Enhancing Data Governance and Security: A Strategic Plan for MaineHealth

Position Paper 1, DSE6003

Adeline Casali

September 7, 2023

# Introduction

## Key Governance, Security, and Ethical Issues

The healthcare industry is one of the most susceptible to security attacks, as the data from healthcare companies is typically private and confidential. This sensitive information is highly valuable to hackers, making the industry a large target for attacks (Seh et al., 2020). This vulnerability is particularly pertinent in the era of big data, as the current trend towards digitizing healthcare has led to an enormous quantity of clinical data availability. Although big data holds the promise of opportunities and the potential to lower costs, improve care, and even save lives, it also introduces a host of data security issues due to the quantity and complexity of the data (Abouelmehdi et al., 2017).

A data breach is defined as “a security incident in which sensitive, protected, or confidential data are copied, transmitted, viewed, stolen, or used by an unauthorized individual” (Khan et al., 2021, p. 2). Data breaches are surprisingly common in the healthcare industry. Between 2005 and 2019, 249.09 million individuals worldwide were impacted by healthcare data breaches (*Data Breach Chronology*, 2023). And in 2022 alone, 51.9 million healthcare records were exposed, stolen, or illegally disclosed (Alder, 2023). Furthermore, these breaches can be extremely costly to the organization. According to IBM’s 2023 Cost of a Data Breach report, the average healthcare data breach in 2022 in the United States cost $10.93 million (IBM, 2023). Avoiding data breaches is one of the leading security goals of this plan.

# Data Governance Framework

## Data Governance Policies

As an overview, data governance involves the utilization of oversight, authority, and control in managing data. Its primary objective is to enhance the worth of data while reducing the expenses and potential risks associated with it (Abraham et al., 2019). Protocols must be put into place the clearly define roles, responsibilities, and accountability for data stewardship, quality, and compliance. These policies should align with industry best practices as well as legal regulations such as the Health Insurance Portability and Accountability Act (HIPAA) and the Health Information Technology for Economic and Clinical Health Act (HITECH) (Kaplan et al., 2023).

## Data Governance Committee

We will establish a Data Governance Committee consisting of key stakeholders from IT, legal, compliance, and business units. This committee with oversee all data-related policies and decisions, as well as be available as a data breach response team in the case of a data security incident.

## Data Inventory and Classification

Data will be inventoried and classified based on sensitivity and confidentiality. Specifically, we will classify data based on three parameters: confidentiality, availability, and integrity. The sensitivity rating will then be calculated based on the algorithm outlined in Carnegie Mellon’s *Guidelines for Data Classification*, allowing the data to be stored as public, private, or restricted (Markiewicz, 2011). In addition to this, a data retention policy will be developed to ensure that we only store data that is necessary for our operations.

## Data Privacy and Ethics

We will develop privacy policies that are focused on adhering to both the legal and ethical requirements of data protection. In addition to this, a code of ethics will be developed for data usage that emphasizes respect for individual privacy, and an Acceptable Use Policy (AUP) will be developed that outlines the acceptable uses of different types of data (National Forum on Education Statistics, 2020).

## Data Anonymization and Differential Privacy

Differential privacy is a mathematical framework designed to guarantee privacy when analyzing or releasing data (Wood et al., 2018). The overall goal of differential privacy is to render maximum security to datasets by reducing the chances of private record identification. This allows statistical information about the data to be publicly available, while still maintaining high levels of confidentiality (Aldeen et al., 2015). For example, in the United States, the combination of date of birth, zip code, and gender uniquely identifies 87% of the country’s population (Sweeney, 2002). It is essential that we preserve the privacy of individuals across all available data.

One way we will work towards this goal of differential privacy is by using data anonymization, specifically by using the k-anonymity method. The k-anonymity method is a commonly used de-identification algorithm that stipulates that each record in a dataset is similar to a minimum of k-1 records in terms of the potentially identifying variables. For example, if k = 5 and gender and date of birth are the potentially identifying variables, then an anonymized dataset would have at least 5 records for each combination of gender and date of birth (El Emam et al., 2009). This helps to ensure confidentiality and privacy across datasets, which is an essential part of our plan.

# Data Security Measures and Protocols

## Data Security Framework

Our data security framework will include numerous strategies to protect our data from breaches. This includes the use of encryption, both locally and in the cloud, regular vulnerability assessments, and intrusion detection systems (SecureAge, 2020). This framework will also be aligned with the National Institute of Standards and Technology (NIST) guidelines and our first line of defense against the evolving landscape of cybersecurity threats (NIST, 2018). By proactively safeguarding our data assets, we mitigate the risk of data breaches and ensure the confidentiality and integrity of patient information.

