know why an informant is credible and on what underlying circumstances this reliable informant relied on when providing the information. The added layer of consideration this holding provided to the reasonableness standard that already existed for probable cause determinations is significant when determining the reliability of ADS to establish probable cause. Through understanding the dubious black box infrastructure of many ADS, can the resulting output be considered a reliable informant under this standard?[54]

C. Spinelli v. United States

In *Spinelli v. United States, 393 U.S. 410 (1969),* the courts added another level of consideration to the judicial guideline created in *Aguilar* by introducing a "sufficiency" component to an officer's affidavit explanation. Similar to Aguilar, William Spinelli was suspected to be partaking in criminal activity, illegal interstate gambling, and evidence

[53] Ibid.
[54] The concept of black box infrastructure will be defined infra in the Black Box Properties section. Reference Black Box Properties section.

uncovered by a warranted FBI search was used against him in trial. The affidavit that authorized the FBI's search warrant was informed by a confidential reliable informant.

Justice Harlan delivered the opinion of the Court which held "the informant's tip, an essential part of the affidavit in this case, was not sufficient (even as corroborated by other allegations) to provide the basis for a finding of probable cause that a crime was being committed."[55] He supported this holding, stating:

> The tip was inadequate under the standard of Aguilar, *supra*, since it did not set forth any reason to support the conclusion that the informant was 'reliable,' and did not sufficiently state the underlying circumstances from which the informant had concluded that petitioner was running a bookmaking operation or sufficiently detail his activities to enable the Commissioner to know that he was relying on more than causal rumor or general reputation.[56]

Spinelli's case raised the question of how to handle insufficient justification provided in an affidavit, particularly concerning when the information is fully or partially corroborated by independent sources. In Spinelli's case, the affidavit did not give any explanation for why the tip should be considered reliable. It merely corroborated other allegations, and thus the judge had no reason to believe the evidence because it could not pass the Aguilar requirement on its own.[57]

---

[55] Spinelli v. United States, 393 U.S. 410, 89 S. Ct. 584, 21 L. Ed. 2d 637, 1969 U.S. LEXIS 2701 (Supreme Court of the United States January 27, 1969, Decided).
[56] DRAPER v. UNITED STATES, 358 U.S. 307, 79 S. Ct. 329, 3 L. Ed. 2d 327, 1959 U.S. LEXIS 1607 (Supreme Court of the United States January 26, 1959, Decided).
[57] Ibid, footnote 45.

The *Spinelli* holding added to the Aguilar requirement by taking the judicial standard a step further and requiring the magistrate to understand how the informant concluded that a crime had been committed. This development created a two-pronged test known as the Aguilar-Spinelli test which basically established an explainability requirement to satisfy probable cause. This is significant when determining the reliability of ADS to establish probable cause under the Aguilar-Spinelli judicial guideline because if ADS cannot explain in an intelligible way how it produced a given output, how would the output suffice under this standard?

D. Ybarra v. Illinois

After *Spinelli*, *Ybarra v. Illinois, 444 U.S. 85 (1979)* extended the judicial guideline for establishing probable cause to include an individualized suspicion requirement, such that officers must have a particularized belief with respect to the person to be searched or seized. Ventura Ybarra was in a tavern where police obtained a search warrant to look for evidence of possession of controlled substances. The officers decided once in the tavern that they would search all patrons present, and upon frisking Mr. Ybarra, officers felt a cigarette pack that ended up containing heroin. The officers charged Mr. Ybarra with unlawful possession of a controlled substance and Mr. Ybarra challenged his conviction.

Justice Stewart delivered the opinion of the court holding:

> Even though police possess a warrant based on probable cause to search a location in which a person happens to be at the time the warrant was executed, a person's mere propinquity to others independently suspected of criminal activity does not, without more, give rise to probable cause to search that person. Where the standard is

> probable cause, a search or seizure of a person must be supported by probable cause particularized with respect to that person. This requirement cannot be undercut or avoided by simply pointing to the fact that coincidentally there exists probable cause to search or seize another or to search the premises where the person may happen to be.[58]

This particularization requirement added another layer of depth onto the probable cause definition by requiring individualized suspicion. Ric Simmons, the Chief Justice Thomas J. Moyer Professor for the Administration of Justice and the Rule of Law at Moritz College of Law at the Ohio State University, describes this particularity requirement as "not merely a statistical likelihood that a suspect is guilty based on his membership in a certain group, but a reference to particular characteristics or actions by the suspect that shows that he specifically is likely to be guilty."[59] "Demographic probabilities" as Arnold H. Loewy, the Judge George R. Killman Jr. Chair of Criminal Law at the Texas Tech School of Law, calls it "are insufficient to create probable cause or reasonable suspicion; the police must also notice something specific to the defendant to create the probability as to him.'"[60] The presence of additional factors that are specific to what the suspect does in terms of the case, not who the suspect is, becomes a critically important concept for establishing probable cause under the Fourth Amendment. This is significant when determining the ability of ADS to establish probable cause under the particularized suspicion requirement of the Fourth Amendment because if ADS is

---

[58] Ybarra v. Ill., 444 U.S. 85, 100 S. Ct. 338, 62 L. Ed. 2d 238, 1979 U.S. LEXIS 151 (Supreme Court of the United States November 28, 1979, Decided).
[59] Simmons, Ric, Quantifying Criminal Procedure: How to Unlock the Potential of Big Data in Our Criminal Justice System (July 29, 2016). 2016 Mich. St. L. Rev. 947 (2016); Ohio State Public Law Working Paper No. 362. Available at SSRN: https://ssrn.com/abstract=2816006
[60] Arnold H. Loewy, Rethinking Search and Seizure in a Post-9/11 World, 80 MISS. L.J. 1507, 1518 (2011)

fundamentally based on generalized statistical probabilities and likelihoods, how could its output satisfy this standard?

E. Illinois v. Gates

By 1979, the probable cause standard encompassed a reasonable, explainable, and individualized ground for an individual to believe that another person was guilty of some crime. However, *Illinois v. Gates, 462 U.S. 213 (1983)* reassessed this standard by taking a closer look at the rigid two-pronged test instituted after *Aguilar* and *Spinelli*. The Bloomingdale, Illinois Police Department received an anonymous letter that alleged Lance and Susan Gates were trafficking drugs. The letter stated when the drugs were being moved, how they were being moved, and where the Gates kept the drugs, among other things. Police officers acted on the anonymous letter's tips and confirmed the drug trafficking allegations.

A search warrant was obtained based on the Bloomingdale police officer's affidavit which included a copy of the anonymous letter. Officers searched the Gateses' home and automobile to find the drugs and other contrabands. Prior to the trial, the Gateses moved to suppress evidence seized during the search, and the trial court approved, ordering the suppression of all items seized. The Illinois Appellate Court affirmed this decision on the holding that the anonymous letter and affidavit were inadequate to sustain a determination of probable cause for issuance of a search warrant under *Aguilar* and *Spinelli* since they failed the two-pronged test.[61] The anonymous letter failed the test because it "provides virtually nothing

[61] Illinois v. Gates, 462 U.S. 213, 103 S. Ct. 2317, 76 L. Ed. 2d 527, 1983 U.S. LEXIS 54, 51 U.S.L.W. 4709 (Supreme Court of the United States June 8, 1983, Decided).

from which one might conclude that its author is either honest or his information is reliable [and] gives absolutely no indication of the basis for the writers' predictions regarding the Gateses' criminal activities."[62]

After receiving briefs and hearing oral arguments regarding the questionable validity of the Bloomingdale Police's search warrant, Justice William Rehnquist delivered the decision of the Supreme Court in favor of the State of Illinois. Justice Rehnquist questioned:

> Whether the rule requiring the exclusion at a criminal trial of evidence obtained in violation of the Fourth Amendment, should to any extent be modified, so as, for example, not to require the exclusion of evidence obtained in the reasonable belief that the search and seizure at issue was consistent with the Fourth Amendment.[63]

This line of questioning lead to the decision to overturn the rigid two-pronged test established under *Aguilar* and *Spinelli* in favor of the more flexible "totality of circumstances" approach. This approach placed significant value on independent police work that corroborated details of an informant's tip, thus shifting the probable cause standard to a "fair probability" on which a reasonable and prudent person would act.

The adjusted standard to establish probable cause under the Fourth Amendment created in *Gates* reflects a shifting affinity towards considering reasonable intuition. A qualitative legal standard such as "fair probability" allows judges to account for facts missing in police affidavits instinctively. Under this definition, probable cause means, after assessing the totality of circumstances, it is with a fair probability that an individual can assume

---

[62] Ibid.
[63] Ibid.

contraband or evidence of a crime will be found if searched. The judicial privilege of accounting for missing facts intuitively is important for establishing probable cause under the Fourth Amendment because it values human instinctive feeling over conscious intelligent reasoning when determining probable cause. This is significant when determining the capability of ADS to establish probable cause under the totality of circumstances approach because, if ADS can only simulate conscious intelligent reasoning and not human instinctive feeling, how could its output fulfill this approach?

