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 Sood, 2012, allowing the data to be stored as public, private, or limited access (Sood, 2012). 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

## Data Backup and Recovery

Mention cloud storage issues!

## Internet of Medical Things (IoMT)

## Incident Response Plan

# 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).

# Conclusion

References: