

**Investigating Perspectives of Genomics Professionals on the Privacy & Security of Genetic  
Data**

A Research Paper submitted to the Department of Engineering and Society

Presented to the Faculty of the School of Engineering and Applied Science  
University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree  
Bachelor of Science, School of Engineering

**Kyle Stephen Caylor**

Spring 2024

On my honor as a University Student, I have neither given nor received unauthorized aid on this  
assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Advisor

Kent Wayland, Department of Engineering and Society

## INTRODUCTION

The Journal of American Medical Informatics Association published a study in 2014 revealing more than 83% of participants have some level of concern about the protection of their genetic data (Rogith et al., 2014). The striking portion of this statistic isn't the 83% of concerned participants, it is that fact that the study was completed more than a decade ago. As private genomic data technologies become increasingly advanced and pertinent to daily life, more discussions are taking place about the ethical regulation of genetic data privacy. As this booming genomics industry nears 22 billion dollars in value by 2030 (LLP & Kumar, 2022), citizens fear the privacy of their genetic data may become increasingly vulnerable.

To ensure the security of genetic information depends on the development of a robust system of guidelines, policies, and regulations. Nevertheless, the complex network of stakeholders claiming equity in the United States healthcare industry can make this process complicated and time-consuming. Patients, healthcare providers, insurers, and lawmakers all vie for their own personal interests as they relate to topics such as coverage and reimbursement policies. The apparent opposition that exists in the relationship between patients and insurers has perpetuated healthcare inequities and continues to do so with the emergence of gene editing technologies and its high out-of-pocket treatment costs. With such a wide variety of stakeholders taking interest in the legal and economic intricacies of our healthcare system, legislative efforts meant to protect the rights of private individuals becomes seemingly apathetic, which can prove to be worrisome when examining the trajectory of technological growth the medical device industry is experiencing.

However, the scope of genetic engineering expands far beyond the hospital room, encompassing ethical questions about gene editing in fetuses and even implications in insurance

discrimination and genetic warfare. These considerations are shaping how the gene editing industry has grown and will continue to grow. It is important to actively engage with stakeholders in the field to better understand the values and perspectives we share about genetic data privacy. This study will investigate the attitudes towards genetic data that are held by one very pertinent group of stakeholders in order to answer the question: *How is genetic data privacy and security viewed amongst genomics professionals?*

## **BACKGROUND & CONTEXT**

The Genetic Information Nondiscrimination Act (GINA), passed by Congress in 2008 as a federal law prohibiting discrimination by employers and health insurers based on an individual's genetic information, was one of the first milestone legislative actions taken to address privacy of and access to genetic information (Feldman, 2011). However, this policy merely protected the utilization of genetic information, and did little to address how or why the information was acquired in the first place. Additionally, employers are a relatively cordial party to be in possession of genetic information. What other factors come into play when someone with more malicious intentions comes into contact with one's genetic data? The popular company 23&Me provides qualitative genotyping to more than 14 million customers in order to detect variants in their genomic data and provide information about their ancestry and genetic predispositions to health problems. Last year on the 29th of April, a data breach went undetected for more than 5 months, leaking valuable and private health reports and raw genetic data of millions of users (French, 2024). Should this large genetic data leak drive more efforts to increase legislation protecting genomic data privacy?

Many of these genomics companies such as 23&Me and AncestryDNA are private entities, and therefore limited legislation exists that actually regulates the protection of the data collected by these companies. The protection and privatization of genetic information can be a gray area in many situations. While each individual should have the right to do whatever they wish with their genetic information, some ethical questions arise when it comes to the distribution of hereditary genomic data. Many genes are shared between family members, meaning the relevant stakeholders of these genetic privacy questions extend far beyond just the individual.