F. Ornelas v. United States

Ornelas v. United States, 517 U.S. 690 (1996), similar to *Gates*, brought up another new perspective that compelled an introspective look at the then current probable cause definition. Detective Michael Pautz of the Milwaukee County Sheriff's Department noticed a car with California license plates in a motel parking lot. He recognized the car as a popular model for drug couriers because it was easy to hide drugs within the interior and noted that California was a "source state" for drugs. He radioed his dispatcher to inquire about the car's owner and was informed the car was registered under either Miguel Ledesma Ornelas or Miguel Ornelas Ledesma.

Upon checking the motel registry, Detective Pautz discovered that an Ismael Ornelas accompanied by a second man checked into the motel with no reservations. He called his partner, Detective Donald Hurrle, who came to the scene to assist Detective Pautz. Together they contacted the local office of the Drug Enforcement Administration (DEA) and asked them to run Ismael and Miguel Ornelas names through the Narcotics and Dangerous Drugs Information System (NADDIS), which is a federal database of known and suspected drug traffickers. Both names appeared in the databases confirming Miguel and Ismael were drug dealers. The officers then summoned Deputy Luedke and the department's drug-sniffing dog, to the scene to replace Detective Pautz.

Detective Hurrle and Deputy Luedke waited at the scene until the petitioners emerged from the motel and got in the car that Detective Pautz had originally noticed. Detective Hurrle and Deputy Luedke approached the car and asked if the petitioners had any illegal drugs or contraband in their possession to which the petitioners answered no. Detective Hurrle asked for identification and was given two California drivers licenses identifying the petitioners as Saul Ornelas and Ismael Ornelas-Ledesma. After confirming their identities Detective Hurrle asked if he could search the car and the petitioners consented. Deputy Luedke noticed a panel above the right rear passenger armrest had a screw that was rusty, which indicated to him that it had been removed at some time. He dismantled the panel and discovered drugs, prompting an arrest of the petitioners.

The petitioners filed pretrial motions to suppress the evidence found during the search alleging the officers violated the petitioners' Fourth Amendment rights by conducting the search without a warrant and detaining them in the parking lot. Chief Justice Rehnquist upheld the magistrate's ruling, which acknowledged the consent given by the petitioners to search the car did not authorize the officers to search inside the panel under Seventh Circuit precedent. They also acknowledged that when the officers approached the petitioner's car, a reasonable person would not have felt free to leave the scene, so the statement of consent could be considered coerced, and the encounter is considered an investigatory stop. Investigatory stops are permissible under the Fourth Amendment if they are predicated by reasonable suspicion to stop the vehicle, and probable cause to perform a subsequent warrantless search of the vehicle. Therefore, for the warrantless search to be legal in the absence of valid consent, a neutral magistrate must support an officer's claims to reasonable suspicion and probable cause.

In arriving at this holding, Justice Rehnquist explains the court's standard for reasonable suspicion and probable cause by stating:

> articulating precisely what 'reasonable suspicion' and 'probable cause' mean is not possible. They are commonsense, nontechnical conceptions that deal with the factual and practical considerations of everyday life on which reasonable and prudent men, not legal technicians, act… The principal components of a determination of reasonable suspicion or probable cause [for investigatory stops and warrantless car searches] will be the events which occurred leading up to the stop or search, and then the decision whether these historical facts, viewed from the standpoint of an objectively reasonable police officer, amount to reasonable suspicion or to probable cause.[64]

This standard follows the totality of circumstances approach which takes all-things-considered into a determination of probable cause from the perspective of an objectively reasonable police officer. Police officers are subjected to many variations of crimes in their line of work that allow them to look at the historical facts of an incident and infer varying levels of guilt from their experience. This inference from experience allows officers to operate within a similar framework that is used in the judicial branch of government, as judges rule based on precedents set in previous cases in order to maintain fair application of the law among all petitioners.

---

[64] Ornelas v. United States, 517 U.S. 690, 116 S. Ct. 1657, 134 L. Ed. 2d 911, 1996 U.S. LEXIS 3391, 64 U.S.L.W. 4373, 96 Cal. Daily Op. Service 3744, 96 Daily Journal DAR 6059, 9 Fla. L. Weekly Fed. S 617 (Supreme Court of the United States May 28, 1996, Decided).

In this case the factual background that informed the police officer's inference to decide whether probable cause existed was the NADDIS data, the model of the car, the issuing state of the license plate, the location of incident, the time of year, the nature of the motel check-in, the body language of the suspect, and finally the loose car door panel. The confluence of these factors created the context under which a reasonable police officer could draw an inference of guilt based on their experience in such a line of work, whereas a layman could view the confluence as mere coincidence and the loose panel as general automotive wear and tear that comes with time. This is significant when determining the capability of ADS to establish probable cause under the totality of circumstances approach because, if ADS was given the same factual background to establish the necessary case-specific context, would it be able to make a reasonable police officer inference, or would it fall short and view the scene with a layman's perspective?

G. Florida v. Harris

Florida v. Harris, 568 U.S. 237 (2013) was another case that compelled an introspective look at the then current probable cause understanding with specific regard to who or what can influence probable cause. The case centered on the reliability of a narcotics dog during a routine traffic stop. Officer Wheetley pulled over Clayton Harris for a routine traffic stop, prompted by an expired license plate, when he noticed an open beer can and Harris's nervous demeanor. This prompted Officer Wheetley to ask Harris's consent to search the vehicle to which Harris refused. Officer Wheetley proceeded to execute a "free air sniff" test with Aldo, his narcotics trained dog, who alerted at the driver's-side door handle. Aldo's alert led Officer Wheetley to conclude he had probable cause to search Harris's vehicle, which turned out not to contain any of the substances Aldo was trained to detect. However, Officer Wheetley did find ingredients for manufacturing methamphetamine and arrested Harris on illegal possession of those ingredients. Harris was released on bail and in a subsequent stop, prompted by a broken brake light, was administered another "free air sniff" test by Aldo. Aldo again alerted at the driver's-side door handle but nothing of interest was found this time.

In Harris's suppression hearing, his attorney focused on Aldo's performance in the field rather than Aldo's extensive drug detection training and its respective merit. The trial court denied Harris's motion to suppress the evidence on the grounds they believed the officer had probable cause, but the Florida Supreme Court (FSC) reversed the decision and held that in every case the officer must present an exhaustive set of records outlining the dog's reliability, specifically the dogs field performance records. Without the dog's field performance records, the FSC held that an officer would be unable to establish probable cause to search the vehicle. This holding is the antithesis of the totality-of-the-circumstances approach and is inconsistent with the "flexible, common-sense standard" of probable cause previously established by the Supreme Court of the United States, thus the ruling of the FSC was reversed.

The FSC held, "[W]hen a dog alerts, the fact that the dog has been trained and certified is simply not enough to establish probable cause."[65] To prove a dog's reliability, the FSC believed that more supporting evidence needed to be produced such as:

> the dog's training and certification records, an explanation of the meaning of the particular training and certification, field performance records (including any unverified alerts), and evidence concerning the experience and training of the officer

---

[65] Florida v. Harris, 568 U.S. 237, 133 S. Ct. 1050, 185 L. Ed. 2d 61, 2013 U.S. LEXIS 1121, 81 U.S.L.W. 4081, 24 Fla. L. Weekly Fed. S 18 (Supreme Court of the United States February 19, 2013, Decided).

> handling the dog, as well as any other objective evidence known to the officer about the dog's reliability.[66]

The FSC continued by stressing the need for “evidence of the dog's performance history,” including documentation revealing “how often the dog has alerted in the field without illegal contraband having been found.”[67] False positives, as believed by the FSC, could help to expose dangerous confounding factors that influence a drug detection dog’s ability to accurately do its job and thus problematizes its reliability.

Justice Kagan delivered the opinion for a unanimous Supreme Court of the United States in which it disagreed with the FSC’s holding based on the rigid evidentiary checklist it required to establish probable cause. She notes that “an alert cannot establish probable cause under the Florida court's decision unless the State introduces comprehensive documentation of the dog's prior ‘hits’ and ‘misses’ in the field.” Justice Kagan contests this notion on the precedent established in previous cases where a gap in factual background for establishing probable cause can be compensated for by other strong indicators of reliability. She continues by refuting the false positives argument presented by the FSC by introducing the concept of false negatives that would be impossible to capture in field data. False negatives could be if a dog failed to alert to a car containing drugs, if a dog alerts for a spot where drugs were previously held in a car that still has residual odor, or if a dog alerts to a car containing drugs but the officer cannot find them and thus deems there are no drugs in the car, among other things. Through this Justice Kagan demonstrates how standard training in controlled environments makes a better metric for a dog’s performance because confounding variables

[66] Ibid.
[67] Ibid.

are limited. This is all to say a flexible all-things-considered approach to probable cause is best and remains the best approach because it gives police and magistrates proper discretion to do their job without bureaucratic handicaps.

