Daniel Akpovi

Katherine Rocca

ENGL 211C

Due Date: 21 June 2024

AI, Privacy, and Data Security

Equifax in 2017 revealed the details regarding 147 million people and became one of the most vivid examples of the consequences of security failures in the world where Artificial Intelligence (AI) controls data. This incident exposed the weaknesses of data processing in contemporary society. At the same time, it revealed a major concern about the emergence of AI as it is as beneficial as it is dangerous to privacy and data security. The advancement in technology through the integration of AI has impacted various fields including healthcare, finance, and much more but at a cost. With AI development and deployment, privacy rights and data security have been violated severely, and there are questions and concerns about the ethical side and the necessity of a strict governance model.

Artificial Intelligence also known as AI, is the ability of a machine, especially a computer, to mimic processes that are associated with human intelligence. They include learning, reasoning, and self-correction. Privacy rights refer to the extent to which individuals have a right to determine what information is disseminated concerning them. Data security aims at shielding the data from destructive forces and unauthorized individuals. The concept of AI goes back as far as the mid-twentieth century, but it is only in recent decades, with developments in machine learning and data analysis that AI has become a part of our existence. As a result, the need of data protection has increased due to the integration of various systems and the emergence of AI

systems as critical tools for processing massive volumes of data (Calo). The importance of privacy and data protection cannot be overemphasized in the development of AI systems, as over time and due to advancements in technology, AI systems deal with increasingly sensitive information, for instance, health records and financial information. The implications of a data breach are severe, ranging from identity theft to financial loss, and the undermining of confidence in digital systems. Therefore, it remains critical to address the relationships between AI, privacy, and data security to protect rights and uphold data systems.

Artificial Intelligence, which many believe will revolutionize industries, exponentially raises the vulnerability of a company's data being hacked. According to a Ponemon Institute report, the average cost of a data breach of AI systems is significantly higher, with breaches costing companies over $4 million (Ibrahim et al.). The breach has been blamed on the fact that AI is complex and sophisticated and tends to complicate security rather than enhance it. As Khan and co-researchers note, due to the large volumes of data that AI systems deal with, they have become rich opportunities for hackers. For instance, data leakage of over 100 million customers' details happened at Capital One in 2019 due to an AI-controlled cloud system malfunctioning (Khan et al.). This compromise not only shows the economic impacts but also portrays artificial intelligence as a system with flaws.

The Capital One breach is most likely to be one of the most suitable examples of using the promise of AI systems. The breach was done through a former employee of the cloud service provider, who embarked on a processual sequence to exploit the misconfigured Web Application Firewall and gain unlawful access to the firm's information. According to a security analyst at Booz Allen Hamilton, Justice Gideon, during a recent personal interview, such breaches underscore the importance of continual monitoring and updating of AI security measures to keep

up with evolving threats (Gideon). This case just proves that there are new opportunities for optimizing business processes and structuring data with the help of AI systems, but at the same time, there are numerous new points for cybercriminals. Shaharyar Khan and co-researchers elucidate further by stating that the exploitation is usually done through more subtle methods like adversarial attacks, in which malicious entities influence the AI software to get around the security system (Khan et al.). Hence, due to their sophistication and self-contained nature, AI systems open doors to a variety of intricate and concealed cyber risks. This duality means that the design, implementation, and even monitoring of AI systems must be reviewed and improved to reduce data breach risks.

In addition, the escalation of AI decision-making in organizational processes escalates the risks of data breaches not only in financial terms but also in reputational and legal ones. Once AI systems are integrated into core activities in many sectors, such as the health and financial sectors, the consequence of a breach is likely to spread throughout sectors and undermine public confidence and investors' as well. For example, Equifax experienced significant backlash from regulators, shareholders, and the public after their data breach, which led to significant penalties and ongoing litigation. This specific instance highlights the need of preventive measures meant to protect personal data by illustrating the possible repercussions of AI-related data leaks (Kabanov and Madnick). Similarly, Ruben Goncalves Miranda and co-authors concur, in their book "Artificial Intelligence and Human Rights," that the need for stringent measures to protect personal data is highlighted by the growing power of private organizations and individuals who can easily access sensitive information for economic benefits, making it crucial to safeguard against these risks (Miranda et al. 25). While organizations continue to incorporate AI systems into their operations, it is fundamental to ensure strong cybersecurity and risk management

measures are in place to minimize the negative effects of cyberattacks on individuals' rights as well as public confidence in emerging innovations.

The use of Artificial Intelligence in surveillance systems has raised massive social issues, including mass surveillance and privacy infringement. According to the Carnegie Endowment for International Peace, 75 countries among 176 use artificial intelligence for surveillance (Feldstein). These include sophisticated facial recognition techniques, effective policing models, and wider data collection and analysis. These technologies are often used without adequate controls and thus make privacy a mere illusion through surveillance. AI does not just refer to the monitoring of activities; it encompasses the analysis of large amounts of data in real time to provide real-time observation and tracking of people.