When trying to investigate a question related to the legislative protection of genetic data – it is important for the reader to be knowledgeable about individual health-related rights such as those outlined in the Health Insurance Portability and Accountability Act of 1996 (HIPAA), which is a federal law that created national standards to protect sensitive patient health information from being disclosed without the patient's consent or knowledge (*Health Insurance Portability and Accountability Act of 1996 (HIPAA) | CDC, 2022*). With the development of medical diagnostic technologies, HIPAA has been amended since its publication to address and include genetic information. In 2013, the HIPAA Omnibus Rule expanded regulations to include genetic information as protected health information (PHI) (*Genetic Information Privacy, 2015*). Additionally, given the complex organizational nature of the healthcare industry, some knowledge of infrastructure (Star, 1999) is valuable in forming opinions about the regulation of genetic data. In particular, the scope of genetic engineering will be thoroughly discussed to outline the stakeholders involved in the decision-making processes associated with the dissemination of genetic information.

Due to the recency in which a vast majority of genomics technologies have emerged, the concept and scope of genetic data privacy perspectives can be considered less extensively researched than other fields. Nevertheless, there are some publications that have helped to expand upon our collective knowledge of the topic. One article published in 2016 entitled “An overview of human genetic privacy” outlined topics such as genomic data sharing, potential genetic privacy risks, and ethical regulations for protecting human genomics data. In this literature the authors even described the field of human genomics as an “active battlefield of data” (Shi & Wu, 2017). Another article published in the *Journal of Law and the Biosciences* investigated the landscape of genetic data privacy regulations. It also commented on the limitations of laws such as HIPAA and GINA in adequately protecting individuals after their genetic information has been disclosed (Clayton et al., 2019). By investigating the perspectives of genomics professionals on topics related to genetic data privacy and security, this paper hopes to collaboratively expand on the work of this prior research and contribute to a more secure ecosystem of genomics.

## **THEORETICAL FRAMEWORK**

In the 1999 work “The Ethnography of Infrastructure”, Susan Leigh Star provides the nuanced framework of infrastructure for understanding the complex dynamics within the genetic data industry. In Star's view, infrastructure encompasses not only physical components but also social practices, policies, and standards that shape the use and governance of technologies. The concept of scope, within this framework, refers to the range and reach of infrastructural systems, including the boundaries they establish and the actors they include or exclude. Applied to the genetic data industry, this notion of scope reveals the interconnectedness of various stakeholders,

from researchers and clinicians to policymakers and technology developers. It highlights how genetic data, as part of an infrastructural system, transcends individual entities and is governed by a network of relationships, regulations, and ethical considerations. Understanding the scope of genetic data infrastructure is crucial for addressing privacy and security concerns, as it involves navigating the intricate web of data flows, access rights, and accountability mechanisms that influence how genetic information is collected, stored, and utilized. By examining infrastructure through the lens of scope, we gain insights into the power dynamics and governance structures that shape the genetic data landscape, paving the way for more informed decision-making and ethical practices within the industry.

## **METHODS**

In this study, the sample population comprises doctors, researchers, and undergraduate researchers affiliated with three genomics labs at the University of Virginia, the Genomics Laboratory, the Pediatrics Genetic Research Lab, the Cytogenetics Laboratories. This diverse cohort represents a range of expertise and perspectives within the genomics field, encompassing seasoned professionals with extensive experience in clinical practice and research, as well as emerging scholars at the undergraduate level engaging with genomics studies. By focusing on individuals actively involved in genomics research at the university setting, valuable insights are gained about the attitudes and perspectives regarding genetic data privacy and security from those directly engaged in the field's advancement. Furthermore, the inclusion of undergraduate researchers adds a unique dimension to the study, allowing for exploration into how future professionals are being introduced to and grappling with the ethical and practical considerations surrounding genetic data privacy within their educational and research experiences.