Moreover, Justice Kagan added that, "a defendant must have an opportunity to challenge such evidence of a dog's reliability, whether by cross-examining the testifying officer or by introducing his own fact or expert witnesses" and they "may contest training or testing standards as flawed or too lax, or raise an issue regarding the particular alert."[68] If all these facts viewed by a reasonably prudent person would lead them to conclude that a search would uncover evidence of a crime, then probable cause can be established. This is significant when determining the capability of ADS to establish probable cause under the flexible all-things-considered approach because, if a defendant has the right to interrogate ADS reliability, how would it go about such a task considering the technical limitations of current technology?[69]

H. Explainability

From the signing of the Bill of Rights to the 21st century, the probable cause definition has grown, shifted and adjusted to the changing times. *Brinegar* determined probable cause to be a reasonable ground for belief of guilt. *Aguilar* and *Spinelli* expanded this to include explainability, making the standard for establishing probable cause include an explanation of the reasonable ground for belief of guilt under their two-pronged test. *Ybarra* went a step

---

[68] Ibid. The courts are divided on whether or not defendants should have the right to evaluate source code independently. It is a legal issue currently being debated. See this paper for more discussion on the topic. Steven M. Bellovin, Matt Blaze, Susan Landau, and Brian Owsley. Seeking the source: Criminal defendants' constitutional right to source code. Ohio State Technology Law Journal, 17(1), December 2020. To appear.
[69] Limitations of current technology will be discussed *infra* in the Limitations of Technical Approaches section. References Limitations of Technical Approaches section.

further and added an individualized suspicion requirement. Then, *Gates* took a step back and reexamined the probable cause definition, abandoning the two-pronged Aguilar-Spinelli test for the totality-of-circumstances approach that required a fair probability on which a reasonable and prudent person would act to establish probable cause. *Ornelas* followed suit with the totality-of-circumstances approach taking all-things-considered into its determination of probable cause. *Harris* rounded the probable cause definition out by encapsulating it in a flexible all-things-considered approach while introducing nuanced factors into the equation.

Together these cases reveal that at the heart of probable cause determinations is the idea of a rational agent explaining why a belief of guilt is reasonable. Explainability is a core function of our legal system intrinsic to its operations, exhibited through the structure of our courts having petitioners explain to the judge what happened, attorneys explaining to juries their positions, and judges explaining to petitioners why the outcome of their cases resulted in their innocence or guilt. In terms of determinations of probable cause, a rational agent's ability to explain the question of why persuasively to another rational agent is the key element in the determination that probable cause exists. The explanation must be sophisticated enough to convince another person that the determination is correct but reasonable enough that the other person can understand the rationale. If an individual comes up with a sophisticated explanation that stems from an advanced understanding of a given truth, but others cannot understand this sophisticated explanation, can this individual's explanation be validated as true? If an explanation is so complicated that only a specialist can understand the rationale, is this explanation considered permissible?

Considering a defendant must be convicted by a jury of his peers and not specialists, accessibility is an essential component of the explainability standards upheld in the courts. Our trial by jury system demonstrates accessibility as an embedded value of explainability in our legal system. Explanations that are accessible to a common man and a specialist alike are the standard by which they can constitute probable cause.

4. LIMITATIONS OF TECHNICAL APPROACHES

Despite all the recent groundbreaking innovations in AI research and computer science, there are still significant limitations that handicap the legal system's ability to trust ADS wholeheartedly. Foremost among them is a failure to meet the explainability standard required to establish probable cause. ML models, beneath all the advanced programming, are simply instantiations of a predetermined policy. In light of recent efforts to solve this predetermined policy problem, ADS remain error-ridden when faced with edge cases and complexity beyond what the model was designed to tackle.

Take satire as an illustrative example of an edge case with which an algorithm might be faced. Twitter could have an ML algorithm that scans its feeds for propaganda. It comes across an article entitled 'CIA's Facebook' Program Dramatically Cut Agency's Costs.'[70] Without the proper context, the algorithm could label this piece as propaganda because it knows Facebook is a technology company, not a CIA program, and thus the article is making a false claim. However, what it misses is the nuanced satirical nature of the piece, criticizing Facebook's invasive practices that violate an individual's privacy. ML cannot distinguish

[70] "CIA's 'Facebook' Program Dramatically Cut Agency's Costs." The Onion, The Onion, 18 Oct. 2017, www.theonion.com/cias-facebook-program-dramatically-cut-agencys-costs-1819594988.

complex and nuanced human concepts such as satire and irony from propaganda and hatred because algorithms lack perspective. These concepts do not fit nicely within the puzzled pattern of human life, meaning they do not follow rigid rules. Rather, they are fluid ideas that change based on the person using or viewing them. Individuals can disagree on where the boundary lies between categorizing a piece of work as propaganda or satire, making confusion an endemic part of these concepts' construction. Thus, if humans can be confused by such concepts, coding algorithms with predetermined policies to detect these concepts is a difficult task.

This illustrative example highlights ADS' core operational scheme which is designed to find the fact patterns upon which they were trained to act. When straightforward fact patterns ADS were trained to find exist, ADS work wonderfully. However, when a fact pattern does not exist or is hard to determine, ADS are horribly conservative, sticking to their pattern despite potential erroneous output. In these instances, the algorithm must fill in its knowledge gap by generalizing between the policy implied by its training data and the new case. The systems are designed to be accurate, not creative, so they struggle when new vectors that they were not trained on are presented. This lack of creativity stunts their ability to be dynamic when faced with changing circumstances that produce novel fact patterns. Change implies there is no pattern to be detected, and thus the systems fail to do their jobs in this instance.

This pattern-centric approach taken by ADS reduces the value of each individual piece of data down to binary categorizations. ADS can find a pattern, but they cannot determine whether the pattern is a good pattern or a bad pattern and how the pattern can affect real people that are represented by the data points. ADS and systems analogous to it have no way of assigning notions of right and wrong to patterns; these notions have to be taught. Teaching systems to understand concepts of right and wrong is a difficult task when there is no absolute consensus on what differentiates something as right as opposed to being wrong. These notions lay on an evolving spectrum of understanding, hence teaching ADS to patternize this understanding is next to impossible.

A. Black Box Properties

Although we can understand how ADS form patterns, ADS still lack explainability. In many cases, they cannot explain to users how they determined one piece of input fit into one pattern as opposed to another piece of input that did not fit into the same pattern. This lack of explainability is attributed to a property of its design known in the computer science world as the Black Box problem. The Black Box problem touches all types and styles of ADS in some form, taking on different properties in each iteration. The problem describes the void between human understanding of machine algorithms and algorithm functionality.

Yavar Bathaee, a litigator at Bathaee Dunne Limited Liability Partnership (LLP.) and a self-proclaimed AI enthusiast, eloquently defines the Black Box problem "as an inability to fully understand an AI's decision-making process and the inability to predict the AI's decisions or outputs."[71] He divides the problem into two parts, Strong Black Boxes and Weak Black Boxes. "Strong black boxes," as he calls them, "are AI with decision-making processes that are entirely opaque to humans. There is no way to determine how the AI arrived at a decision, what information is outcome determinative to the AI, or to obtain a ranking of the variables processed by the AI in the order of their importance. This form of black box cannot be analyzed ex post by reverse engineering the AI's outputs."[72] Weak Black Boxes, on the

[71] Bathaee, Y. (2018). THE ARTIFICIAL INTELLIGENCE BLACK BOX AND THE FAILURE OF INTENT AND CAUSATION. Harvard Journal of Law & Technology, 31(2). Retrieved from https://jolt.law.harvard.edu/assets/articlePDFs/v31/The-Artificial-Intelligence-Black-Box-and-the-Failure-of-Intent-and-Causation-Yavar-Bathaee.pdf.

[72] Ibid.

other hand, are also opaque to humans but can be reverse engineered or probed to determine a loose ranking of the importance of the variables the AI takes into account.[73] "This in turn may allow a limited and imprecise ability to predict how the model will make its decisions."[74]

Both types of Black Boxes have the capability of functioning outside the creators' initial goals in ways the creators are not able to understand or predict. The lack of transparency generated by ADS' Black Box feature is referred to as a system's complexity. Depending on the problem, an engineer is trying to solve, the complexity of an algorithm can get incredibly dense. Coupled with all the advancements in ADS that are allowing for computational power beyond what was previously comprehensible, the Black Box problem of ADS is only growing. The Black Box problem helps explain why ADS output cannot establish probable cause under the Fourth Amendment because the lack of transparency generated within these systems cannot meet the explainability requirement needed for probable cause. ADS algorithms rely on geometric relationships that humans struggle to visualize; thus, human audits cannot understand the machine's decision-making process and therefore render the output inexplainable.