## Data Backup and Recovery

Data loss can have dire consequences in the healthcare sector by impacting both patient care and business continuity. For this reason, we will establish automated, regular data backup and disaster recovery plans. These plans will work to minimize data loss in the event of a system failure or breach. In addition to this, we will include encryption on all data sent to the cloud, as well as authentication parameters when accessing or deleting information from cloud storage (Sood, 2012).

## Internet of Medical Things (IoMT)

Another aspect of data security that must be considered is the security and privacy of the Internet of Medical Things (IoMT). The use of these devices is becoming increasingly common in the healthcare setting and includes things such as personal devices used to access, deliver, and process medical data, and body sensors for remote monitoring. Although these devices have the potential to increase efficiency and patient safety, their protection methods must be considered. We will incorporate three main types of security to ensure their protection: device security, connectivity security, and cloud security (Hatzivasilis et al., 2019). In addition to our data privacy policy, this security should ensure the privacy and protection of IoMT patient data.

## Incident Response Plan

Despite our best efforts, no organization is immune to data breaches. To address this, we will develop and regularly update an incident response plan. This plan will outline the precise steps to be taken in the event of a data breach, which will ensure a swift and effective response. A fast and efficient response to a data breach is crucial, as slow responses could result in both reputational damages and legal fines (Khan et al., 2021). Our response strategy is especially pertinent when it comes to reputation, as research has shown that not only do data breaches lead to disappointed customers and investors, but they can also result in decreased market values and cause billions of dollars in stock losses (Gordon et al., 2011). Our response plan will be based on the Federal Trade Commission’s *Data Breach Response: A Guide for Business* report and include response strategies such as quickly securing systems to prevent additional data losses, consulting a legal team, and notifying the appropriate parties to align with federal regulations (*Federal Trade Commission*, 2023).

# Data Access, Usage, and Sharing Mechanisms

## Access Controls

We will create a formalized process for requesting access to data, and enhance access controls through role-based access management (RBAC) to ensure that only authorized personnel can access specific datasets (Zhu et al., 2015). In addition to this, we will also implement multi-factor authentication (MFA) for access to personal and sensitive systems (Banyal et al., 2013).

## Data Usage Policies

An Acceptable Data Use policy will be defined and implemented, which defines the intended purposes for which data can be used. Furthermore, data usage will be monitored and audited to detect any unauthorized or suspicious activities (Hosseinzadeh et al., 2020).

## Secure Data Sharing

Secure data-sharing mechanisms will be established, ensuring that all data shared with external partners, vendors, and research institutions is completed in compliance with data privacy regulations. This includes the implementation of data anonymization techniques, where necessary (National Science and Technology Council, 2023).

## Employee Training

Training will be provided to new employees within the first week of starting the position and current employees on a yearly basis. This training will cover the company’s Acceptable Use Policy (AUP) as well as other privacy rules such as HIPAA, and how to report a data breach. This training will be interactive and include discussion, to ensure that employees are engaged and understand objectives and policies (National Forum on Education Statistics, 2020). It’s crucial to ensure that employees have a solid understanding of data privacy laws and consequences, as research suggests that 80% of data leaks are due to staff (Gregory, 2011).

# Conclusion

In conclusion, the strategic plan presented in this document for enhancing data governance and security at MaineHealth underscores our commitment to safeguarding patient data and ensuring its quality, all while adhering to regulatory requirements. Our collaborative approach involving key stakeholders through the creation of a Data Governance Committee allows for a shared responsibility in data stewardship. Furthermore, our comprehensive data governance policies and data security framework reinforce our dedication to patient privacy and security against evolving cyber threats. Ultimately, MaineHealth remains committed to an ongoing and adaptive process, ensuring the security and ethical handling of data to serve our patients and communities effectively.

References:

Abouelmehdi, K., Beni-Hssane, A., Khaloufi, H., & Saadi, M. (2017). Big Data Security and privacy in Healthcare: A Review. *Procedia Computer Science*, *113*, 73–80. https://doi.org/10.1016/j.procs.2017.08.292

Abraham, R., Schneider, J., & vom Brocke, J. (2019). Data governance: A conceptual framework, Structured Review, and research agenda. *International Journal of Information Management*, *49*, 424–438. https://doi.org/10.1016/j.ijinfomgt.2019.07.008

Aldeen, Y. A., Salleh, M., & Razzaque, M. A. (2015). A comprehensive review on Privacy Preserving Data Mining. *SpringerPlus*, *4*(1), 1–36. https://doi.org/10.1186/s40064-015-1481-x

Alder, S. (2023, January 24). *2022 Healthcare Data Breach Report*. HIPAA Journal. https://www.hipaajournal.com/2022-healthcare-data-breach-report/

Banyal, R. K., Jain, P., & Jain, V. K. (2013). Multi-factor Authentication Framework for cloud computing. *2013 Fifth International Conference on Computational Intelligence, Modelling and Simulation*. https://doi.org/10.1109/cimsim.2013.25

