Prospectus

University of Virginia Facilities Management Website
(Technical Topic)

Facebook Data Breach of 2018
(STS Topic)

By

Shivani Surti

November 1, 2021

Technical Project Team Members: Dwij Gandhi and Jane Weng

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.

Signed: _____Shivani Surti_____

Approved: _____ Date _____
Benjamin Laugelli, Department of Engineering and Society

Approved: _____ Date _____
Paul McBurney, Department of Computer Science

Introduction

The adoption of technology into businesses and companies has rapidly increased, leading to improvements in productivity and performance and a reduction in human errors. In fact, in 2018, 41.4% of companies have reported having a dedicated digital transformation team to help them adopt and implement new technologies within the organization (Newman & Blanchard, 2018). With the rapidly changing work atmosphere, it is imperative that companies keep pace with the times. The move from paper to online databases promotes better data storage and efficiency in retrieving that data. Currently, the University of Virginia's recycling department maintains its data storage through paper forms that get translated over to Excel sheets. The data collection to storage pipeline is a slow and manual process that takes up time that could be used to solve other business issues. This time is crucial considering the department's ambitious goals for the future. My team and I propose to help the Recycling department by creating a website that would allow the department to keep track of all the waste the department has collected. This website would help the Recycling Department focus more on reaching their mission instead of putting hours into organizing the data the department collects. The visualization of data my team and I hope to create as part of the website would help the department quickly view and analyze the data to help make quicker judgments on whether the department is reaching its goals.

However, building a functioning website and database for the Recycling Department will not suffice; other factors must also be considered. As technology becomes a part of our everyday life, it has become just as essential to ensure our data is secure**.** Breaches in the database or website could have negative consequences for the department. For example, an outsider could manipulate the data or steal information about the users of the website. Many actors/factors must be considered to make the technical project successful; such include securing the identity of the

website users, the users of the website and organization, and the number of funds available. Each of these factors would influence the design and outcome of the project. If precautions are not taken, a data breach like the one Facebook suffered in 2018 could happen again. The Facebook Data Breach was a result of a failed network that compromised the integrity of the organization.

My team and I are building a heterogeneous network in our technical project, so looking at the failure of a related socio-technical network would be constructive. By not acknowledging both the technical and social aspects of the design, the reader fails to comprehend the network involved in creating a successful technical product. By focusing only on the technical aspect of the design, the technical choices behind a product will fail to be considered. By focusing solely on the social aspect, a person would fail to understand what compromises were made to get a middle ground between expectations and the reality of technological practices.

The broader goals set out in my research question are socio-technical in nature and, as such, require proposals that address both technical and social aspects. By applying design and database concepts to my technical project, I will address the issue of developing technical specifications for the Recycling Department's website that would help it improve productivity and work efficiency. By examining the Facebook Data breach of 2018, I will understand what human and non-human factors influenced the failure of Facebook's network.

Technical Report

University of Virginia's recycling department is responsible for and helps facilitate waste and recycling management. It's goal is to reduce landfill waste and give these waste items a second life. The recycling division collects aluminum (and other metals), cardboard, electronics, glass, plastics and film plastics, and white and mixed office paper (Recycling, 2021). These waste items get sent to local recycling programs. The University's goal for 2030 is to minimize its

carbon footprint, which means promoting and accurately measuring the waste collected (Recycling Services, 2021). The department currently uses Excel sheets and paper forms to keep track of the volumes of waste collected. Waste collection drivers pick up this waste and report how many bags of waste they have collected on paper forms. These papers get passed to supervisors who translate them to Excel sheets.

With the University's ambitious goals, it is even more crucial to have well-documented information on how much waste is collected overall and from each facility. There are many facilities at the University and manually calculating the volume of waste for each of those facilities is a tedious task that requires a lot of human resources and attention to detail. Paper forms and the lack of a trash bag-to-volume automated conversion tool result in human error due to the inconsistency in calculating the volume of waste. There is a significant probability of inconsistent data throughout the pipeline (from recording the data to translating it). One mistake in the forms by a driver could jeopardize the accuracy of the volume output. A lack of structure in the Excel sheets makes it harder to get a clear snapshot of how much waste has been collected from each facility. Currently, the Excel sheets are structured as one book with different sheets according to pick up date, type of waste, and different facilities. This format becomes rather inconvenient and confusing when trying to interpret the data.

This technical project aims to develop a modern organizational tool that can be used by drivers and management to maintain accurate data storage. The objective is for people to easily access information about each facility without the hassle of looking through Excel sheets. Also, moving driver forms from paper to digital would help organize the documents and allow information to be quickly available. To achieve this goal, my team and I propose creating a database and website for the recycling department that would allow it to move from paper forms

and Excel sheets to digitalized resources and management systems making maintaining information easier.