B. Adversarial Machine Learning

Due to ADS's Black Box nature, figuring out why a system decided to categorize a data point in one way as opposed to another can be difficult. But, coders smart enough to reverse engineer the decision process and understand the machine's decision can manipulate the system's design. Adversarial Machine Learning is a technique used to deceive models. It

[73] Ibid.
[74] Ibid.

works similar to an optical illusion where coders can intentionally design an input that forces the model to make a mistake.[75]

A famous adversarial example is described in a 2015 paper published at the International Conference on Learning Representations entitled, *Explaining and Harnessing Adversarial Examples*.[76] Google researchers Ian Goodfellow, Jonathon Shlens and Christian Szegedy, began with an image of a panda. By adding an imperceptibly small change to the panda image they were able to change the panda image's classification to gibbon with high confidence. To the human eye the image was still clearly a panda, but to the machine, the image was best classified as a gibbon. Because the machine cannot articulate what made it reclassify the panda to a gibbon, nor can it consider why this question matters, the machine makes this judgment presuming it is correct although a human would know otherwise. In this benign example it is easy to miss the significance of this flaw, but when considered in terms of an autonomous car who misidentifies a stop sign and speeds through a busy intersection, the significance of this flaw is magnified.[77]

Adversarial examples illustrate important limitations of ADS, revealing how they can be unknowingly fooled and thus vulnerable to miscalculations. Small imperceptible difference makes ADS unreliable because it is hard for human audits to understand why the system changed its classification when no change is visible to the human eye. There is an incommensurability between the way in which the human brain and ADS interpret the world, which allows for adversarial flaws. Therefore, in a legal sense, machine statements of

[75] Goodfellow, I. (2019, March 7). Attacking Machine Learning with Adversarial Examples. Retrieved from https://openai.com/blog/adversarial-example-research/
[76] Goodfellow, I. J., Shlens, J., & Szegedy, C. (2015). Explaining And Harnessing Adversarial Examples. Google Inc. Retrieved from https://arxiv.org/pdf/1412.6572.pdf
[77] Eykholt, K., Evtimov, I., Fernande, E., Li, B., Rahmati, A., Xiao, C., … Song, D. (2018). Robust Physical-World Attacks on Deep Learning Visual Classification. CVPR.

probable cause can produce inaccurate determinations because their interpretation of the world is vastly different from that of humans.

C. Bad Classifiers

Another limitation of ADS is the possibility for bad classifiers in ML algorithms. Classifiers, in this case, refer to the mathematical function that maps input data to a category. Bad classifiers thus map undesirable correlations that are difficult to detect. In a contrived example produced by a 2016 paper, '*Why Should I Trust You?': Explaining the Predictions of Any Classifier*, Marco Tuli Riberiro, Carlos Guestrin, and Sameer Singh, all researchers with connections to the University of Washington, hand-selected twenty images of wolves and huskies to train a model.[78] The training images of wolves all deliberately had snow in the background while the husky images did not. An additional sixty images were given to the system and the classifier predicted wolf if there was snow or a light background at the bottom and husky otherwise regardless of animal color, position, pose, etc.[79] This bad classifier delineated snow and light backgrounds as the distinguishing factor to determine whether the image was a husky or a wolf.

Any human observer would know that a wolf is not a wolf because of its environment. Rather, a wolf is a wolf because of its physical characteristics, anatomy, and other specific traits unique to the wolf species. A husky similarly is a husky because it has husky physical characteristics, anatomy, and other specific traits unique to the dog species. A human observer would also know that a husky can be found in wintery environments, so making that the distinguishing classifier would be a bad classification.

---

[78] Ribeiro, M. T., Singh, S., & Guestrin, C. (n.d.). "Why Should I Trust You?" Explaining the Predictions of Any Classifier. Retrieved from https://www.kdd.org/kdd2016/papers/files/rfp0573-ribeiroA.pdf.
[79] Ibid, footnote 64.

Although this example was explainable, the room for how nuanced machine thinking can be with bad classifiers is a vulnerable limitation of ADS. Such poor classification metrics speak to ADS's capability to optimize for wrong utility functions that, in complex systems, would be opaque to humans and can result in harm. Since ADS cannot determine whether they are using a bad classifier or not in their computing the system's output would not be able to constitute probable cause.

D. Proxy Variables

Similar to bad classifiers, proxy variables, as a technical limitation, prove to be a substantiable legal restriction on ADS's capability to establish probable cause because they are discriminatory in effect. Proxy variables are variables that have a close correlation to the goal of the program but inherently represent something unrelated and often discriminatory. They arise from confounding variables that can produce anomalous correlations and thus faulty output. Confounding variables can be anything such as race or age that are not directly inputted into the system but are accounted for through proxy variables such as zip code or credit score, making the output legally discriminatory and accordingly impermissible in a court of law.

Algorithms are very good at discovering proxy variables, although they are not the first to use them for discriminatory practices. Take qualified voters in America throughout fluctuating historical time periods for a perfect example of proxy variables discriminating against marginalized populations. The Constitution does not definitively spell out who is eligible to vote. Over time four amendments have been passed prohibiting the disenfranchisement of certain marginalized demographics. The Fifteenth Amendment, ratified in 1870, prohibited the disenfranchisement of citizens "on the account of race, color, or previous condition of servitude."[80] The Nineteenth Amendment, ratified in 1920, prohibited disenfranchisement of citizens "on the account of sex."[81] The Twenty-Fourth Amendment, ratified in 1964, abolished the poll tax qualification for Federal Elections.[82] Finally, the Twenty-Sixth Amendment, ratified in 1971, prohibited the disenfranchisement of citizens on the account of age, setting the minimum age to 18 years old.[83] Aside from these federal prohibitions the rest was largely left up to the States. The States leveraged proxy variables such as property ownership, religious tests, free status, poll taxes, literacy tests, and recently incarceration as means to keep marginalized populations from voting and maintain disenfranchisement in their elections. Although none of these qualifications directly related to discriminatory categories, because of institutional inequalities and systematic racism these qualifications correlated exceedingly well with race, ethnicity, age, and gender, among other protected categories. Thus, the implications of these proxy variables had discriminatory effects.

Outside of voting the practice of redlining encompassed a number of proxy variables that correlated exceedingly well with minorities and had discriminatory implications on their communities. Redlining is the process of systematically denying various public and private services to residents of specific neighborhoods through raising prices. Services such as banking, insurance, mortgages, health care, or even retail businesses are denied to residents because their neighborhoods are deemed 'riskier' than similar neighborhoods of different racial compositions, and thus prices in their neighborhood for service providers are higher. Bill Dedman, an investigative journalist for Newsday, won the Pulitzer Prize for his series of

---

[80] United States Constitution Amendment XV
[81] United States Constitution Amendment XIX
[82] United States Constitution Amendment XXIV
[83] United States Constitution Amendment XXVI

articles, *The Color of Money*, where he divulges how banks and other mortgage lenders in Atlanta were discriminating against black neighborhoods. Despite giving loans to the poorest white neighborhoods in Atlanta, Dedman discovered that many banks and other mortgage lenders did not lend in middle-class or more affluent black neighborhoods. He explained how the banks made use of proxy variables such as:

> … poor quality housing and lack of home sales in black neighborhoods, fewer applications from blacks, and limitations in the federal lending data… real estate agents, appraisers, federal loan programs… banking officials said they would make more loans to blacks if real estate agents sent them more black applicants. Real estate brokers who work in black neighborhoods confirmed that they often don't send black homebuyers to banks or savings and loans, but said that is because those institutions have not been responsive and do not solicit their business.[84]

Dedman's discovery of proxy variables leveraged by many banks and other mortgage lenders in Atlanta revealed a cyclical feedback loop that continuously excluded black neighborhoods from further development and thus perpetuated a toxic brand of racism. Similar to exclusionary voting qualifications, although none of the bank's lending factors directly related to race, institutional inequalities and systematic racism correlated these factors exceedingly well with attributes of being non-white. Thus, the implications of these proxy variables had discriminatory effects.

Both exclusionary voting qualifications and discriminatory redlining reveal how proxy variables have been embedded in public policy long before computer algorithmic polices started shining interrogative spotlights on them. ADS had nothing to do with proxy variables

[84] Dedman, Bill. "Atlanta Blacks Losing in Home Loans Scramble: Banks Favor White Areas by 5-1 Margin." The Atlanta Journal-Constitution, 1 May 1988. The Color of Money, http://powerreporting.com/color/1a.html.

used in the Jim Crow South or in the banks of redlined Atlanta, yet they were still there. This phenomenon is a call to look introspectively and retrospectively at our society to understand that bias is not something ADS created. However, it is something ADS encodes. Although societal and political norms concerning fairness have progressed over the natural course of human history, the evolution of these norms should not be conflated with their erasure.

