Position Paper 2: Balancing Data Security and Usage

DSE6003

Adeline Casali

September 27, 2023

In today’s digital age, organizations face the intricate problem of balancing safeguarding data and harnessing its potential. This is particularly pertinent due to the rapid growth of data in the modern world. Companies often have incredible amounts of data available to aid in decision-making, but this access also comes with the risks of privacy and data breaches. The consequences of a data breach are tremendously high and include both monetary ramifications as well as reputation damages. Due to these repercussions, safeguarding data is crucial, but it also needs to be accessible as there is an increasing reliance on data in business operations in the modern age. This essay argues that it is possible to find the equilibrium between data security and utility, but this balance is unique to each organization and depends on several factors. These factors include industry-specific nuances, the influence of organizational culture, and the organization’s risk tolerance.

On one side of this problem is the argument for limiting access through increased data security measures. There are a multitude of reasons as to why this should be a paramount concern for organizations. Firstly, legal regulations such as the Health Insurance Portability and Accountability Act (HIPAA) and the Health Information Technology for Economic and Clinical Health Act (HITECH) must be complied with, or organizations will risk civil and criminal penalties. Secondly, the average cost of a data breach in the United States in 2023 reached a whopping $9.48 million, and in certain industries, such as healthcare, the average cost in the United States was even higher at $10.93 million (IBM, 2023). However, money is not the only thing at stake when it comes to a data breach. Trust, reputation, and consumer confidence are all impacted by a data breach, and studies have shown that individuals are more likely to engage with organizations and share their data when they trust that their information will be secure (Morey et al., 2015).

And finally, companies have an ethical requirement to protect the data they work with, particularly if it contains sensitive personal data. Robust data security measures protect individual privacy by ensuring that personal information remains confidential and is not misused or exposed without consent. However, as information is removed and security measures are put into place, the utility of data is reduced (Mivule & Turner, 2013). For example, implantable medical devices (IMDs) are often equipped with wireless interfaces for easy access for timely medical interventions. However, it is not feasible to use typical security measures such as passwords and certificates on these devices due to frequent emergencies in medical settings, making them susceptible to eavesdropping and unauthorized access and control (Rostami et al., 2013). In this case, although a decrease in security measures increases the device's utility, it leaves the device vulnerable to remote attacks that could severely harm patients' privacy and health.

On the other hand, is the case for broader access to data. Data has incredible power within an organization, and data insights can provide a strong competitive advantage. Not only do they aid in data-driven decision-making, but they also provide information on market trends and customer understanding. In addition, the use of a data-driven approach in an organization fosters both innovation and research by allowing for personalized solutions to problems as well as informing research questions and evidence-based policy (Marshall et al., 2015). Furthermore, data-driven industries contribute significantly to economic growth. One study found that companies that utilize data-driven decision-making have 5-6% higher productivity and output growth than those that do not value a data-driven approach (Brynjolfsson et al., 2011). Another study from IBM in 2014 found that organizations are 36% more likely to beat their competitors in revenue growth and operational efficiency when they use big data and analytics within their innovation processes (IBM, 2014).

A critical example of how data access is necessary is seen in the response to the COVID-19 pandemic in 2020. Shared data was a critical aspect of our work to contain the virus, whether that be by access to data for vaccine development or for contact tracing from a public health perspective. Contact tracing is a commonly used public health tool to manage the spread of infectious diseases by determining people exposed to a disease through interaction with an infected individual. A contact tracer then notifies the at-risk individual of their risk of exposure and provides them with guidelines to help reduce the spread of the disease even further (CDC, 2020). However, although contact tracing has been designed to protect privacy while benefiting the public, concerns have been raised regarding privacy, legality, and equity issues due to the large volume of contact tracing carried out as part of the pandemic response in the United States (Wacksman, 2021). It is critical for public health agencies like the Centers for Disease Control (CDC) to remain adaptable and balance the goals of rigorous data protection and utilization to achieve public health goals.