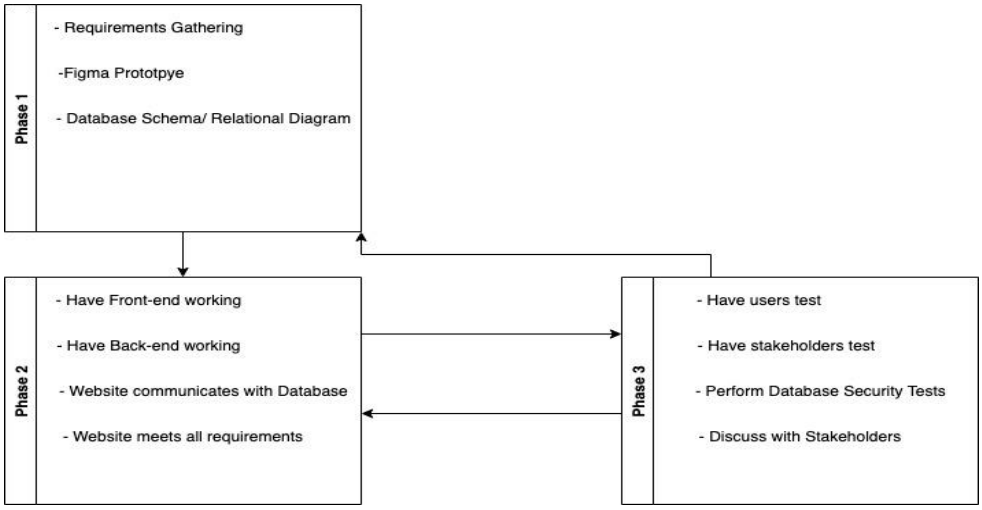

Figure 1: Diagram representing the Phases of the project.

This project will be broken down into phases for improved organization and management according to the software development lifecycle reflected in Figure 1. A software development lifecycle is an organized process for creating a piece of software from getting requirements, to the finished product; it includes continuous testing and implementation. Phase one of this technical project will be to make a Figma prototype and database diagrams based on meetings and discussions with the potential users of the website, along with the management responsible for maintaining the website. Database schemas will be created based on the fundamental principles of database design. A schema is the building block of the database that defines which tables will be created and what attributes they will have. Principles of database design include concepts such as establishing primary keys (an attribute of a table that will be its primary identifier), foreign keys (a unique attribute that has been established in another table) and creating relationships between tables (relationships help bring information in each of the tables

together in a meaningful way). A table in a database is a list of rows and columns that store information based on the defined schema (Database design basics, 2021).

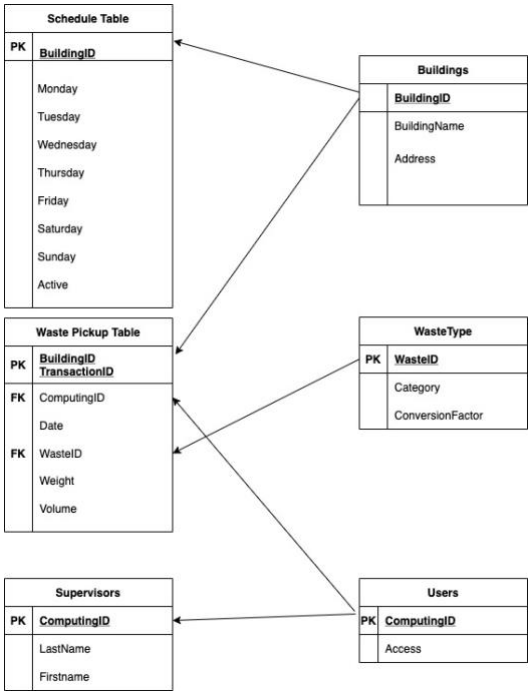

Figure 2: Relational Diagram of Potential Database Structure

The Figma prototype will be presented to the website users before development to make sure the design matches their expectations in terms of features wanted and usability of the design. The database structure must support all the information currently stored in the Excel sheets. Therefore, it needs careful evaluation and approval from stakeholders. A potential structure of the database design can be seen in Figure 2. The box represents the tables in the database, and the arrows represent how the tables are connected.

Phase two will be coding the website design and connecting the database to the website. In phase two, the goal is to also populate the database with actual information that will help with testing the website and database. The database will be created using Microsoft Azure's SQL Server, and the website will be created using .NET Core's Model View Controller (MVC) framework. Microsoft Azure's SQL Server is a cloud-based database service that is scalable and intelligent. In simpler terms, the database that I will be making for the Recycling Department will be hosted on Azure's server (Azure SQL Database, 2021). .NET Core's MVC framework is an architectural pattern that navigates user requests in an ordered manner and is implemented using code (Overview of ASP.NET Core MVC, 2021). The reasoning behind choosing these specific technologies is that the University already uses Azure services and .NET Core

frameworks in their current applications. By using these technologies, the technical project will be consistent and easily integrate with their current codebase. Each phase will be repeated until the product passes all testing requirements and satisfies the stakeholder's demands. This analysis will help determine if the website meets users' needs and if further improvements are needed.