The survey was administered electronically via email to the selected participants at genomics labs at the University of Virginia. The survey initially sent to the faculty directors of the laboratories, who subsequently distributed it out to their comprehensive list of lab workers. The survey consisted of a series of eight structured questions designed to elicit responses on a variety of topics related to genetic data privacy, including concerns about data breaches, confidence in current security measures, perceptions of the need for legislative intervention, and attitudes towards the use of secure technology services. Participants were asked to rate their responses on a scale of 1 to 5, with 1 indicating the lowest level of agreement or concern, and 5 indicating the highest. Additionally, participants were able to answer “unsure” to any of the questions if they felt uncomfortable or inadequately equipped to answer any particular question. No personal identifiable information (PII) was collected throughout the completion of the survey, and all survey responses were recorded anonymously. The full survey questionnaire can be found in Appendix A. The survey responses, all numerical ratings on a scale of 1 to 5, will be analyzed quantitatively. The survey results were analyzed by reviewing participant responses across questions to identify trends and patterns. Responses were quantified by calculating the percentage of participants who responded with each option. This allowed for a clear analysis of the distribution of opinions and the prevalence of certain viewpoints within the sample population. This approach will enable the identification of key trends within the responses, shedding light on the priorities among genomics professionals at the University of Virginia.

## **RESULTS & DISCUSSION**

In total, survey responses were collected from 45 genomics professional participants. The first question on the survey asked participants about their overall concern for the potential of

genetic data security breaches. Of the 45 survey responses, only two participants responded as having no concern at all about the security of genetic information, and approximately 27% of participants indicated that they were more than moderately concerned about this security. These responses, especially when coming from genomics professionals themselves, indicate a high level of concern with the ability of existing infrastructure to protect the genetic data of individuals. Addressing these concerns is essential to foster a climate of confidence and ethical responsibility in the handling and utilization of genetic data in the healthcare industry.

The second question on the survey addressed the existing efforts governments have made to ensure the security of genetic information. It asked participants about their confidence in these measures in minimizing genetic data storage, which is the principle of limiting data collection and retention to the bare minimum necessary to accomplish a given purpose (“What Is Data Minimization and Why Is It Important?,” 2024). This principle is a key foundation to many data privacy laws and regulations, as it seeks to ensure no unnecessary information is being collected and stored about an individual’s genetic data. More than 30% of participants replied “unsure” to this question, which highlights a possible collaborative gap that exists between legislators and those who actually work in the field of genomics. It is imperative that the policy-making process to protect genetic information be a joint effort between those who hold political power and those who are more knowledgeable about the actual implications of a genetic data security. This idea once again relates back to the Star’s concept of scope, as it demonstrates the breadth of the field of genomics and highlights a lack of collaboration between the aforementioned stakeholders. Additionally, thirteen participants indicated only slight confidence in the existing efforts that are in place to protect genetic information, highlighting a necessity for more efforts to be made to



protect this valuable information, a concept which was reinforced by the following question on the survey.

The third question of the survey sought to elucidate perspectives on the need for more comprehensive and rigorous legislation protecting genetic information. More than half of all participants indicated a significant need to expand on the existing data security measures, with six additional responses citing this need as urgent. As genetic technology continues to advance and accelerate, there is an urgent need for stricter legislation to ensure genetic data privacy. This legislation would protect individuals from potential misuse or exploitation of their most valuable information. Enhancing these legal frameworks could ensure more comprehensive protection for individuals, ultimately helping to foster trust in the field of genomics and commercial services such as AncestryDNA, while helping to mitigate the risks of unauthorized access. The fourth and fifth questions shifted the focus of the survey from existing and prospective legislative policies to the responsibility genomics professionals, primarily patient-facing providers, have in communicating the potential value that genetic information holds.

The fourth question inquired about how proactive genomics professionals should be in educating patients and the general public about the importance of genetic data privacy. Approximately 70 percent of participants indicated that these professionals should be very proactive in these efforts. Genomics professionals have a fundamental responsibility to educate both their patients and the general public about the significance and impact of genetic data security. This education can help ensure individuals make informed decisions regarding the utilization and protection of their genetic information. By increasing awareness and understanding of existing policies and potential risks of misuse, genomics professionals can empower individuals to confidently navigate the complex topic of genetic data privacy with

confidence and autonomy. The following question sought to form a link between the extent to which the public is informed and the adoption of secure genomics practices. Of all the responses, 34 participants indicated some level of connection between these two variables. A public well-informed about the value their genetic data can help influence the adoption of more secure practices, helping hold higher accountability for the misuse or misdistribution of genetic data. When properly informed, individuals are likely to demand higher standards of privacy protection, encouraging more comprehensive policies that strengthen the integrity of genetic data protection measures.

The last three questions on the survey once again shifted the focus towards how accountability and competition might influence the sphere of genetic data privacy and regulation. The sixth question asked about how a lack of accountability for genetic data breaches would hinder further advancements in genomics technologies. The results indicated no clear correlation between accountability for genetic data breaches and advancements in genomics research. This suggests that while concerns about data breach accountability may be present among respondents, they may not perceive it as a significant factor that hinders growth in the field. The seventh question read as follows: *How likely are genomics professionals to prioritize data security over convenience when selecting technologies for genomic analysis?* The responses were well-distributed, with about 25% of participants responding “Likely”, 25% responding “Neutral”, and 25% responding “Unlikely”. This indicates that weighing the importance of data security against the practicality and ease of use of available technologies may be more than black and white; it reinforces the complexity in the decision-making process of professionals when selecting which technologies they use. The last question on the survey asked about how industry competition might influence the prioritization of effective data security measures, and more than

60% of respondents indicated a moderate to significant influence. These responses suggest that fiscal pressures and the desire to remain competitive in a rapidly advancing industry are powerful forces that also influence the decision-making processes of professionals.

## **CONCLUSION**

Overall, the survey results were indicative of a lack of confidence in the ability of existing measures to protect an individual's genetic information. They suggested that this lack of protection enables outside factors such as market competition to influence the decision-making process of genomics professionals. Additionally, the results highlighted the vital importance that exists in having a well-informed public. The public's knowledge of and advocacy for genetic data legislation can play a significant role in shaping future policy decisions.

Genetic information is an extremely precious resource. DNA is the fundamental blueprint that distinguishes one individual from another, and thus it should be protected equally, or even more securely than, something as valuable as a passport or social security number. As the field of genomics technology continues to progress, the true utility of DNA will expand greater than ever imagined. Diagnostic technologies will better help identify genetic diseases, generate treatment plans, and even prevent the diseases from ever coming to fruition. Forensic analysis will improve to help solve crimes and increase national safety. Genetically modified organisms, and possibly even humans, will even grow stronger and more resistant to illnesses and injury.

However, it is imperative that adequate safeguards are in place long before this advanced world becomes a reality. This is because genetic information is also an extremely dangerous resource, and as beneficial technology advances, so will threatening technology. Genetic information can be weaponized in the form of engineering bioweapons or pathogens made to

target specific vulnerable populations, posing a significant threat to global security and public health. Future research should be done to analyze the frequency with which genetic information is misused or protected inadequately. Computer science professionals should work collaboratively with genomics professionals to investigate how diagnostic technologies are used and safeguarded at both the public and private level, hopefully identifying any holes through which the information may be leaked or maldistributed. Furthermore, legislators at both the state and federal levels should be encouraged to work hand in hand with genomics professionals to develop more comprehensive genetic data privacy laws and regulations. This will hopefully help to foster a world in which advancements in genomic technologies are remembered for their advancements and not in their shortcomings.

