The Art of Accuracy in Artificial Intelligence

Sheila Niaz

Imagine building a company that everyone believed would change the world. While on a trip to Asia, Elizabeth Holmes was shaken when she witnessed a SARS outbreak, and, upon returning home, did not leave her room for a week while working on a patent idea. Born in 2003 as a result of this flurry of work was Theranos, a Silicon Valley start-up of legendary proportions that promised a revolutionary blood analyzer with the ability to run hundreds of tests from just a finger prick in the comfort of one's own home.[1] Holmes enlisted engineers to design her invention called the Edison; no version of this machine, however, would ever be accurate enough to provide completely reliable results nor capable of performing the full range of tests that Holmes originally claimed.[2] She knew the product did not work, yet insisted on running real tests—on people who had serious illnesses—and reporting those inaccurate results. To grasp the severity of her actions, consider that doctors usually use blood tests to increase or decrease a patient's dosage of medication or diagnose conditions which may require immediate action. Doctors rely heavily on lab results, and inaccurate results could be fatal. The only reason the trials went ahead was because Holmes had been outright lying to investors and patients about how accurate the results were.[3] During the trial period, Theranos performed tests on third-party machines traditionally used for blood testing, and desperately tried to get accurate results on their own Edison machine.[4] They already promised their Edison machines to the world, and their 140 million dollar contract with Walgreens, which required them to launch in February 2013, was four months overdue. Some engineers at Theranos noticed the issues and, after raising concerns that only fell on deaf ears, quit. They spoke with journalists and shared that Theranos had duped everyone by testing samples on existing third-party machines, not Theranos' own Edison machine. Eventually, the Wall Street Journal exposed Theranos in the fall of 2015.

A woman once held as a Silicon Valley visionary is now facing federal fraud charges. Holmes fooled everyone, receiving over 600 million dollars of funding in the process. She and her company knew it did not have the results promised, and they knowingly forged the accuracy of the results. From a development framework, Theranos did not seem to care about the accuracy of their work; rather, they simply aimed to launch a product that contained the engineering components traditionally required in development: data confidentiality, integrity, and availability. Although Theranos built a product that meets these three, widely accepted engineering requirements, they cannot escape liability for the inaccurate output of their machines. Even though developers[5] traditionally do not take into account the legal ramifications of accuracy of their data, new policies should be developed that require accuracy as one of the necessary engineering

[1] Jennifer Couzin-Frankel, *The rise and fall of Theranos*, 360 Science 720–720 (2018).

[2] In re Ariz. Theranos, Inc., Litig., 308 F. Supp. 3d 1026, 1041.

[3] *Id.* at 1048.

[4] Eleftherios P. Diamandis, *Theranos phenomenon: promises and fallacies*, 53 Clinical Chemistry and Laboratory Medicine (CCLM) (2015). This included stacking six mini Edison machines on top of each other to get a higher output of tests. The additional heat generated actually hindered the accuracy of tests further, but Theranos did not have time to work on their technology

[5] A programmer, coder, or software engineer is a person who creates computer software. The term computer programmer can refer to a specialist in one area of computers, or to a generalist who writes code for many kinds of software.

properties in software development. Even if accuracy is not added as a necessary fourth component to the "confidentiality, integrity, and availability" trio, and therefore do not hold developers accountability for inaccuracies, other people, such as policy makers and lawyers, will call to action on their own by noting the implications and instituting legal policies accordingly. However, if the requisite knowledge of the developers is not accurately communicated to policy makers and lawyers, how will policy makers and lawyers create effective policies that will ensure developers do, in fact, control for accuracy? If the developers do not have sound, logical engineering reasons to improve a process, it is unlikely they will take legal policies seriously, as industry standards (social norms in software development) are powerful forces that are often unaccounted for during legislative proceedings.

Foremost, I will discuss how poor accuracy can lead to catastrophic security flaws. Poor accuracy can lead to: 1) security flaws; 2) security vulnerabilities; and 3) security inefficacies. Reporting data results as present when, in fact, none exist, sends users looking for something that cannot be found. Even worse, this could cause a security vulnerability to be overlooked and leave an unanticipated security flaw to be discovered by a malicious party. While the security vulnerabilities caused by Theranos' data inaccuracies were not exploited, they could have been if the project had continued forward. Theranos attempted to use networking communication to send data to blood labs and notify the consumers. The data inaccuracies resulted in a security vulnerability due to the sheer lack of effective mitigation of easily anticipated attacks that could breach standard security procedures due to the inaccurate outputs the artificial intelligence ("AI") is taught to accept. Because the blood test results Theranos recorded were inaccurate, it is possible that security could have been impacted as a downstream consequence of the compromised results that were shared via network communications in attempts to correct the accuracy issues. Developers may think inaccurate data has no impact on how their AI is viewed; after all, how is the developer supposed to control the data quality? Their AI is only intended to process data; it is not their fault the data source was unreliable. However, developers should realize that if the data their artificial intelligence is consuming is inaccurate, there are substantial security implications such as allowing unauthorized users access to data through various AI methods of verification. In addition, inaccurate data will bog down the efficiency of AI and create undesirable biases. Consequently, developers should be more readily willing to protect their AI from consuming data that will create biases that do not accurately reflect logical results.