Bias is endemic to every known community. People make decisions on what is or is not possibly relevant for an algorithm to be trained on, which means that fallible, biased people are inserting their own values into the "objective" algorithms. "Bias in, bias out," an adapted version of the computer-science idiom "garbage in, garbage out," coined by Sandra G. Mayson, an Assistant Professor of Law at the University of Georgia School of Law, encapsulates this notion.[85] Algorithms' predictions are only as good as the data on which they are trained, so if they are trained on garbage they will produce garbage, and if they are trained on biased input they will produce biased output.

ProPublica, an independent non-profit newsroom, revealed the implications of algorithms trained on biased data producing biased output in its in-depth review of how proxy variables can manifest in risk assessment scores. The risk assessment scores investigated in ProPublica's report concern scores that inform decisions about who can be set free at every stage of the criminal justice system, from assigning bond amounts to priming judges during criminal sentencing.[86] ProPublica examined risk assessment scores assigned to over 7,000 people arrested in Broward County, Florida, as their case study. These scores were calculated by a product created by a for-profit company called Northpointe. ProPublica's

---

[85] Mayson, Sandra Gabriel, Bias In, Bias Out (September 28, 2018). 128 Yale Law Journal 2218 (2019),; University of Georgia School of Law Legal Studies Research Paper No. 2018-35. Available at SSRN: https://ssrn.com/abstract=3257004.
[86] Ibid, footnote 3.

report detailed how machine bias was disproportionately impacting Black America, explaining that it is "difficult to construct a score that doesn't include items that can be correlated with race such as poverty, joblessness, and social marginalization."[87] Ruha Benjamin, an associate professor at Princeton University, comments on this problem of machine bias disproportionately affecting Black America in her book, *Race After Technology*, where she explains how ML is trained on data manufactured through histories of exclusion and discrimination.[88] In her book, she coins the phrase "New Jim Code," which illuminates how biased data reinforces notions of White Supremacy and deepens social inequity. She implores her readers to consider the decisive question, do algorithms reduce existing inequities or make them worse?[89] Similar to bad classifiers, the room for how nuanced machine thinking can be when such biased data is authorized to produce such random proxy variables is a vulnerable limitation of ADS.

As mentioned previously, algorithms are very good at discovering proxy variables, such as the ones described by ProPublica in their report, but they have no way of accounting for them to produce fair and just outcomes. Proxy variables act as confounding factors that subject algorithms to anomalous and even spurious output. This output can result in discrimination, thus encoding bias in a mathematical sense. ADS have no contextual conception of what makes a proxy variable a proxy variable and thus have no corrective measure to address this kind of discrimination. The intent of the algorithm is to find correlations that facilitate arriving at their goal, not to understand the cause of said correlations and how they may produce discriminatory effects. Thus, this potential for

---

[87] Ibid, footnote 3.
[88] Benjamin, Ruha. Race after Technology: Abolitionist Tools for the New Jim Code. Polity, 2019.
[89] Ibid.

discriminatory output is socially and legally unacceptable, inhibiting ADS from establishing probable cause that could be permissible in a court of law.

D. Human v. Machine Decision Making Process

All the previously mentioned limitations of ADS contain some aspect of how machines arrive at decisions in a manner that is radically different from that which humans employ. This divide in the decision-making process is vital when considering the capacity for machines' decision output to be viable in court. To better understand this difference, consider how humans arrive at decisions and contrast that with how machines arrive at decisions in order to formulate the differentiating factor(s) that omit machine decision output from legal validity.

The human capacity of "judgment" describes the various steps individuals use when trying to reach beyond evidence encountered to draw conclusions from that evidence.[90] Thus by its nature, human judgment requires a level of extrapolation. To deal with this extrapolation, humans employ different thinking strategies dependent on the format of the data set and the type of evidence being considered.[91] Research in psychology, the scientific field of study that explores the human mind and its functions, suggests that humans make judgments by relying on a small set of shortcuts called judgment heuristics.[92] The two major heuristics we deploy are the availability heuristic and the representativeness heuristic.

The availability heuristic is the strategy of judgment that uses how easily an example comes to mind as the basis for assessing how common that example is in the world.[93] When an

[90] Gleitman, H., Gross, J. J., & Reisberg, D. (2011). Psychology (8th ed.). page 348 ISBN 978-0-393-11726-4.
[91] Ibid, at page 353.
[92] Ibid, at page 348.
[93] Ibid, at page 349.

individual is faced with a decision, an individual wants their conclusion to rely on not just one observation but on a pattern drawn from various observations that summarize multiple experiences. This summary generally requires a comparison among frequency estimates or an assessment of how often an individual has encountered a particular example. These frequency estimates are central for human judgment, but often the human mind has difficulty recalling an objective record of experience. Humans think of specific cases relevant to the particular judgment at hand, and if the example comes to mind easily, individuals can conclude that the circumstance is a common one. On the other hand, if it takes an individual a longer period of time to arrive at the judgment outcome, an individual can conclude that the circumstance is nuanced. For many judgment calls, the availability heuristic strategy works because events that are frequent in the world are likely to be frequent in our personal experience and are therefore well represented in our memories.[94] However, there are circumstances in which this strategy is misleading because the organization of memory creates a bias in what is easily available, leading to an error in frequency, a distorted perception, and inadequate precautions exercised in the present case. [95]

The representativeness heuristic relies on broader knowledge to make some forecast about the decision at hand. This heuristic of human judgment hinges on the categorization of examples assuming that each member of a category is representative of the category and each category is relatively homogenous so that every member resembles every other member.[96] The general uniformity in categories allows us to extrapolate from our experiences what to expect next time and thus allows us to make judgments off this expectation.[97] However,

[94] Ibid, at page 348.
[95] Ibid, at page 350.
[96] Ibid, at page 348.
[97] Ibid, at page 351.

overuse of this heuristic can be dangerous because making one member of a category representative of an entire group ignores atypical cases and can lead to erroneous conclusions.

Although these heuristics provide critical insights for the human decision making process, individuals often rise above these heuristic shortcuts and rely on other more laborious but often more accurate judgment strategies.[98] Humans can think within a dual-process theory of judgment meaning that they can utilize two different types of thinking, one that is fast and efficient in a wide range of circumstances, the other that is slower and takes more effort but is less risky and often avoids errors.[99] Similar to ADS, humans can form patterns from presented situations through the aforementioned heuristics that act as a policy for the human brain to instantiate.

However, what differentiates the human decision-making process from that of ADS is humans' ability to recognize patterns outside of what they were taught through utilizing their dual-process theory of judgment. If input from a presented situation correlates with experiences the human mind has seen in the past, the mind can deploy the fast and efficient type of thinking because it is a familiar case, similar to that of ADS. If input from a presented situation is outside what the human mind has previously experienced, humans can take a step back and evaluate the situation, taking the necessary time to come to the best decision possible with the information with which they have been presented combined with the information they already know, vastly different from that of ADS. This reflective action of recognizing what an individual does not know and allocating the needed time and care to figure out the best course of action is a differentiating factor between ADS and human

[98] Ibid, at page 352.
[99] Ibid, at page 352

decision-making. Humans reflectivity permits refinement and adjustment of their decision-making models for missing information and allows them to change their models on the spot before they make a decision that affects others. This reflective step allows humans to produce an output that is closer to being the right decision where ADS would be blatantly wrong.

This unique trait of the human decision-making process arises from humans' capability to consider inefficient suboptimal considerations that factor empathy, kindness, and ethics into the equation when necessary. Humans understand that not all questions have a corresponding right answer, or a definitively correct decision output associated with the particular problem. Many times, decisions are convoluted, and the right decision for one person may not be a suitable decision for another person. Considerations that factor in curiosity, exploration, curation, love, or an experience are all inherently inefficient variables that machines struggle to find the value in, yet humans know to be invaluable resources. The human brain also exercises imagination, allowing it to imagine possibilities beyond a training set. These possibilities may seem illogical or improbable but can be conceivable within a given context that if the circumstances were right, they would be able to grapple with. This gives humans a unique ability to deal with fringe cases with which ADS immensely struggle.

As a direct implication of the divide between ADS decision making and that of humans, ADS would not be able to establish probable cause under the Fourth Amendment because human judgment in judicial discretion is an essential part of probable cause determinations, and current technology cannot adequately substitute for wisdom produced by the human mind. Kiel Brennan-Marquez, an associate professor of Law and William T. Golden Scholar at the University of Connecticut School of Law, affirms this view in his paper, *Plausible Cause: Explanatory Standards in the Age of Powerful Machines*, where he argues that statistical accuracy of ADS is not enough of an explanation to substitute for judicial scrutiny.[100] He explains the concept of value-pluralism, which assesses "which values are at stake in a given decisional environment and ask[s], where necessary, if those values have been properly balanced."[101] Our judicial system's practice of navigating complex lines of value-pluralism:

> enables judges (1) to consider the plurality of values implicated by the exercise of state power and (2) to resolve conflicts between those values in a context-sensitive way. Atday's end, the rationale for individualized review, costly and inefficient as it may be, is that in some settings we cannot be sure in advance which values will be implicated by the exercise of power. And when that is true, decision making resists automation. Decisions must be subject—or at least susceptible—to case-by-case evaluation in order to ensure that no particular value or set of values subsumes others.[102]

Human dual-process theory of judgment allows us to grapple with value-pluralism in ways that ADS cannot because of their unimaginative and rigid policy instantiations that are a predetermined and essential part of their structural framework. Despite trans-humanist views of modern advances in technological approaches to solving this problem of predetermination, ADS still fundamentally lack a capacity to make prudent judgments, which is a crucial aspect of a judge's decision. Therefore, their output cannot establish probable cause under the Fourth Amendment.