## BIBLIOGRAPHY

- Clayton, E. W., Evans, B. J., Hazel, J. W., & Rothstein, M. A. (2019). The law of genetic privacy: Applications, implications, and limitations. *Journal of Law and the Biosciences*, 6(1), 1–36. <https://doi.org/10.1093/jlb/lbz007>
- Feldman, E. A. (2011). The Genetic Information Nondiscrimination Act (GINA): Public Policy and Medical Practice in the Age of Personalized Medicine. *Journal of General Internal Medicine*, 27(1), 743–746. <https://doi.org/10.1007/s11606-012-1988-6>
- French, L. (2024, January 26). *23andMe: Raw genetic data stolen in months-long cyberattack*. SC Media. <https://www.scmagazine.com/news/23andme-raw-genetic-data-stolen-in-months-long-cyberattack>
- Genetic Information Privacy*. (2015). Electronic Frontier Foundation. <https://www.eff.org/issues/genetic-information-privacy>
- Health Insurance Portability and Accountability Act of 1996 (HIPAA) | CDC*. (2022, June 28). <https://www.cdc.gov/phlp/publications/topic/hipaa.html>
- LLP, S. M. R., & Kumar, S. (2022, September 12). *Genome Editing Market to Hit \$21.36 Billion by 2030, With a CAGR of 17.33% Asserts Strategic Market Research*. GlobeNewswire News Room. <https://www.globenewswire.com/en/news-release/2022/09/12/2514071/0/en/Genome-Editing-Market-to-Hit-21-36-Billion-by-2030-With-a-CAGR-of-17-33-Asserts-Strategic-Market-Research.html>
- Rogith, D., Yusuf, R. A., Hovick, S. R., Peterson, S. K., Burton-Chase, A. M., Li, Y., Meric-Bernstam, F., & Bernstam, E. V. (2014). Attitudes regarding privacy of genomic information in personalized cancer therapy. *Journal of the American Medical Informatics Association*, 21(e2), e320-325. <https://doi.org/10.1136/amiajnl-2013-002579>
- Shi, X., & Wu, X. (2017). An overview of human genetic privacy. *Annals of the New York Academy of Sciences*, 1387(1), 61–72. <https://doi.org/10.1111/nyas.13211>
- Star, S. (1999). The Ethnography of Infrastructure. *American Behavioral Scientist*, 43(3), 377–391. <https://doi.org/doi/10.1177/00027649921955326>
- What is Data Minimization and Why is it Important? (2024). *Kiteworks | Your Private Content Network*. <https://www.kiteworks.com/risk-compliance-glossary/data-minimization/>

## Appendix A

<b>Question #</b>	<b>Question</b>	<b>Response Choices</b>
1	On a scale from 1 to 5, how concerned are you about the potential risks of genetic data leaks and breaches?	(1) Not concerned at all (2) Slightly concerned (3) Moderately concerned (4) Very concerned (5) Extremely concerned (6) Unsure
2	In 2018, the European Union enacted the General Data Protection Regulation Law, which addressed themes of informed consent, data breaches, and data storage minimization. How confident are you in these current efforts employed to protect genetic data?	(1) Not confident at all (2) Slightly confident (3) Moderately confident (4) Very confident (5) Extremely confident (6) Unsure
3	To what extent do you believe there is a need for stricter legislation to protect genetic data privacy?	(1) No need at all (2) Slight need (3) Moderate need (4) Significant need (5) Urgent need (6) Unsure
4	How proactive do you think genomics professionals should be in educating patients about the importance of genetic data privacy?	(1) Not proactive at all (2) Slightly proactive (3) Moderately proactive (4) Very proactive (5) Extremely proactive (6) Unsure
5	How much do you think public awareness about genetic data privacy impacts the adoption of secure practices among genomics professionals?	(1) Negligible impact (2) Minimal impact (3) Moderate impact (4) Significant impact (5) Crucial impact (6) Unsure
6	To what extent do you believe the lack of accountability for genetic data breaches hinders advancements in genomics research?	(1) Not a hindrance at all (2) Slight hindrance (3) Moderate hindrance (4) Significant hindrance (5) Major hindrance (6) Unsure
7	How likely are genomics professionals to prioritize data security over convenience when selecting technologies for genomic analysis?	(1) Very unlikely (2) Unlikely (3) Neutral (4) Likely (5) Very likely (6) Unsure

8	To what extent do you believe that industry competition impacts the prioritization of genetic data security measures?	(1) No influence (2) Minor influence (3) Moderate influence (4) Significant influence (5) Dominant influence (6) Unsure
---	---	--

Appendix A. Survey Questionnaire - Attitudes Concerning Genetic Data Privacy & Security