Instances of data inaccuracy are a serious security vulnerability, as AI capabilities heavily rely on massive data sets. As AI capabilities increase, security capabilities need to keep pace, otherwise user data remains at risk. There is a gap in current data protection law and this gap requires intervention. AI, security vulnerabilities, and data accuracy provide a purpose for thinking about the required intervention in relation.

With respect to data inaccuracy, first, I will discuss the current state of information and data security, focusing on the "CIA triad." Next, I will cover the important difference between data integrity and data accuracy. Subsequently, I will explore why a broader conception of information and data security is needed with respect to AI and Internet of Things ("IoT"); specifically, why the accuracy of data should be of concern to developers and any party evolved in the life cycle of data. I will go on to explain the important difference between security and privacy, which are often used interchangeably in everyday life and in legal scholarship.

Afterwards, I will report how accuracy in data processing must be considered a vital benchmark of data security in the realm of AI and big data consumption. Following that, I will examine whether it is possible for AI to be objective, as AI biases come from data. Additionally, the following subsections will provide an overview and analysis of current laws and policies surrounding AI and data security, such as regulating bias in artificial intelligence, accuracy versus explainability, accuracy in the Fair Credit Reporting Act, the General Data Protection Regulation's Right to be Forgotten and its implementation, and the future of AI legislation. Finally, I will conclude by discussing the disconnect between computer scientists and lawyers, which is where accuracy issues manifest.

I. CIA Triad

Confidentiality, integrity, and availability, renowned as the "CIA triad," are the foundational building blocks of information security.[6] The model is designed to guide policies for information security within an organization and serve as a standard for robust and wholesome development. An attack[7] on any one of these items should be anticipated by the developer's product and information system, and based on which of these items is being compromised the most, efficient security controls should be crafted accordingly.

A. Confidentiality

Confidentiality is used to maintain secrecy and ensure it remains undisclosed to unintended parties or individuals.[8] Measures undertaken to ensure confidentiality assures that sensitive information is accessed only by an authorized person and kept away from those not authorized to possess them. Access must be controlled to only limit those authorized to view the data in question. It is common for data to be categorized according to the amount and type of damage that could be done should it fall into unintended hands. Sometimes safeguarding data confidentiality may involve special training for those privy to such information. Such training would typically include security risks that could threaten this information, such familiarizing authorized people with risk factors and how to guard against them. Maintaining strong password-related best practices and information about social engineering[9] methods would work to prevent users from bending data-handling rules with good intentions and potentially disastrous results. Encryption, user IDs and passwords, and two-factor authentication are standard procedures that have become the norm in maintaining data confidentiality. Other options include artificial intelligence dense methods, such as biometric verification.[10]

B. Integrity

[6] Marko Cabric, *Confidentiality, Integrity, and Availability*, Corporate Security Management 185–200 (2015).

[7] An attempt to gain unauthorized access to information resource or services, or to cause harm or damage to information systems.

[8] Unit 7: Maintaining Security and Confidentiality, Educational Testing 116–119.

[9] A form of attack that relies heavily on human interaction and often involves manipulating people into breaking normal security procedures.

[10] Any means by which a person can be uniquely identified by evaluating one or more distinguishing biological traits, such as, fingerprints, facial geometry, retina and iris patterns, voice waves, and DNA.

Integrity involves maintaining the consistency and trustworthiness of data over its entire life cycle.[11] Data must not be changed in transit, and steps must be taken to ensure that data cannot be altered by unauthorized people (for example, in a breach of confidentiality). In the information security context, integrity means that when a sender sends data, the receiver must receive exactly the same data as sent by the sender. Any changes of the data during transit would mean the integrity has been compromised. Measures to protect integrity include utilizing include file permissions and user access controls. Version control[12] may also be used to prevent erroneous changes or accidental deletion by authorized users. In addition, some means must be in place to detect any changes in data that might occur as a result of non-human-caused events such as a server crash. Some data maintains verification of integrity by means of cryptographic checksums.[13] It is important for organizations to have backups be available to restore the affected data to its correct state.

C. Availability

Availability is the condition wherein information is available to the authorized parties whenever required.[14] Serious implications can occur if data is unavailable to systems. Software and system engineers know it is essential to have plans and procedures in place to prevent or mitigate data loss as a result of some unanticipated event. Ensuring rigorously maintained hardware by performing hardware repairs immediately when needed, maintaining a correctly functioning operating system environment that is free of software issues, and keeping current with all necessary system upgrades are important in maintain data availability. Providing adequate communication bandwidth[15] and preventing the occurrence of bottlenecks[16] are also equally important. Safeguards against data loss or interruptions in connections must include unpredictable events such as natural disasters and fire. To prevent data loss from such occurrences, backup copies may be stored in a geographically-isolated location. Extra security equipment or software such as firewalls and proxy servers can guard against downtime and unreachable data due to malicious actions such as denial-of-service attacks and network intrusions.

II. Accuracy v. Integrity

Most people assume data integrity and data accuracy are synonyms of one another; however, this assumption is false. Data integrity speaks to the validity of the transmitted data.[17] Error checking processes and validation measures are heavily relied on to ensure the integrity of data that is transferred or reproduced without the intention of alteration. Data accuracy is a subset of data quality, and data quality is often used in describing how reliable and consistent the data

[11] Jay-Louise Weldon, *Maintaining Data Base Integrity*, Data Base Administration 133–144 (1981).

[12] The task of keeping a software system consisting of many versions and configurations well organized.