*Cost of a data breach 2023*. IBM. (2023). https://www.ibm.com/reports/data-breach

*Cybersecurity framework*. National Institute of Standards and Technology (NIST). (2018, April 16). https://www.nist.gov/cyberframework

*Data Breach Chronology*. Data Breach Chronology | Privacy Rights Clearinghouse. (n.d.). https://privacyrights.org/data-breaches

*Data breach response: A guide for business*. Federal Trade Commission. (2023, August 10). https://www.ftc.gov/business-guidance/resources/data-breach-response-guide-business

*Data Encryption and the NIST Framework Comprehensive and inherent approach*. SecureAge. (2020). https://www.secureage.com/secureage/pdf/Data-Encryption-NIST-Framework-Whitepaper.pdf

El Emam, K., Dankar, F. K., Issa, R., Jonker, E., Amyot, D., Cogo, E., Corriveau, J.-P., Walker, M., Chowdhury, S., Vaillancourt, R., Roffey, T., & Bottomley, J. (2009). A globally optimal K-anonymity method for the de-identification of Health Data. *Journal of the American Medical Informatics Association*, *16*(5), 670–682. https://doi.org/10.1197/jamia.m3144

Gordon, L. A., Loeb, M. P., & Zhou, L. (2011). The impact of information security breaches: Has there been a downward shift in costs? *Journal of Computer Security*, *19*(1), 33–56. https://doi.org/10.3233/jcs-2009-0398

Gregory, A. (2011). Data governance — protecting and unleashing the value of your Customer Data assets. *Journal of Direct, Data and Digital Marketing Practice*, *12*(3), 230–248. https://doi.org/10.1057/dddmp.2010.41

Hatzivasilis, G., Soultatos, O., Ioannidis, S., Verikoukis, C., Demetriou, G., & Tsatsoulis, C. (2019). Review of Security and privacy for the internet of medical things (IOMT). *2019 15th International Conference on Distributed Computing in Sensor Systems (DCOSS)*. https://doi.org/10.1109/dcoss.2019.00091

Hosseinzadeh, A., Eitel, A., & Jung, C. (2020). A systematic approach toward extracting technically enforceable policies from data usage control requirements. *Proceedings of the 6th International Conference on Information Systems Security and Privacy*. https://doi.org/10.5220/0008936003970405

Kaplan, B. A., Gilroy, S. P., DeHart, W. B., Brown, J. M., & Koffarnus, M. N. (2023). Data handling: Ethical principles, guidelines, and recommended practices. *Research Ethics in Behavior Analysis*, 191–214. https://doi.org/10.1016/b978-0-323-90969-3.00006-2

Khan, F., Kim, J. H., Mathiassen, L., & Moore, R. (2021). Data Breach Management: An integrated risk model. *Information & Management*, *58*(1), 103392. https://doi.org/10.1016/j.im.2020.103392

Markiewicz, D. (2011, September 15). *Guidelines for Data Classification*. Carnegie Mellon University. https://www.cmu.edu/iso/governance/guidelines/docs/DataClassificationGuidelines\_FINALv1.0.doc

National Science and Technology Council. (2023, March). National strategy to advance privacy-preserving data sharing and analytics. https://www.whitehouse.gov/wp-content/uploads/2023/03/National-Strategy-to-Advance-Privacy-Preserving-Data-Sharing-and-Analytics.pdf

Seh, A. H., Zarour, M., Alenezi, M., Sarkar, A. K., Agrawal, A., Kumar, R., & Ahmad Khan, R. (2020). Healthcare data breaches: Insights and implications. *Healthcare*, *8*(2), 133. https://doi.org/10.3390/healthcare8020133

Sood, S. K. (2012). A combined approach to ensure data security in cloud computing. *Journal of Network and Computer Applications*, *35*(6), 1831–1838. https://doi.org/10.1016/j.jnca.2012.07.007

SWEENEY, L. (2002). K-anonymity: A model for protecting privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, *10*(05), 557–570. https://doi.org/10.1142/s0218488502001648

U.S. Department of Education. (2020). *Forum guide to data governance - national center for education ...* National Forum on Education Statistics. https://nces.ed.gov/forum/pdf/Forum\_Guide\_to\_Data\_Governance.pdf

Wood, A., Altman, M., Bembenek, A., Bun, M., Gaboardi, M., Honaker, J., Nissim, K., O’Brien, D., Steinke, T., & Vadhan, S. (2018). Differential Privacy: A Primer for a non-technical audience. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3338027

Zhu, Y., Huang, D., Hu, C.-J., & Wang, X. (2015). From RBAC to ABAC: Constructing flexible data access control for cloud storage services. *IEEE Transactions on Services Computing*, *8*(4), 601–616. https://doi.org/10.1109/tsc.2014.2363474