One of the most well-known examples of AI use in policing is the Chinese government surveillance with the help of facial recognition technology. In Beijing and Shanghai, the Chinese government has put in place AI-powered surveillance systems that can recognize individuals and follow them over vast areas in real-time. Scholar David Karpa from the University of Bremen and co-researchers report that these systems are not only for the surveillance of public areas but also for the implementation of a social scoring system in the country where people's behaviors are supervised and rated (Karpa et al.). They infringe on individual privacy rights as the citizens are monitored without their consent. This is followed by the obscurity and irresponsibility in the collection, storage, and use of information, leading to limitation of the freedom of speech and self-actualization. In the same vein, the scholar Peter Asaro states that with the increasing automation of bureaucratic processes, there are significant concerns about the fairness, accountability, and transparency of the algorithms that influence crucial decisions affecting people's life opportunities and rights (Asaro 41). This growing trend of applying artificial

intelligence in surveillance requires the formulation of policies that adequately deal with issues concerning national security while at the same time not forgetting basic human rights to privacy.

However, it is essential to note that an AI system does not lack biases that are to blame for the misuse of personal data by the systems. Research has shown how these AI algorithms keep on perpetrating and sometimes even enhancing the bias of the training data (Mittelstadt et al.). For instance, one of the instruments of risk assessment used in the criminal justice system relating to recidivism was found to be racist. It was highlighted that Black defendants were considered high risk, at nearly double the number of white defendants with similar criminal histories (Vincent and Viljoen). Similarly, in another study, it was realized that the facial recognition algorithms were at least ten to 100 times less accurate in identifying African American and Asian faces than White faces (Cavazos et al.). Such biases in AI result in unfair treatment and are a sign of a greater issue with data exploitation – data that is obtained and used for self-serve social prejudice.

It is equally important to consider such issues as how biased AI can be in terms of the privacy and security of data. It is more so when AI systems perform prejudicial actions against some groups because it undermines the confidence that people have placed in technology and other organizations. For example, the employment filters that operate as a negative effect on some groups of people can lead to unjust rejection of applicants and result in socioeconomic discrimination. Changwu Huang and co-authors illustrate that other studies demonstrated that discrimination in AI for recruitment was apparent in the 67% reduced call-back rate for applicants of color (Huang et al.). Secondly, personal data within such prejudiced systems might be used for unauthorized profiling and surveillance, which is even less desirable in terms of privacy. This is especially the case when AI is applied in policing or credit risk assessment, in

which disparity in data results in higher prejudice against affected categories. Based on the above differences, awareness and the development of the right ethical AI policies that would ensure that individuals have the right to privacy and be informed are needed.

Moreover, it has been observed that the systematic bias in AI maintains systemic inequality and social exclusion. Prejudice is present not only in unequal treatment but also in decision-making input, in particularly delicate fields like law enforcement, health, and education. According to Jaqueline G. Cavazos and co-authors, research showed that the templates of AI algorithms in diagnostics favor racism and discriminate against patients of color unfairly (Cavazos et al.). Also, the algorithms implemented in education systems, like the standardized testing and admissions, unfairly lock out minority students and promote systems of oppression. Furthermore, the employment of racially discriminatory predictive policing models means that such communities are undesirably policed and surveilled to a greater extent (Asaro). These systemic implications highlight the importance of developing comprehensive approaches to address prejudice in AI systems. Such approaches include, but are not limited to, the acquisition of diverse datasets for training, making algorithms open source, and constant auditing to ensure that all AI-enabled systems are fair and equitable in their decision-making.

Critics of the opinion that AI is a threat to data privacy and security claim that AI dramatically improves data security as compared to data threat elimination methods. AI systems are capable of scanning through vast data in real-time apart from identifying discrepancies that may point to security threats. According to the pioneer in AI research, Andrew Ng, AI-enabled security systems can identify potential attacks and prevent them since the systems can learn behavior patterns that are malicious and pose a threat to the information (Ng). Through

Darktrace, for instance, it can be argued that the proactivity from AI leads to a decrease in the number of data breaches while increasing data security.

However, setting that, AI can identify threats with enhanced capabilities, but it has its strengths and weaknesses and can be misused. For instance, AI can be misused to automate and scale up cyberattacks, create sophisticated phishing campaigns, generate deepfakes for social engineering, and even develop self-evolving malware that can adapt to avoid detection. It is essential to understand that the same tools that are used to safeguard information can be used maliciously by hackers. Paraskevas charged that, for instance, the 2018 British Airways data breach incident that saw the personal and financial data of over 380, 000 customers leaked was enabled by vulnerabilities in the organization's AI systems (Paraskevas). Also, AI is malleable through approaches like adversarial strategies by which slight modifications to inputs can mislead an AI system into producing wrong outputs. This shows that, while AI can bring security benefits, it also creates new threats and challenges that can be leveraged for gain on a substantial scale in terms of security breaches and data misuse.