[13] A value that represents the number of bits (the smallest unit of data in a computer) in a transmission message and is used by information security professionals to detect high-level errors within data transmissions.

[14] Deepak Khazanchi, *Information Availability*, Handbook of Research on Information Security and Assurance 230–239 (2009).

[15] The capacity of a wired or wireless network communications link to transmit the maximum amount of data from one point to another over a computer network or internet connection in a given amount of time.

[16] A point in the enterprise where the flow of data is impaired or stopped entirely.

[17] CHAPTER 4. What Is Data Governance?, Data Integrity and Data Governance 82–95 (2018).

is.[18] It refers to whether the data values stored for an object are the correct values. To be correct, a data value must be the right value and must be represented in a consistent and unambiguous form. For example, one's birthday is October 18, 1987. If a database stored one's birthday as October 18, 1995, the data would be inaccurate because it is the wrong value. In other words, data integrity involves the intent of keeping the data reliable and data accuracy is the plain binary result of its factualness.

Most developers only think of CIA, but not accuracy. Data cannot be considered secure unless it is, in part, accurate, which current standards do not account for. In the case of Theranos, had there been an obligation for the data results to be accurate, their fraudulent actions would not have been as excessive and longwinded. The engineers at Theranos had achieved their golden standard of producing a product that defines three objectives of security: maintaining confidentiality, integrity (likely debatable), and availability. However, even with these three important items, because accuracy was not listed as an objective item or valued at Theranos, the company's ultimate demise resulted as a direct cause of inaccuracies.

III. Artificial Intelligence and Security Vulnerabilities

Artificial intelligence capabilities rely on large volumes of data sets,[19] which itself poses challenges to the CIA triad because of the sheer volume of information that needs to be safeguarded, the multiplicity of sources it comes from, and the variety of formats in which it exists. Furthermore, because the main concern of big data is collecting and making some kind of useful interpretation of the information, responsible data oversight is often lacking, as Edward Snowden exposed when he reported on the NSA's collection of massive volumes of American citizens' personal data.[20]

As a subset of artificial intelligence, the Internet of Things[21] introduces privacy concerns that should require protection of the information of individuals from exposure in the IoT environment, in which almost any physical or logical entity can be given a unique identifier[22] and the ability to communicate autonomously over the Internet or similar network. The data transferred by a certain endpoint may not result in any privacy issues on its own. However, when even divided data from multiple endpoints is gathered, organized, and analyzed, it can yield sensitive information.

IoT security is another challenge because the IoT consists of so many Internet-enabled devices other than computers, which often go unpatched and are often configured with default or weak passwords. Unless adequately protected, IoT things could be used as a separate attack avenue. For example, it was recently established that a network could be compromised through a Wi-Fi-enabled light bulb, such as the Philips Hue bulb.[23] As more and more products are

[18] Introduction to Data Quality, Data-Centric Systems and Applications Data Quality 1–18.

[19] Otherwise known as "big data".

[20] Joseph Verble, *The NSA and Edward Snowden*, 44 ACM SIGCAS Computers and Society 14–20 (2014).

[21] Otherwise known as "IoT", a system of interrelated computing devices that are provided with unique identifiers and the ability to transfer data over a network without requiring human interaction.

[22] A numeric or alphanumeric string that is associated with a single entity within a given system.

[23] I. Nikolaidis, *How Secure is Your Wireless Network? Safeguarding Your Wi-Fi LAN [Book Review]*, 18 IEEE Network 4–5 (2004).

developed utilizing artificial intelligence with the capacity to be networked by neutral or malicious sources, it is more important than ever for developers to consistently consider security implications in product development.

IV. Security v. Privacy

There are important differences between security and privacy. Security and privacy are often used interchangeably in everyday life and in legal scholarship.[24] However, like computer scientists, the public should be able to draw a technical distinction between the study of "security" and the study of "privacy." Privacy involves difficult normative decisions about competing claims to legitimate access to use and alter information.[25] It requires selecting from among different philosophies and choosing how various rights and entitlements should be ordered. Security is the implementation of those privacy choices—it mediates between information and privacy selections.[26] Privacy establishes a normative framework for deciding who should legitimately have the capability to access and alter information, whereas security is simply the implementation of those choices. It is important for individuals to separate privacy from security because the two have very different, and important, consequences.

The world's top data mining[27] company is Acxiom. The company's databases contain information on over half a billion consumers and process over a billion such records each day.[28] Acxiom earns over a billion dollars annually by selling this data to companies that want to market their products more effectively. It is fair to equate Acxiom as the definition of "big data."

In 2003, Acxiom provided a defense contractor with the social security numbers of passengers who flew on JetBlue flights.[29] The passengers' information quickly became public. The disclosure led to intense criticism of the company and to a complaint to the Federal Trade Commission. In addition, in 2002 and 2003, hackers penetrated Acxiom's computers, accessing the records on millions of American consumers.[30] Acxiom failed to detect the breaches; in fact, law enforcement detected the attacks. Even more embarrassingly, while law enforcement was investigating the initial case, it discovered a second group of hackers who had broken into Acxiom's server three times in the last year.

