**Data Function Areas - Ethics**

Chase Rawson

Department of Technology, Granite State College

DATA 520: Data Mining, Cleaning, and Visualization

Dr. Tony Sulpizio

January 11, 2023

Throughout the world, data collecting has been a priority for almost every business and organization. Each entity has changed its decision-making by using data collected from its clientele to make crucial decisions. These decisions change the entity’s operations for better or for worse. With the role of technological innovation, these entities have collected data and made decisions in easier ways. However, these decisions, and the ways the data was collected, bring ethical problems affecting their clients. Data ethics “encompasses the moral obligations of gathering, protecting, and using personally identifiable information and how it affects individuals.” (Cote, 2021) It is important to ensure data privacy and ethical uses of data, with never-ending consequences ranging from social security information to financial information being leaked, violating federal and state laws. Many different principles must be followed to ensure data ethics.

One of the essential principles of data ethics relates to privacy. When a consumer fills out a form online, there is an excellent chance that personally identifiable information will be collected, such as your name, birthdate, address, Social Security card, and financial information. Organizations such as healthcare companies and banks have millions of names and birthdates saved on servers to protect themselves against hackers. These two industries must follow specific laws to protect against sensitive data being released to the public, even if the data owners, the consumers, give them consent to collect and analyze their information. The Payment Card Industry Data Security Standard, also known as PCI DSS, enforces data ethics against companies possessing payment card information. As described by SUNY Plattsburgh, “The PCI DSS Compliance Committee is responsible for developing, implementing, monitoring, and evaluating a comprehensive and coherent PCI DSS compliance program and coordinating overall institutional efforts.” (“PCI DSS compliance”, n.d.). Secure data architecture and file encryption have been used to prevent attacks that occurred on purpose or accidentally. Purposefully or accidentally. Data privacy concerns many consumers, so they need to understand how their data is used.

Transparency when handling data ensures that data analysts use consumer data for any purpose that the consumer approves. When collecting data, it is essential to notify them of the purpose, such as how you plan to collect their data, how you plan to store it, and how you plan to use it. To ensure these morals are followed, government agencies and organizations, such as OECD and the FTC, have issued guidelines that detail the data-collecting process. FTC’s guidelines are based on privacy by design, in which they recommend that organization design and implement their own data privacy rules that must follow each of the five principles, Notice/Awareness, Choice/Consent, Access/Participation, Integrity/Security, and Enforcement/Redress (Henderson et al., 2017). Say, for example, a company wants to use the data collected to personalize their client’s website experience; their rules of notifying them of data collected must be based on the FTC principle of Choice, “consumers must be given options with respect to whether and how personal information collected from them may be used for purposes beyond those for which the information was provided.” (Henderson et al., 2017). Violating any of these principles, depending on the type of data a company may possess, would most likely result in financial punishment from multiple governments depending on the nationalities of the consumers. Moreover, most times, the outcomes of data analysis and data leaks result in far more damage to the consumer, questioning whether there was any purpose to collecting data in the first place.

Before collecting data, it is important to determine the purpose by asking, “why you need it, what you’ll gain from it, and what changes you’ll be able to make after analysis. If your intention is to hurt others, profit from your subjects’ weaknesses, or any other malicious goal, it’s not ethical to collect their data.” (Cote, 2021). When data collecting, misrepresenting or underrepresenting the sample of data collected brings forth the role of bias. Having bias when data is collected, created, or selected and not resolving these issues before presenting the findings can also be malicious. As a data analyst, it is ethical to “provide evidence that any products they have developed do not exhibit bias or potential harm against any demographic subgroups such as race, gender or ethnicity or subgroups defined by genetic markers or socio-economic status.” (“Our guiding ethical principles”, n.d.). For example, a data analyst wants to collect information about one’s health. The purpose of collecting health information is vague; it would not be appropriate to ask for specific, sensitive information like one’s mental health when asking about their experience at a hospital. With these three principles in use, collecting personal information may seem challenging, but it can improve our lives.

One example of these three principles in use comes from Coca-Cola. Their privacy policy statement states, “We, The Coca-Cola Company and its affiliates, take your privacy seriously. Our mission is to refresh the world and we can’t do that without your trust. This Privacy Policy explains what information we collect about you, why we collect the information, as well as how we collect and use the information so that we can deliver the best, most refreshing experience that we can.” (Coca-Cola®, n.d.). Coca Cola’s policy statement shows transparency by acknowledging they collect their user’s data. The company shows privacy by, “standard physical, technical and administrative measures designed to reduce the risk of loss, misuse, unauthorized access, disclosure or modification of your Personal Information.” (Coca-Cola®, n.d.). They also list their intentions to use collected data for business and essential purposes to comply with all laws and regulations.

The rules around data ethics have given more purpose to collecting data, storing data, and preventing personally identifiable information of millions of people from getting leaked. Enforcing rules that promote transparency, privacy, and the purpose of collecting data helps our economy and promotes a just world. With data ethics, the handling of data has protected millions in our world and eliminated the possibility of organizations with damaged reputations.

**References**

*Coca-Cola®*. Discover Real Magic. (n.d.). Retrieved January 12, 2023, from <https://us.coca-cola.com/privacy-policy>

Cote, C. (2021, March 16). *5 principles of data ethics for business*. Business Insights Blog. Retrieved January 12, 2023, from <https://online.hbs.edu/blog/post/data-ethics>

Henderson, D., Earley, S., Sebastian-Coleman, L., Sykora, E., & Smith, E. (2017). *Dama-Dmbok: Data Management Body of Knowledge* (2nd ed.). Technics Publications.

*Our guiding ethical principles*. School of Data Science and Analytics. (n.d.). Retrieved January 12, 2023, from <https://datascience.kennesaw.edu/about/ethics.php>

*PCI DSS compliance*. State University of New York (SUNY) College at Plattsburgh. (n.d.). Retrieved January 12, 2023, from <https://www.plattsburgh.edu/about/offices-divisions/administration-finance/administrative-services/legal-compliance/pci-dss/>