F. Thought Experiments

To further illuminate the limitations of current technological approaches let us examine two distinct ML techniques within a legal context. Consider the illustrative example

---

[100] Brennan-Marquez, Kiel, Plausible Cause: Explanatory Standards in the Age of Powerful Machines (September 5, 2016). Vanderbilt Law Review, Vol. 70 (2017). Available at SSRN: https://ssrn.com/abstract=2827733 or http://dx.doi.org/10.2139/ssrn.2827733
[101] Ibid.
[102] Ibid.

of an imaginary judge who is given an affidavit. The judge feeds the affidavit into a supervised learning algorithm which has been trained on affidavits labeled by humans and learned what characteristics correspond with particular affidavit labels. The algorithm checks if the inputted affidavit matches any previously labeled affidavits that were a part of the program's training data. The algorithm then can find commonalities across similar classes of crimes through affidavits' stylized language patterns. For example, drug case's affidavits typically include the phrases "trafficking", "conspire", "possess with intent to distribute", "in violation of 21 U.S.C." among other phrases that can help the algorithm cluster them together.[103] After finding the similarities, the supervised system produces an answer that fits the label it was trained on and accompanies that answer with a confidence level. The algorithm was trained on a corpus of legal data including other affidavits that at least resemble the new affidavit being input into the algorithm, but the algorithm was not trained on the nuances of this case. It may have been trained on previous cases that involved home, car, or office searches but its training set did not include computer searches. How meaningful would the affidavits' similarities be in this case?

Consider the difference between the physical search of a home, car, or office compared to that of a digital computer search. A physical search and a digital search are approached in different ways, require different specialists, and discover different pieces of evidence. Although the varying evidence may lead to the same conclusion, the narrative each piece of evidence tells is important for informing the judges outcome. If drugs were found in a home, the owner of the drugs could be anyone living in the house or any visitors to the house, with

[103] Brandon G. Johnson US District Court Eastern District of Tennesee Greenneville Dviison Affidavit for search warrant, Wendy G. Boles U.S. Distirct Court Eastern District of Tennesse at Knoxville Affidavit, Jacquelyn Gilliam U.S. Distirct Court for the Eastern Distrcit of Michigan Affidavit, Derek Dunn U.S. District Court for the Dirtict of New Hampshier Affidavit in support of applications for search warrant

varying degrees of likelihood. If evidence of drug sales were found on a computer, the sale could be linked to an individual username which could have been hacked or used by someone other than the individual linked to the account, but this likelihood is vastly different from the likelihood associated with the drugs found in the house. It is evident to any reasonable person that the computer search should not match any other type of search because no other type of search is similar enough to be clustered together with a computer search. However, an algorithm may disagree and label the computer search as a match with another type of search because it does not consider the semantics. How useful would the supervised learning output be here to deal with the nuance of a computer search?

In a similar manner, the imaginary judge feeds the same affidavit into an unsupervised learning algorithm which evaluates the affidavit by different metrics. Dissimilar to the supervised learning algorithm that compares the given affidavit with affidavits on which it was trained, the unsupervised learning algorithm looks at the given affidavit and clusters it with other affidavits it finds to be similarly matched based on its own metrics. After it matches a cluster of affidavits, humans have the option to go in and label the clusters. Based on the cluster in which the given affidavit falls, the judge will be able to make a historically consistent decision by following the precedent set by the given affidavits' cluster.

The algorithm could run for years, producing reasonable outcomes and keeping the judge aligned with precedent. However, some day a defendant could challenge the algorithm's metrics necessitating an audit of the system. The algorithm could have been clustering affidavits based on key terms such as 'loitering,' 'traffic violation,' or 'uncooperative.' These terms may suggest and, in many cases, may be linked to criminal activity, but the algorithm has no way to check if these clusters are structured on proxy variables that often marginalize minority communities. With this understanding, how would the imaginary judge who feeds the aforementioned affidavit into an unsupervised learning algorithm be able to audit the system to ensure the algorithm clusters the affidavits by a permissible legal standard and not by proxy variables?

A system where this process was in part digitized could prove these thought experiments viable. With society's growing lean towards digitalization, a future where this possibility is a reality raises interesting questions about whether ADS's output can constitute probable cause to issue a search warrant. In the thought experiment produced above, neither supervised nor unsupervised learning algorithms declared definitively if there is or is not probable cause. They merely assisted the judge's determination by evaluating aggregated historical data. But the validity of this assistance is problematic when considering the limitations of both technical approaches in conjunction with the current operations of modern-day courtrooms in America.

Federal courts very rarely reject affidavits for search warrants; instead, they instruct officers how to rework the initial affidavit so it could be accepted on the next submission. This creates a shortage of rejected affidavits that can no longer teach the algorithm the standard for permissible and impermissible evidence. Fortunately, the courts have a good faith exception, which permits evidence if the law was not clear on a particular point, thus providing the algorithm with affidavits that were not sufficient, but evidence that was permissible. This exception provides a better source of rejected affidavits for the algorithm to determine the standard for permissible and impermissible evidence. On account of these two realities of legal proceedings, there are very few cases where evidence is suppressed because of insufficient affidavits; hence there would not be enough training data for the algorithm to run efficiently and satisfactorily. Digitizing this process with current technology would undermine our constitutional commitment to provide individualized and equal justice for all.

5. CONCLUSION

In *Kyllo v. United States, 533 U.S. 27 (2001)*, Danny Lee Kyllo was indicted for growing marijuana in his house through police use of a thermal-imaging device from the street. The case became immediately controversial as the thermal-imaging device was seen as a direct invasion of Kyllo's privacy and infringement on his Fourth Amendment rights. The Supreme Court agreed the warrantless use of a thermal-imaging device aimed at a private home from a public street constituted an unlawful search within the meaning of the Fourth Amendment and reversed the decisions of both the Oregon District Court and the Court of Appeals for the Ninth Circuit to deny Kyllo's motion to suppress the seized evidence.

Justice Scalia delivered the opinion of a divided court offering crucial insight into how to deal with technology in the future constitutionally. He addressed the elephant in the room, asserting, "It would be foolish to contend that the degree of privacy secured to citizens by the Fourth Amendment has been entirely unaffected by the advance of technology."[104] Justice Scalia then fixated his opinion on the pertinent question of "what limits there are upon this power of technology to shrink the realm of guaranteed privacy."[105] He came upon the broad idea of "general public use" that because the thermal-imaging device was not in "general public use," it constituted an intrusion into a constitutionally protected realm of citizen life. Justice Stevens acutely points out in his dissent that in Justice Scalia's "general public use" explanation, he created an "all-

---

[104] Kyllo v. United States, 533 U.S. 27, 121 S. Ct. 2038, 150 L. Ed. 2d 94, 2001 U.S. LEXIS 4487, 69 U.S.L.W. 4431, 2001 Cal. Daily Op. Service 4749, 2001 Daily Journal DAR 5879, 2001 Colo. J. C.A.R. 2926, 14 Fla. L. Weekly Fed. S 329 (Supreme Court of the United States June 11, 2001, Decided).
[105] Ibid.

encompassing rule for the future" that shackles the Court to a "prematurely devised constitutional constraint."[106] This dissent begs the question of what happens when thermal-imaging devices or any technology for that matter become "general public[ly] use[d]" products. Justice Scalia's timely question concerning the limits of technology on space for individual privacy is complicated by the growing ubiquitous nature of technological devices that are inherently privacy invasive.

There is something dangerous about buying into the idea that technology's pervasiveness in generally used public products can constitutionally evolve at the expense of citizen's right to privacy. I would oppose Justice Scalia's "general public use" rule, bearing in mind how the average citizen interacts with advanced complicated technological systems. Citizen's encounter ADS every day through their habitual use of Google, Netflix, Facebook, and Amazon, among other platforms. Yet many citizens do not entirely understand how Google's search engine works, how Netflix recommends movies, how Facebook populates its timeline, how Amazon suggests items, and what the implications of using these platforms are for their privacy. Simply because many ADS may appear familiar to individuals does not make them reasonable tools for invasions of traditionally shielded areas of private life. Still, Justice Scalia's "general public use" rule, would determine ADS do not constitute an intrusion into a constitutionally protected realm of citizen life because ADS are universally employed throughout society. How will the courts rectify this tension moving forward?