Astoundingly, the data mining giant had exposed sensitive consumer data three times—once by a deliberate choice and twice through incompetence. While legal scholars tend to blend privacy and security, as stated above, they are separate issues. Acxiom should have anticipated potential attacks and actively worked to keep their systems robust and secure. The law says crimes with intent or malice should be punished more harshly than crimes of negligence, however as a result of the lack of anticipation for the predictable and foreseeable attacks that occurred with Acixom, corporate incompetence should be weighed more heavily or at least

[24] Jon L. Mills, *Privacy: the lost right*, 46 Choice Reviews Online (2009).

[25] Julie C. Inness, *Information, Access, or Intimate Decisions About Our Actions? The Content of Privacy*, Privacy, Intimacy, and Isolation 56–69 (1996).

[26] Susan Landau, *Security and Privacy: Facing Ethical Choices*, 12 IEEE Security & Privacy 3–6 (2014).

[27] The practice of examining large databases in order to generate new information.

[28] Acxiom: Fuelling Marketing with Big Data, Big Data in Practice 103–109 (2016).

[29] In re JetBlue Airways Corp. Privacy Litig., 379 F. Supp. 2d 299, 305.

[30] Florian W. Bartholomae, *Networks, Hackers, and Nonprotected Consumers*, SSRN Electronic Journal (2013).

equally with corporate malice.[31] Acxiom was at the forefront of artificial intelligence technology utilizing data mining methods, and did not take adequate steps to ensure the safety of the data. As artificial intelligence capabilities increase, security capabilities need to keep pace. It does not matter how cool and new Acxiom's data mining algorithm is, Acxiom should be prepared with security practices in place to protect the data it uses. Some may argue that Acxiom cannot be blamed for their lack of preparedness because data mining had not popularized at the time, therefore how would they know how to protect against attacks if no one has been attacked before? By implementing a new method of reading and profiting from data, Acxiom assumed the risk of handling such data and all its consequences. If Acxiom was not willing to think ahead for possible breaches or attacks, then they were not responsible enough to handle the dense role of being at the forefront of implementing new technologies. Acxiom had tremendous security failures that were overshadowed by their meek privacy ones. Companies and individuals should be held accountable for not implementing security practice to protect its data and placing its users at risk, and the law should punish security failures more readily and harshly than privacy ones.

V. Data Processing: Garbage In, Garbage Out

Accuracy in data processing must be considered a vital benchmark of data security in the realm of AI and big data consumption. In computer science, "garbage in, garbage out" describes the concept that flawed, or nonsense input data produces nonsense output or "garbage." If machine learning algorithms are being fed inaccurate, "garbage" data, then one can readily anticipate that the AI's analysis will output"garbage" results. For artificial intelligence, the data quality of the output depends on the quality of the input. With bad data,[32] applications with AI capabilities, such as chatbots[33] or personal assistants, will produce results that are inaccurate, incomplete, or incoherent. Having good, reliable data is especially important for AI subsets like machine learning, which gain greater capabilities over time by analyzing large sets of data, learning from them, and ultimately making adjustments that make the applications more intelligent.

Before feeding data sets into a machine learning application, the data should already be vetted, accurate, consistent, and useful enough for the model to learn from. Data can be gathered from any number of sources, and with that comes the possibility that the data is not complete or fully accurate. To ensure your data is high quality, and therefore useful, it needs to be pre-processed before being used in a model. Otherwise, the risk of putting "garbage" in stands. The data should also be properly cleaned; data cleaning[34] must be carried out when potential issues have been identified with the data set. With unclean or otherwise "garbage" data, machine learning software will not be able to produce results that are accurate or complete. This, in turn, builds models that learn from bad examples. Clean, accurate data ensures data-dependent tasks will not produce "garbage" visuals, models, and organization. By ensuring the datasets collected

[31] Richard T. Sylves & Louise K. Comfort, *The Exxon Valdez and BP Deepwater Horizon Oil Spills*, 56 American Behavioral Scientist 76–103 (2012).
[32] An inaccurate set of information.
[33] A computer program designed to simulate conversation with human users, especially over the internet.
[34] The process of taking incomplete data sets and filling in missing values or removing them altogether, along with removing noisy data and outliers.

are accurate, machine learning tools are more favorable to operating as intelligently as the person who took the time to care for its information.

Now, the elephant in the room is exposed–who is this person that should vet the information and be held responsible for the verification process? Presently, data from the information holder is received by the database, without any sort of screening or vetting. Developers work to sort the data by data types, but this sorting task does not involve any sort of screening, nor should it. The developer's primary role is to make sure the original data can be processed and produce output per the programmed logic; would it make sense to hold responsible the developers for the originating data's accuracy? People believe the developers should be the control of the originating data accuracy; however, I think this theory is too outlandish and not well thought-out. It is too easy to point the finger at the developer in the lifecycle of data. The developer's core security foundation requires the data be kept confidential, maintain its integrity, and be available to the owner. Placing the responsibility of data accuracy on the developer may compromise the integrity component of the engineering properties. The developer should not have the ability to manipulate data, as doing such removes the authentication of what was collected and jeopardizes the potential usefulness of the analytical result. Why should the developers be made to make to the significant decision of what is accurate or inaccurate? This idea of placing responsibility on the developer will greatly increase developer's work and may overall inhibit the development process due to their time being consumed from the vetting phase. It may be more viable to place responsibility on the developer for the data output's accuracy, but even so, the same issue remains for the developer; data is just data, and it is not their place to interpret it and distinguish right from wrong. The developer's subjective influence should play a role in development matters, even though it does not have any other viable phase for it to exist in, as it could not still be expected to yield a fair and objective product with such influences. Developers are simply a stakeholder in the process, not the single point or source of influence.