Consequently, balancing the two aspects of data utility and privacy is a critical problem modern organizations face. While over-restriction can cause frustration and hinder innovation and growth, over-permissiveness carries the risks of vulnerabilities to data breaches and legal and ethical challenges. Finding the middle ground between the two aspects requires various strategies combined with an evaluation of company values and risk tolerances. One strategy that can be used to find this balance is establishing clear data usage policies and data governance frameworks. These frameworks provide the structure for responsible data management. Another important tactic is utilizing collaborative decision-making between organizational infrastructure managers and functional managers to devise strategies that optimize both security and usage. This equilibrium will likely look different depending on the organization’s industry. For example, healthcare organizations work with highly private and sensitive information, which carries a significant risk of damage in the case of a data breach. This will require a prioritization of privacy and security when determining data policies and frameworks. Other industries, such as marketing, may value company-wide access to their data over intense security, as their risk tolerance for a breach is much higher. Overall, each organization needs to determine its own company values and risk tolerances when establishing its internal policies and security measures.

Additionally, contemporary approaches such as the k-anonymity method and differential privacy can be used as tools for businesses to balance data privacy and utility. This is particularly relevant in situations where they need to share aggregate or statistical information while preserving the privacy of individual contributors. Both k-anonymity and differential privacy provide a mathematical framework and techniques to add noise or perturbation to data in a controlled manner, making it difficult to reverse-engineer or identify the individual data points while still allowing for meaningful analysis (Wood et al., 2018; El Emam et al., 2009). These privacy-preserving analytical methods are just a few examples of modern techniques available that can be used to approach the complex problem of evaluating and balancing data privacy. However, even with the latest data privacy algorithms, such as differential privacy, there is a significant loss of data utility that comes with increased confidentiality (Mivule & Turner, 2013).

In conclusion, although it will always be challenging, finding a balance that achieves both data security and access objectives is possible. Balancing data privacy and usage is not a one-size-fits-all approach, and finding the balance is a multifaceted problem that varies across organizations and requires building a culture of data responsibility. Industry-specific considerations profoundly influence the feasibility of striking this equilibrium, the prevailing organizational culture, and an organization's risk tolerance. While the task is complex, it is manageable. Additionally, by valuing data as a strategic asset and leveraging technology for security, companies can better adapt and improve their methodologies. The technology available to address this challenge is constantly evolving, and companies must evolve their own methods in tandem to keep up. For example, technological assets such as machine learning and artificial intelligence (AI) are being increasingly utilized. Machine learning algorithms can be used to check and optimize the balance between data anonymity and data utility (Esmeel et al., 2020). In addition, organizations that utilize AI and automation in security measures save an average of $1.76 million compared to companies that do not (IBM, 2023). This illustrates the evolving nature of data management in the digital age. Companies need to be agile in adapting to both evolving threats and opportunities to ensure that a balance is maintained. Overall, organizations must recognize the uniqueness of their context, align their data practices accordingly, and safeguard their data in an increasingly data-driven world. The equilibrium between data security and usage is intrinsically nuanced and context-dependent, reflecting the ever-evolving nature of data management.

References:

Brynjolfsson, E., Hitt, L. M., & Kim, H. H. (2011). Strength in numbers: How does data-driven decisionmaking affect firm performance? *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1819486

CDC. (2020). *Contact Tracing*. Centers for Disease Control and Prevention. https://www.cdc.gov/museum/pdf/cdcm-pha-stem-lesson-contact-tracing-lesson.pdf

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

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

Esmeel, T. K., Hasan, M. M., Kabir, M. N., & Firdaus, A. (2020). Balancing data utility versus information loss in data-privacy protection using K-anonymity. *2020 IEEE 8th Conference on Systems, Process and Control (ICSPC)*. https://doi.org/10.1109/icspc50992.2020.9305776

*Innovative analytics*. IBM. (2014). https://www.ibm.com/thought-leadership/institute-business-value/en-us/report/innovateanalytics

Marshall, A., Mueck, S., & Shockley, R. (2015). How leading organizations use big data and analytics to innovate. *Strategy & Leadership*, *43*(5), 32–39. https://doi.org/10.1108/sl-06-2015-0054

Mivule, K., & Turner, C. (2013). A comparative analysis of data privacy and utility parameter adjustment, using machine learning classification as a gauge. *Procedia Computer Science*, *20*, 414–419. https://doi.org/10.1016/j.procs.2013.09.295

Morey, T., Forbath, T., & Schoop, A. (2015, May 1). *Customer data: Designing for transparency and trust*. Harvard Business Review. https://hbr.org/2015/05/customer-data-designing-for-transparency-and-trust

Rostami, M., Burleson, W., Koushanfar, F., & Juels, A. (2013). Balancing security and utility in medical devices? *Proceedings of the 50th Annual Design Automation Conference*. https://doi.org/10.1145/2463209.2488750

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