To start, if ADS become explainable, then the limit bounds of their use in society may change. Explainable AI, as a research discipline for computer scientists and academics, is a

---

[106] Ibid.

growing field as the need for comprehensive automated answers increases. The inaugural Association for Computing Machinery (ACM) Conference on Fairness, Accountability, and Transparency took place in 2018 and has taken place in each successive year since convening to discuss issues including explainable AI. The conference draws academics from all around the world to explore multi-disciplinary approaches to computing machinery ethics. Resulting conversations focused on explainable AI mention potential solutions to the interpretability problem faced when human experts seek to understand AI output. However, none of the proposed solutions are entirely explainable, despite revealing a promising start.

The question of how much individuals can trust explainable AI outputs would still remain, as explainability does not correct all the other vulnerabilities inherent within ADS such as bias and discrimination. Societal norms change with every new generation born, but ADS encode generational bias without a proper mechanism to adjust to changing generational norms that work as an anecdote to bias. When new generations realize the discrimination faced by their ancestors, they find a corrective apparatus to mitigate such bias. Explainable ADS would need to work as such a corrective apparatus continuously retraining itself on new data that accurately reflects changes in culture. New data may look different to the system since it was stripped of old biases, but explainable ADS must be capable of adapting to such flux while maintaining satisfactory functionality. This ability to adapt to changing norms must extend to handle new Supreme Court interpretations that update the standard of acceptability in society. Until fully operational explainable AI is common practice in ADS, the above analysis and this conclusion remain unchanged as they are both rooted in the technology of today.

As made evident above, bias is one aspect of ADS's lack of explainability that disqualifies its output from establishing probable cause under the Fourth Amendment, as anything discriminatory in nature is impermissible in criminal court. Bias is not something ADS created. Rather, it is an endemic feature of communities that ADS encode in their algorithms and perpetuate in their output. Congress needs to pass laws that align with the Supreme Court's stance on bias starting with the Algorithmic Accountability Act of 2019 that was introduced to the 116th Congress in April 2019 and referred to the Subcommittee on Consumer Protection and Commerce.[107] This bill seeks to regulate bias in automated decision-making systems by requiring companies to audit their ML algorithms for discrimination in an impact assessment.[108] The impact assessment will provide necessary transparency about how companies approach the question of fairness in their models and thus will bring accountability back into the fold.

The Algorithmic Accountability Act and other bills similar to it will help alleviate the strain massive innovation placed on the legal frameworks tasked with protecting citizens from the unintended consequences of the Fourth Industrial Revolution. Progressions towards the near future where ADS are used in more contexts connected to criminal justice are not out of the question, and Congressional support through bills will help address the immediate risks posed by such a future. There is a strong constitutional case to be made, as demonstrated *supra*, that current ADS do not meet the rigid legal explainability standard required to establish probable cause under the Fourth Amendment because they cannot provide enough reliable evidence to justify their output. Therefore, under present legal interpretations and available technology, probable cause determinations established from correlations found by ADS are impermissible in United States Courts.

---

[107] "All Info - H.R.2231 - 116th Congress (2019-2020): Algorithmic Accountability Act of 2019." Congress.gov, 11 Apr. 2019, www.congress.gov/bill/116th-congress/house-bill/2231/all-info.

[108] Hao, Karen. "Congress Wants to Protect You from Biased Algorithms, Deepfakes, and Other Bad AI." MIT Technology Review, MIT Technology Review, 2 Apr. 2020, www.technologyreview.com/2019/04/15/1136/congress-wants-to-protect-you-from-biased-algorithms-deepfakes-and-other-bad-ai/.

6. BIBLIOGRAPHY

"*All Info - H.R.2231 - 116th Congress (2019-2020): Algorithmic Accountability Act of 2019*." Congress.gov, 11 April 2019, www.congress.gov/bill/116th-congress/house-bill/2231/all-info.

Aguilar v. Texas, 378 U.S. 108, 84 S. Ct. 1509, 12 L. Ed. 2d 723, 1964 U.S. LEXIS 994 (Supreme Court of the United States June 15, 1964, Decided).

Aleksander, Igor. "*Decision and Intelligence*." Volume 6. London: Kogan. 2013. Page. p. 185. ISBN 9781850914075.

Alkhatib, Ali. "*Street-Level Algorithms: A Theory at the Gaps Between Policy and Decisions*." Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, edited by Michael Bernstein, ACM, 2019, pp. 530:1–530:13. CHI '19.

ALPAC Report. "*Language and Machines — Computers in Translation and Linguistics*." A Report by the Automatic Language Processing Advisory Committee, Washington, DC, 1966.

Angwin, J., Larson, J., Kirchner, L., & Mattu, S. (2019, March 9). "*Machine Bias*." Retrieved from https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

Barron v. Baltimore., 32 U.S. 243, 8 L. Ed. 672, 1833 U.S. LEXIS 346 (Supreme Court of the United States February 16, 1833, Decided).

Bathaee, Yavar. "*The Artificial Intelligence Black Box And The Failure Of Intent And Causation*." Harvard Journal of Law & Technology, vol. 31m ser. 2, Spring 2018, pp. 890-934.

Beavers, Anthony. "*Alan Turing: Mathematical Mechanist*". In Cooper, S. Barry; van Leeuwen, Jan (eds.). Alan Turing: His Work and Impact. Waltham: Elsevier. 2013. pp. 481–485. ISBN 978-0-12-386980-7.

Bellovin, Steven M., Matt Blaze, Susan Landau, and Brian Owsley. "*Seeking the source: Criminal defendants' constitutional right to source code.*" Ohio State Technology Law Journal, vol. 17 ser. 1, December 2020. To appear.

Bellovin, Steven M., Preetam K. Dutta, and Nathan Reitinger. "*Privacy and synthetic datasets*." Stanford Technology Law Review, vol. 22 ser. 1, 2019.

Benjamin, Ruha. "*Race after Technology: Abolitionist Tools for the New Jim Code*." Polity, 2019.

Brennan-Marquez, Kiel. "*Plausible Cause: Explanatory Standards in the Age of Powerful Machines*" Vanderbilt Law Review, vol. 70, 2017.

Brinegar v. United States, 338 U.S. 160, 69 S. Ct. 1302, 93 L. Ed. 1879, 1949 U.S. LEXIS 2084 (Supreme Court of the United States June 27, 1949, Decided).

Brownlee, Jason. "*A Gentle Introduction to Matrix Operations for Machine Learning*." Machine Learning Mastery. 9 August 2019. https://machinelearningmastery.com/matrix-operations-for-machine-learning/.

Castelluccia, Claude, and Le Métayer, Daniel. "*Understanding algorithmic decision-making: Opportunities and challenges.*" Scientific Foresight Unit. Directorate-General for Parliamentary Research Services. European Parliament. 2019. https://www.europarl.europa.eu/RegData/etudes/STUD/2019/624261/EPRS_STU(2019)624261_EN.pdf.

"*CIA's 'Facebook' Program Dramatically Cut Agency's Costs*." The Onion. 18 October 2017. www.theonion.com/cias-facebook-program-dramatically-cut-agencys-costs-1819594988.

Dedman, Bill. "*Atlanta Blacks Losing in Home Loans Scramble: Banks Favor White Areas by 5-1 Margin.*" The Atlanta Journal-Constitution. 1 May 1988. The Color of Money. http://powerreporting.com/color/1a.html.

DRAPER v. UNITED STATES, 358 U.S. 307, 79 S. Ct. 329, 3 L. Ed. 2d 327, 1959 U.S. LEXIS 1607 (Supreme Court of the United States January 26, 1959, Decided).

Entick v. Carrington (1765) 19 Howell's State Trials 1029; 95 ER 807

Eykholt, Kevin. Evtimov, Ivan… Song, Dawn. "*Robust Physical-World Attacks on Deep Learning Visual Classification*." arXiv. Cornell University. 2018.

Felton, Ryan. "Michigan unemployment agency made 20,000 false fraud accusations – report." The Guardian. 18 December 2016. https://www.theguardian.com/us-news/2016/dec/18/michigan-unemployment-agency-fraud-accusations.

Florida v. Harris, 568 U.S. 237, 133 S. Ct. 1050, 185 L. Ed. 2d 61, 2013 U.S. LEXIS 1121, 81 U.S.L.W. 4081, 24 Fla. L. Weekly Fed. S 18 (Supreme Court of the United States February 19, 2013, Decided).

Gleitman, H., Gross, J. J., & Reisberg, D. "*Psychology*." 8th edition. 2011. ISBN 978-0-393-11726-4.

Gray, David C. and Citron, Danielle Keats. "*The Right to Quantitative Privacy.*" Minnesota Law Review. vol. 98, 2013. U of Maryland Legal Studies Research Paper, 2013-23.