The interpretation of data is typically reserved for the data analyst,[35] whose training greatly differs from the software developer. Perhaps, the data analyst can play a more significant role in the lifecycle of data. Presently, the data analyst reads into what the developer and their AI produces, which means the data analysts role comes up much later than the initially collection of the dataset. That said, it does not seem quite right to insert the data analyst at two separate points in the process, though their input could not induce any harm. The responsibility should not fall of on the developer or the analyst.

Perhaps, history holds the key for finding the solution of who should be responsible for vetting–after all, computers have only been around for 70 years. How was accuracy being checked without computers? Traditionally, editors played a key role in checking for data accuracy. Once authors created their paper, it underwent a formal editing process with a whole other entity reviewing for accuracy. Perhaps, a similarly simple solution can be applied to review digital data. Inserting a new entity between the information holder and the database that will undergo the data cleaning process would resolve the vetting problem. However, this entity would

[35] A data analyst interprets data, patterns, and trends; turning it into information which can offer ways to improve a business, thus affecting business decisions.

need to be well-trained in the data's context and not allow subjective influences to take ahold. AI biases come from data, and the entity controlling the data holds a powerful influence of the AI's output.

VI. Objective Artificial Intelligence: Is It Possible?

AI is a reflection of its originator. To understand how bias is becoming embedded into AI, one must examine how it is being built. First, AI developers design the applications by forming the questions the AI will solve for and deciding how it will do so. Next, the AI is fueled with the developer's logic (i.e. models) and the originating source of knowledge (i.e. data). Finally, the AI outputs results, responses, and insights based on the developer's design decisions and the knowledge it has been provided. Because the AI is applying its logic to massive amounts of data, the biases within the logic and the knowledge become amplified. Abruptly, the developer is presented with a huge projection of a trait they may not have even known they possessed. AI, typically in the form of machine learning models, due to its reliance on big data, reveals the originator's biases through the patterns of interaction and forms of discrimination that are imbedded into it because the originator is not able to consistently be conscious of or eradicate their own implicit biases as they are creating these systems.[36]

These biased AI results are not just a reflection of present-day biases, but also of historical biases, which have decided who can live in desirable areas, acquire a strong education, and obtain decent work.[37] In addition to the conscious and unconscious discrimination AI developers embed into their designs, because minorities make up the largest percentage of low-income households, they often do not have access to the internet, devices, and apps necessary for contributing to the datasets that fuel AI systems.[38] Since AI began back in the 1940s, minorities have predominantly been underrepresented in the technology industry. For example, in 2015, Google's machine learning models which power their search engine resulted in African Americans when prompted to find images of gorillas. Google's infamous photo search disaster is a direct result of the lack of diversity in its datasets.[39] This specific incident occurred not because developers were actively trying to reinforce racial insults, but because the developers likely commenced training their machine learning models with their own data which is likely mostly comprised of White and Asian ethnicities as most of Google's developers are of those ethnicities. Because of the non-diverse sources of the data, it is likely that Google did not use photos that represented a wide diversity of races, or even gave a thought about potential biases as the developers were likely hurried to achieve acceptable accuracy for every possible search query.

The most basic method of solving for bias issues is to start by gathering more diverse datasets for the machine learning models to process. Having diverse data is the starting point for creating AI that takes all sorts of diverse bodies into account. How bias ends up being handled in

[36] Leon Bottou, *How big data changes statistical machine learning*, 2015 IEEE International Conference on Big Data (Big Data) (2015).

[37] Matthew Hutson, *Even artificial intelligence can acquire biases against race and gender*, Science (2017).

[38] Household Internet access, Hows Life in the Digital Age? (2019).

[39] Leandro Rodrigues Manso Silva, Carlos Augusto Duque & Paulo Fernando Ribeiro, *Power Quality waveform recognition using Google Image Search Engine (iPQ-Google)*, 2016 17th International Conference on Harmonics and Quality of Power (ICHQP) (2016).

AI postures one of the major threats to the success of the utopian vision for objective AI. Bias is rooted in ethics and values—two things that do not have a simple right or wrong answer.[40] They are imaginary laws created and enforced by humans, and humans do not all share the same opinions on what those laws should be.

Who is responsible for determining what laws AI should abide by? The smartest security specialists? The richest technology company? The most powerful country? What if AI is created to be absolutely politically-correct at all times–is the perspective not biased in and of itself? What if AI prevents creating biased content or ideas, and everything just becomes the same? A standardized world would result in none of the beauty or inspiration that arise from human differences. It is easy to state that an objective world would be a fairer one, but where does equity come into play? These questions do not have a simple answer. As they are programming logic to analyze these data sets, developers must engage in continual self-critique and analysis of the extent to which their personal biases inform their work.

VII. Artificial Intelligence and the Law

While it appears that developers do not prioritize accuracy, it is not the same case for policy makers and lawyers. Policy makers are well aware of challenges artificial intelligence poses on data protection law and are increasingly anxious about the need to balance innovation and data protection in artificial intelligence and other data-driven technologies.[41] Within the following subsections, I will be providing an overview and analysis of current laws and policies surrounding AI and data security.