Moreover, claiming that privacy risks are worth bearing for the sake of future AI innovation disregards privacy as a human right and data protection. In the case of the European Union, the General Data Protection Regulation enshrines the protection of personal data specifically stating that privacy is a human right. On the same note, ethical issues offer an argument as to why the aspect of privacy should be given top priority. Some philosophers like Immanuel Kant opined that it is immoral to treat people as mere tools to achieve certain goals, but as rational beings with rights to personal liberty and a right to privacy. The purpose of ethics and legal measures for individuals is undermined when these principles are disregarded in favor

of innovation. Hence, while AI business solutions create numerous advantages, there must be solid methods to counterbalance them that do not harm privacy and data protection.

AI advancement and general use has caused rampant privacy rights and data security violations; therefore, society requires to reconsider the integration of these technologies. Lacking intervention, AI surveillance breach and leaks endanger civil liberties, which might result in an environment where personal privacy is consistently violated. It is, therefore, high time to enact stricter and more sensible laws and regulations that would prioritize the security and privacy of personal data. As for future development if current tendencies will remain as they are, people might find personal data protection endangered by the advantages of AI. However, with strong privacy and strict regulation, Artificial Intelligence can be used to improve human life while not violating human rights, thus creating a world where technology grows hand in hand with ethics.

Works Cited

Calo, Ryan. “Artificial Intelligence Policy: A Primer and Roadmap.” *U. Bologna L. Rev*., vol. 3, 2018, p. 180.

Cavazos, Jacqueline G., et al. “Accuracy Comparison across Face Recognition Algorithms: Where Are We on Measuring Race Bias?” *IEEE Transactions on Biometrics, Behavior, and Identity Science*, vol. 3, no. 1, 2020, pp. 101–11, https://doi.org/10.1109/TBIOM.2020.3027269.

Feldstein, Steven. *The Global Expansion of AI Surveillance*. Carnegie Endowment for International Peace Washington, DC, 2019. *Google Scholar*, https://blog.fdik.org/2019-09/WP-Feldstein-AISurveillance_final1.pdf.

Gideon, Justice. Personal interview. 11 June 2024.

Huang, Changwu, et al. “An Overview of Artificial Intelligence Ethics.” *IEEE Transactions on Artificial Intelligence*, vol. 4, no. 04, Aug. 2023, pp. 799–819. www.computer.org, https://doi.org/10.1109/TAI.2022.3194503.

Ibrahim, Amani, et al. “The Challenges of Leveraging Threat Intelligence to Stop Data Breaches.” *Frontiers in Computer Science*, vol. 2, 2020, p. 36, https://doi.org/10.3389/fcomp.2020.00036.

Kabanov, Ilya, and Stuart Madnick. *A Systematic Study of the Control Failures in the Equifax Cybersecurity Incident*. 3957272, 2020. *Social Science Research Network*, https://doi.org/10.2139/ssrn.3957272.

Karpa, David, et al. “Artificial Intelligence, Surveillance, and Big Data.” Diginomics Research Perspectives, edited by Lars Hornuf, Springer International Publishing, 2022, pp. 145–72. *DOI.org (Crossref)*, https://doi.org/10.1007/978-3-031-04063-4_8.

Khan, Shaharyar, et al. “A Systematic Analysis of the Capital One Data Breach: Critical Lessons Learned.” *ACM Transactions on Privacy and Security*, vol. 26, no. 1, Nov. 2022, p. 3:1-3:29. ACM Digital Library, https://doi.org/10.1145/3546068.

Miranda, Gonçalves, Rubén., et al. *Artificial Intelligence and Human Rights*, Dykinson, S.L., 2021. *ProQuest Ebook Central*, https://ebookcentral.proquest.com/lib/odu/detail.action?docID=6851958.

Mittelstadt, Brent Daniel, et al. “The Ethics of Algorithms: Mapping the Debate.” *Big Data & Society*, vol. 3, no. 2, Dec. 2016, p. 205395171667967. DOI.org (Crossref), https://doi.org/10.1177/2053951716679679.

Ng, Andrew. “Andrew Ng: How AI Could Empower Any Business | TED Talk.” *TED Talk*, 2022, https://www.ted.com/talks/andrew_ng_how_ai_could_empower_any_business?language=en&subtitle=en.

Paraskevas, Alexandros. “Cybersecurity in Travel and Tourism: A Risk-Based Approach.” Handbook of E-Tourism, edited by Zheng Xiang et al., Springer International Publishing, 2022, pp. 1605–28. *DOI.org (Crossref)*, https://doi.org/10.1007/978-3-030-48652-5_100.

Peter M. Asaro, "AI Ethics in Predictive Policing: From Models of Threat to an Ethics of Care," in *IEEE Technology and Society Magazine*, vol. 38, no. 2, pp. 40-53, June 2019, doi: 10.1109/MTS.2019.2915154.

Vincent, Gina M., and Jodi L. Viljoen. “Racist Algorithms or Systemic Problems? Risk Assessments and Racial Disparities.” Criminal Justice and Behavior, vol. 47, no. 12, Dec. 2020, pp. 1576–84. *DOI.org (Crossref)*, https://doi.org/10.1177/0093854820954501.