Goodfellow, Ian. Shlens, Jonathon, and Szegedy, Christian. "*Explaining and Harnessing Adversarial Examples.*" International Conference on Learning Representations. Google Research. 2015

Goodfellow, Ian. Papernot, Nicolas… Clark, Jack. "*Attacking Machine Learning with Adversarial Examples*." OpenAI. 24 February 2017.

Hao, Karen. "*Can You Make AI Fairer than a Judge? Play Our Courtroom Algorithm Game*." MIT Technology Review. 2 April 2020.

Hao, Karen. "*Congress Wants to Protect You from Biased Algorithms, Deepfakes, and Other Bad AI*." MIT Technology Review. 2 April 2020.

Hardesty, Larry. "*Explained: Neural Networks*." MIT News. MIT News Office. 14 April 2017.

Illinois v. Gates, 462 U.S. 213, 103 S. Ct. 2317, 76 L. Ed. 2d 527, 1983 U.S. LEXIS 54, 51 U.S.L.W. 4709 (Supreme Court of the United States June 8, 1983, Decided).

Joh, Elizabeth E., "*Policing by Numbers: Big Data and the Fourth Amendmen*t." Washington Law Review. vol. 89 rev. 35, 2014, pp. 35-68.

Kerr, Orin S. "*Impressions from the Oral Argument in Kansas v. Glover*." Reason, 5 November 2019.

Kerr, Orin S. "*Reasonable Suspicion From Driver to Car: A Few Thoughts on Kansas v. Glover*." Reason. 31 October 2019.

Kerr, Orin S. "*Why Courts Should Not Quantify Probable Cause*" The Political Heart of Criminal Procedure: Essays on Themes of William J. Stuntz (Michael Klarman, David Skeel, and Carol Steiker, eds), pages 131-43 Cambridge 2012

Khan, M., Jan, B., & Farman, H. "*Deep Learning: Convergence to Big Data Analytics*." Springer Singapore. 2019. doi: 10.1007/978-981-13-3459-7

Knight, Will. "*Military Artificial Intelligence Can Be Easily and Dangerously Fooled*." MIT Technology Review. 2 April 2020.

Kim, J. (Ed.). "Fourth Amendment." Law. Cornell University. 2017. https://www.law.cornell.edu/wex/fourth_amendment

Klein, Daniel. "Mighty mouse." MIT Technology Review. 19 December 2018.

Kowalski, Robert A. "*The early years of logic programming*" The ACM Digital Library. Association for Computing Machinery. January 1988.

Kyllo v. United States, 533 U.S. 27, 121 S. Ct. 2038, 150 L. Ed. 2d 94, 2001 U.S. LEXIS 4487, 69 U.S.L.W. 4431, 2001 Cal. Daily Op. Service 4749, 2001 Daily Journal DAR 5879, 2001 Colo. J. C.A.R. 2926, 14 Fla. L. Weekly Fed. S 329 (Supreme Court of the United States June 11, 2001, Decided).

Lacambra, Stephanie. "*Predictive Policing one pager*." Electronic Frontier Foundation. 30 October 2017.

Lee, Timothy B. "*How neural networks work-and why they've become a big business.*" arstechnica. WIRED Media Group. Condé Nast. 2 December 2019.

Loewy, Arnold H. "*Rethinking Search and Seizure in a Post-9/11 World*" Mississippi Law Journal. vol. 80, no. 4, Summer 2011.

Lighthill, James. "*Artificial Intelligence: A General Survey*." Artificial Intelligence: a paper symposium. Science Research Council. 1973.

Maini, Vishal. "*Machine Learning for Humans, Part 2.1: Supervised Learning*." Machine Learning for Humans. Medium. 19 August 2017.

Maini, Vishal. "*Machine Learning for Humans, Part 3: Unsupervised Learning*." Machine Learning for Humans. Medium. 19 August 2017.

Mayson, Sandra G. "*Bias In, Bias Out*" Yale Law Journal. vol. 128, 2218. 2019. University of Georgia School of Law Legal Studies Research Paper No. 2018-35.

McCarthy, Andrew C. "*Roger Stone's sentence: How the judge will really decide to 'lock him up'*." The Hill. Capitol Hill Publishing Corp. News Communications INC. 14 February 2020.

McCarthy, J., Minsky, M., Rochester, N., Shannon, C.E., "*A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*." 31 August 1955. http://raysolomonoff.com/dartmouth/boxa/dart564props.pdf.

McCauliff, C.M.A. "*Burdens of Proof: Degrees of Belief, Quanta of Evidence, or Constitutional Guarantees?*" Vanderbilt Law Review vol. 35: 1293. 1982.

Mnih, Volodymyr. Silver, David… Graves, Alex. "*Playing Atari with Deep Reinforcement Learning.*" DeepMind Technologies. arXiv. Cornell University.

Obermeyer, Ziad. Powers, Brian… Vogeli, Christine. "*Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations.*" Science, vol. 366, no. 6464, pp. 447-453. 25 October 2019.

Ornelas v. United States, 517 U.S. 690, 116 S. Ct. 1657, 134 L. Ed. 2d 911, 1996 U.S. LEXIS 3391, 64 U.S.L.W. 4373, 96 Cal. Daily Op. Service 3744, 96 Daily Journal DAR 6059, 9 Fla. L. Weekly Fed. S 617 (Supreme Court of the United States May 28, 1996, Decided).

Pan, Sheri B. "*Get To Know Me: Protecting Privacy And Autonomy Under Big Data's Penetrating Gaze*." Harvard Journal of Law & Technology, vol. 30, ser. 1, Fall 2016, pp. 239–261.

Pasquale, Frank A., "*Professional Judgment in an Era of Artificial Intelligence and Machine Learning*" University of Maryland Digital Commons. Faculty Scholarship. Publication 1614. 2019.

Poole, David. Mackworth, Alan. Goebel, Randy. "*Computational Intelligence: A Logical Approach.*" Oxford University Press. New York. ISBN 978-0-19-510270-3. 1998

PSA. "*Risk Factors and Formula*." Public Safety Assessment (PSA). Laura and John Arnold Foundation. www.psapretrial.org/about/factors.

Puente, Mark. “*Police Leaders Debate Merits of Using Data to Predict Crime.”* Government Technology, e.Republic. 5 July 2019.

Ribeiro, Marco T. Singh, Singh. and Guestrin, Carlos. “’*Why Should I Trust You?’: Explaining the Predictions of Any Classifier.*” arXiv. Cornell University. 2016.

SAS. “*Machine Learning: What it is and why it matters*.” SAS Institute. https://www.sas.com/en_us/insights/analytics/machine-learning.html.

Schwab, Klaus. The Fourth Industrial Revolution. Portfolio Penguin, 2017.

Semayne’s Case All ER Rep 62 5 Co Rep 91 a Cro Eliz 908 Moore KB 668 Yelv 29 77 ER 194

Simmons, Ric. “*Quantifying Criminal Procedure: How to Unlock the Potential of Big Data in Our Criminal Justice System*” Michigan State Law Review vol. 947. pp. 947-1017. 2016. Ohio State Public Law Working Paper no. 362.

Slobogin, Christopher. “*Let's Not Bury Terry: A Call for Rejuvenation of the Proportionality Principle*.” St John's Law Review. vol. 72:1053, pp. 1053-1096. 1998. Vanderbilt Law School Faculty Publications.

Spinelli v. United States, 393 U.S. 410, 89 S. Ct. 584, 21 L. Ed. 2d 637, 1969 U.S. LEXIS 2701 (Supreme Court of the United States January 27, 1969, Decided).

Tetlock, Philip E. Gardner, Dan. “*Superforecasting: The Art and Science of Prediction*.” Philip Tetlock Consulting Inc. Connaught Street, Inc. Broadway Books. Crown Publishing Group. Penguin Random House L.L.C. New York. 2015.

Thompson, Derek. “*Should We Be Afraid of AI in the Criminal-Justice System*?” The Atlantic, Atlantic Media Company, 20 June 2019.

Turing, Alan. "*Computing Machinery and Intelligence.”* Mind. vol. LIX issue 236. pp. 433–460. 1 October 1950. Mind Association. Oxford Academic. Oxford University Press.

United States v. Burr, 25 F. Cas. 2, 1807 U.S. App. LEXIS 489 (Circuit Court, D. Virginia April 1, 1807).

United States Constitution

Ybarra v. Ill., 444 U.S. 85, 100 S. Ct. 338, 62 L. Ed. 2d 238, 1979 U.S. LEXIS 151 (Supreme Court of the United States November 28, 1979, Decided).

Zimmeck, Sebastian. Bellovin, Steven M. “*Privee: An architecture for automatically analyzing web privacy policies.*” In 23rd USENIX Security Symposium (USENIX Security 14), pages 1--16, San Diego, CA, August 2014. USENIX Association.