A. Regulating Bias in Artificial Intelligence

Current non-discrimination laws cover specific sectors such as housing, employment, credit, and specific groups of people who might be the victims of discrimination because of their race, gender, religion, national origin, age, or disability. AI technologies being used are no exemption from these rules. The Equality Act[42] should be considered towards amending current discrimination laws to include sexual orientation and gender identity as protected characteristics of discrimination. However, even though such developments in discrimination law would impact the use of AI systems, they are not part of AI governance. Perhaps motivating companies to go beyond anti-discrimination law to assess possible discriminatory impacts on other groups and in other contexts would be more helpful. But, how would it be decided which contexts and which groups would be further assessed? In addition, why is it these investigations only occur when new AI systems are involved, and why are older statistical techniques not being analyzed? Overall, restructuring of anti-discrimination laws will be necessary; however, it is not a matter solely discussed in privacy or AI law.

[40] Nick Bostrom & Eliezer Yudkowsky, *The ethics of artificial intelligence*, The Cambridge Handbook of Artificial Intelligence 316–334.

[41] Data Protection and Privacy by Design, A User's Guide to Data Protection: Law and Policy.

[42] A bill in the United States Congress, that, if passed, would amend the Civil Rights Act to prohibit discrimination on the basis of sexual orientation and gender identity in employment, housing, public accommodations, public education, federal funding, credit, and the jury system.

B. Accuracy vs. Explainability

Under certain circumstances, explanations for how calculated results came to be are required under governing law. Machine learning models create new concerns of explainability[43]. Even when the underlying algorithm is transparent to the user, the derived models are often difficult to understand and explain because the pattern of interactions is incredibly complex and often uses clusters of factors that make no intuitive or abstract sense. The Defense Advanced Research Projects Agency is familiar with the negotiation between accuracy and explainability, and has implemented research efforts aimed at increasing the level of explainability for each level of accuracy.[44]

The compromise between accuracy and explainability need not be dictated by a universal approach represented in law. The most ideal compromise would differ by area because the risks and benefits of analytic techniques depend less on their intrinsic characteristics and more on their field of use. For example, within a context of health care, data scientists and medical professionals are aware of the dangers of relying on correlations that might reflect confusing treatment variables which may have dire results on their patients. The correlations reported in the data are real, but it does not reflect real-world knowledge that doctors know from practice. For this reason, doctors know to always consider multiply sources and contexts before implementing a solution on their patient; the context of the AI output is only one variable.

For this reason, some researchers intentionally use less accurate statistical models that allow them to see clearly the effect of each factor on the variable being predicted.[45] As a result, the researchers can avoid problems such as a confounding treatment variable hidden in the models created by more accurate, but less comprehensible, machine learning techniques. It would be risky to use all-encompassing privacy or AI legislation to reduce these complex, context-dependent questions about explainability and accuracy.

C. Accuracy and The Fair Credit Reporting Act

The federal Fair Credit Reporting Act[46] ("FCRA") regulates consumer reporting agencies and consumer reports. Certain rights exist under this federal law, which include the right to access one's credit report file, the right correct any inaccuracies in your credit report, and the right to seek damages against those who violate the law. The FCRA mandates that consumer reporting agencies use reasonable procedures for collecting, maintaining, and distributing information. As of July 1, 2010, the FCRA requires anyone supplying information to consumer reporting agencies—including original creditors and debt collectors—to have reasonable policies and procedures for ensuring the accuracy and integrity of the information they report.[47] A person can only request and receive data about another person from a credit reporting agency with a valid reason. Valid, as defined by the FCRA, specifies those who have a true need for access,

[43] The extent where the AI's output values are related to its model prediction in such a way that humans understand.

[44] Joichi Ito, *The Limits of Explainability*, Joi Itos Web (2018).

[45] Alex John London, *Artificial Intelligence and Black-Box Medical Decisions: Accuracy versus Explainability*, 49 Hastings Center Report 15–21 (2019).

[46] 15 U.S.C. § 1681.

[47] Julie Hollar, *Fairness & Accuracy in Reporting (FAIR)*, Encyclopedia of Activism and Social Justice.

such as creditors, potential creditors, insurers, employers, landlords, and certain other businesses, such as utility companies.[48]

Under the FCRA, and aside from just inaccurate information, people have the right to dispute both the accuracy and the completeness of items in their file. The distinction between accuracy and completeness is important. For example, a credit report might state accurately that a creditor has sued a person. However, this information may be incomplete because the person later paid off the debt or is not actually liable for it. It is possible to dispute the information about the lawsuit because it is incomplete, and such inaccurate, incomplete, or unverifiable information usually must be corrected within 30 or 45 days.[49]

If someone collects inaccurate data on a person, who is imposed with the burden to correct it? It seems unfair to make the person whose data was collected to correct it–after all, they were not the party to submit the data, yet they are the party being punished for its inaccuracies. The FCRA was enacted by Congress in response to abuses in the consumer reporting industry. Employers had begun to place increased reliance on consumer reporting agencies to provide background checks on potential employees. It was found that agencies were frequently reporting inaccurate information that was harmfully impacting individuals. The FCRA was intended to offer consumers with a remedy for the damage caused by inaccurate credit reports.[50] 15 U.S.C. § 1681(e)(b) requires that a credit reporting agency follow reasonable procedures to safeguard full accuracy in their credit reports. The courts of appeals has been unable to determine who bears the burden of proof in actions brought under § 168l(e)(b); should a plaintiff show that the agency did not use reasonable procedures in maintaining a plaintiff's credit report file, or should the agency show that it did use reasonable procedures? In addition, the people's privacy is a major concern of § 1681(e)(b). However, while it was intended to prevent unreasonable and careless invasions of consumer privacy, it did not intend to prevent the spreading of critical credit data.[51] Congress implemented a variety of measures when it ratified the FCRA to ensure that credit reporting agencies disclose accurate information. Particularly, 15 U.S.C. § 168l(e)(b) provides that "[w]henever a consumer reporting agency prepares a consumer report it shall follow reasonable procedures to assure maximum possible accuracy of the information concerning the individual about whom the report relates." When accuracy is in dispute, the FCRA requires specific procedures a credit reporting agency must follow to ensure a thorough investigation. If an agency either intentionally or negligently fails to follow reasonable procedures, the inflicted consumer has a right to sue. However, the statute does not disclose who bears the burden of proving that the agency did not implement proper reasonable procedures in generating the credit report.

[48] Alexandra Power Everhart Sickler, *The (Un)Fair Credit Reporting Act*, SSRN Electronic Journal(2016).

[49] Kevin J. Smith & Lindsay Colvin, *An Employers Guide to the Fair Credit Reporting Act*, 43 Employment Relations Today 93–98 (2017).

[50] Development and Regulation of Consumer Credit Reporting in the United States, The Economics of Consumer Credit (2006).

[51] G. Allan Van Fleet, *Judicial Construction of the Fair Credit Reporting Act: Scope and Civil Liability*, 76 Columbia Law Review 458 (1976).

Unsurprisingly, when a statute is silent regarding an aspect of legislation, conflicting interpretations will inevitably develop. Consequently, the circuits are undecided on the issue. Some circuits have inferred that the burden lies on the reporting agency, to prove that they indeed followed reasonable procedures. In contrast, other circuits have patently placed the burden on the plaintiff to show that the agency failed to follow reasonable procedures.

Litigation is a primary tool to allow for consumers to protect their privacy.[52] When credit reporting agencies are confronted with the potential of expensive litigation, they have a very pressing incentive to comply with consumer protection laws. Nevertheless, the courts should still recognize the difficulties a plaintiff has in bringing forth a claim under the FCRA, and configure the law appropriately.

Equifax, TRW, and Trans Union are the three main credit bureaus that together control about ninety-nine percent of the market.[53] Because of their role as data collectors, these agencies also have sole or better access to critical information and have the files and the personnel who regularly make records. Plaintiffs are clearly disadvantaged when bringing suit under the FCRA, because they have to go against a sizable and powerful adversary with virtually endless financial resources and access to all of the information a plaintiff would need to be win their case. Courts should acknowledge this disparity of power and allow claims under § 1681(e)(b) to advance by the plaintiff simply providing the inaccurate report. This way, the plaintiff definitively will hold the burden of proof, but it will be a lighter burden.[54] In addition, credit reporting agencies can then easily rebut the case. If agency indeed acted properly, they can rightfully produce records and testimony and easily exonerate themselves.

The courts should take extra precautions when infringing on consumer's privacy rights. While the courts do not have the authority to alter the language of the statute, they can apply the statute liberally and in a way that protects individual privacy rights. In addition, Congress should recognize that the most vital instrument in protecting consumer privacy is the consumer. Thus, Congress should not take away the means that consumers need to bring successful claims against credit reporting agencies when their privacy rights are violated and provide consumers a fair opportunity to litigate their claims.

D. The Right to Be Forgotten: General Data Protection Regulation

The most renowned feature of the General Data Protection Regulation ("GDPR") is the Right to Erasure (i.e. the right to be forgotten). However, its implementation is not as simple as it sounds. Article 17 of the GDPR states that data subjects have the right to have their personal data removed from the systems of controllers and processors under a number of circumstances, such as by removing their consent for its processing.[55] Outwardly, complying with statute may be an

[52] Jon L. Mills, *Strategies and Remedies to Protect Privacy*, Privacy 269–304 (2008).

[53] Leo Onyiriuba, *Authorization and Responsibility for Bank Credit in Emerging Markets*, Emerging Market Bank Lending and Credit Risk Control 503–512 (2016).

[54] Douglas Walton, *Solving the Problems of Burden of Proof*, Burden of Proof, Presumption and Argumentation 176–210.

[55] Andrew Denley, Mark Foulsham & Brian Hitchen, *Does the GDPR apply to you?*, GDPR – How to Achieve and Maintain Compliance 7–14 (2019).

overwhelming task, and to add to the complexity, there are many cases where conflicting regulations will prevent the data processor from complying with the request.[56]

In particular circumstances, individuals have the right to have their data 'erased' where the management of the data fails to satisfy the requirements of the GDPR. The right can be exercised against data controllers, who are required to respond from any allegation without excessive postponement, and when the data is no longer necessary for the purpose for which they were initially collected or if the individual withdraws consent to processing, as well as if there is no other reason for processing. Overall, erasure requests will be allowed where data are 'unlawfully' processed, which is potentially difficult to truly define, as there are many reasons why data could be processed unlawfully under the GDPR.[57] For example, the data may be inaccurate, or an element of an information notice may not have been provided to the individual. These examples provide that it is not always evident whether the right is merited for the data to be erased.

If the data controller has made the data owner's personal data public and has been requested to erase the data, the data controller must also inform other controllers who are processing the data that the data owner has requested erasure of those data. The duty is intended to support an individual's rights in a digital atmosphere by taking reasonable steps to meet the data owner's request. Nevertheless, the duty can be wide-reaching and extremely difficult to implement. How is the data controller supposed to identify other data controllers that have used the data owner's data, which has been made publicly available in the public domain? To erase all the personal data, the controller would have to notify anyone to whom it has disclosed such data; this almost seems like an outlandish task. Fortunately for data controllers, non-digital documents which are not stored via the internet, such as unsearchable online data, are not classified as personal data in the GDPR and are therefore not subject to the right to erasure.[58] In addition, some personal data sets are almost impossible to remove from an individual's record, as the data could be backed up on a local machine or server. While such uneditable data sets are in-scope of the erasure right, they still would be considered outside the breadth for erasure removal procedures due to their unalterable nature.

A valuation should be done to decide what an organization can and cannot do where it is infeasible to erase all of an individual's personal data. Data retention and other legal exceptions will likely always apply. However, this does not justify the data controller storing the records on any accessible digital environment.[59] The best protection of personal data would be to archive it into a more protected and locked down system that meets the retention requirements and also goes as far as possible at meeting the data owner's wish for it to be erased. The core of GDPR is

[56] A Brief History Of Data Protection, EU GDPR: A Pocket Guide, Schools edition 11–17.

[57] Paul Voigt & Axel Von Dem Bussche, *Practical Implementation of the Requirements Under the GDPR*, The EU General Data Protection Regulation (GDPR) 245–249 (2017).

[58] Managing Personal Data Internationally, EU General Data Protection Regulation (GDPR): An Implementation and Compliance Guide - Second edition 249–263.

[59] Georgios Papaioannou & Ioannis Sarakinos, *The General Data Protection Regulation (GDPR, 2016/679/EE) and the (Big) Personal Data in Cultural Institutions: Thoughts on the GDPR Compliance Process*, Lecture Notes in Computer Science Maturity and Innovation in Digital Libraries 201–204 (2018).

to keep the data controller focused on best serving the rights of the data owner as much as possible.

Erasure is not a binary area, with right or wrong defined. Every organization, individual, and technology platform used entails different facts and circumstance which will require a case by case analysis. For example, some technologies allow for more strict control of removal of individual records, whereas some provide none at all. The primary goal is to focus on the reasoning that would be given to a judge in court. Would the data controller be confident that it had a justifiable position on taking the proper measures by the data owners? This analysis is best left for judges or regulators that have all the facts documented and ready for dissemination.

E. The Future of AI Legislation

The new privacy laws currently being deliberated by Congress would create new obligations that will apply to AI programs regarding disclosure, consent, access, correction, and reasonable use of personal data. However, policy makers should be mindful about formulating new requirements that apply only to AI systems. It is well-known that AI is deeply context-dependent, and progress in AI technologies will not be achieved by trying to regulate such technology's growth. However, by examining, its use in each different area of application and setting out specific guidelines governing that specific use may be necessary if some regulation is required.[60]

In the interim, the Trump administration's recent Executive Order on AI[61] proposes an interesting idea of instructing agencies to develop regulatory and non-regulatory methods of use concerning AI technologies, per the advisement by the Office of Management and Budget. The pertinent jurisdiction of congressional committees supervising such a process would be a good start in ensure that the assessment is completed accurately, and it may serve to offer additional understanding about the different ways in which further legislation is needed.

VIII. Conclusion

From the perspective of the policy maker or lawyer, it appears that accuracy issues have a pretty simple solution–that is, to make the developers responsible for the accuracy of the data. Lawyers solve problem by reading facts, seeing facts that trigger a rule of law, and applying the law to solve the problem. However, this conventional approach of problem solving lacks a key element–social norms. It would be impossible to implement an effective solution with taking social norms into consideration because people will not be willing to change their ways simply because they are told to. For example, some people like to think stealing information (such as copyrighted movies, music, and T.V. shows) is acceptable, and will continue to steal such information until society effectively communicates otherwise.

[60] S.j. Blodgett-Ford, *Future privacy: A real right to privacy for artificial intelligence*, Research Handbook on the Law of Artificial Intelligence 307–352 (2018).

[61] The executive order outlines one of the five main directives: the Office of Management and Budget and the National Institute of Standards and Technology should establish guidelines and standards to enable the regulation of AI technologies, with the aim of enabling innovation while protecting privacy and national security interests.

Computer scientists do not understand the law, nor are they trained to. There is a disconnect between computer scientists and lawyers, and this disconnect is where accuracy issues manifest. Data confidentially, integrity, and availability are the primary goals of any software developer when programming any sort of application or software. What occurs if the above three elements are met, but the data itself is not accurate? Instances of data inaccuracy have proved to be a serious security vulnerability, as AI capabilities heavily rely on massive data sets. As AI capabilities increase, security capabilities need to keep pace, otherwise user data remains at risk. There is a gap in current data protection law, and this gap requires intervention. AI, security vulnerabilities, and data accuracy provide a purpose for thinking about the required intervention in relation. For these reasons, policy makers and computer scientists have yet to figure out how to best maintain data accuracy in artificial intelligence consumption.