STS Research Topic

Facebook experienced a data breach in 2018 that compromised the data of millions of users. Aleksandr Kogan, the developer of a personality quiz featured on Facebook, called "This Is Your Digital Life" sponsored by Cambridge Analytica, had collected information on millions of Facebook users and their friends. The personality quiz would collect personal information (e.g., name, birthday) and Facebook activity, such as which pages a user follows and what posts they like. About 87 million users' information was passed to Cambridge Analytica by Aleksandr Kogan. Cambridge Analytica used this information to promote Donald Trump in the presidential elections by exploiting a Facebook feature called "dark post." A "dark post" is an ad explicitly tailored for the user. Using this information, Cambridge Analytica made personalized ads that would emotionally compel people to vote for Donald Trump (ur Rehman, 2019). By making ads that people could relate with or promising policies people would want to see in the future, those people would be likely to vote for Donald Trump in the Presidential Election. This activity only came to light when a former researcher at Cambridge Analytica named Christopher Wylie exposed the activity to journalists. Articles about the data breach were first released by The Guardian (Curtis, 2018).

For this data breach, Facebook's negligence is often blamed. In the early stages of Facebook, app developers were allowed to collect information about Facebook users and their friends (Tarran, 2018). These permissions were given to help support and encourage app

developers; however, they proved detrimental. Many would argue that Facebook should have been stricter as millions of users use Facebook daily, meaning a lot of information about these users can be collected. The leakage of that data violates the user's privacy and personal information rights (Bogle, 2018). While it is true that Facebook should have been more careful, other factors involved, in addition to Facebook's negligence, contributed to the data breach. If we continue to think that Facebook's negligence was responsible for the data breach, we will not understand the role that these other factors played alongside Facebook in the project's failure. By focusing narrowly on the lack of attention from Facebook, we fail to account for the mounting pressure of the election. Also, it underscores the contribution of the actors involved: Aleksandr Kogan, Cambridge Analytica, and Christopher Wylie.

Drawing on the Actor-Network Theory, I argue that Facebook's negligence in conjunction with election pressures and conflict between Aleksandr Kogan and Facebook resulted in the data breach. These actors were especially problematic to the network as they were the actors that used private user information to influence the elections. Actor-network theory is described as "a combination of social and technical engineering in an environment filled with indifferent or overtly hostile physical and social actors" (Law, 1987)**.** Actor-network Theory is the interplay of social and technical actors that contribute to the project. Power in this theory is defined by how strong the bonds between the actors are in a network. The stronger the network, the more power that network has. The process of forming and maintaining these actor-networks is called translation (Cressman, 2009). Heterogenous engineering is designing technology that would impact and transform the social and technical world (Callon, 1987). I will use these principles to analyze the Facebook data breach of 2018. To support my argument, I will analyze

evidence from exclusive interviews given by Facebook, Cambridge Analytica, and Christopher Wylie, public reports released by Facebook, and public reports released by Cambridge Analytica.

Conclusion

The deliverable for the technical problem described in this paper will be creating a .NET Core Website with the integration of an Azure SQL Server for database storage. The STS paper will be focused on understanding how and why the Facebook Data Breach of 2018 occurred. The actor-network theory will be applied to identify the actors involved and understand their contributions to the failure of Facebook's network that resulted in a data breach. The combined results of this technical report will serve to address the issue of technological adaptation and data security issues from a socio-technical perspective. This analysis will help guide and improve the technical problem I am trying to address with the University's recycling department.

Word Count: 2174

Bibliography

Azure SQL Database. (2021). Retrieved from Microsoft Azure: https://azure.microsoft.com/en-us/products/azure-sql/database/#overview

Bogle, A. (2018, March 26). Facebook after Cambridge Analytica: Is this the beginning of the end?: Facebook is not impenetrable. It has to face pressure from regulators, advertisers and users ready to log off. *ABC Science Online; Ultimo*.

Callon, M. (1987). Society in the making: The study of technology as a tool for sociological analysi. In M. Callon, *The social construction of technological systems: New directions in the sociology and history of technology* (pp. 83-103). London: MIT Press.

Cressman, D. (2009). *A Brief Overview of Actor-Network Theory: Punctualization, Heterogeneous Engineering & Translation.* Simon Fraser University.

Curtis, J. (2018). Cambridge Analytica: Millions of Facebook profiles 'mined illegally'. *IT Pro; London*.

Database design basics. (2021). Retrieved from Microsoft: https://support.microsoft.com/en-us/office/database-design-basics-eb2159cf-1e30-401a-8084-bd4f9c9ca1f5

Law, J. (1987, April). On the Social Explanation of Technical Change: The Case of the Portuguese Maritime Expansion. *Technology and Culture, 28*(2), 235.

Newman, D., & Blanchard, O. (2018). *2018 Digital Transformation Index.* Futurum.

Overview of ASP.NET Core MVC. (2021). Retrieved from Microsoft Docs: https://docs.microsoft.com/en-us/aspnet/core/mvc/overview?view=aspnetcore-5.0

Recycling. (2021). Retrieved from UVA Sustainability: https://sustainability.virginia.edu/programs/recycling

Recycling Services. (2021). Retrieved from Facilities Management: https://www.fm.virginia.edu/services/recycling.html

Tarran, B. (2018, June). What can we learn from the Facebook—Cambridge Analytica scandal? *Significance; London, 15*(3), 4-5.

ur Rehman, I. (2019). Facebook-Cambridge Analytica data harvesting: What you need to know. *Library Philosophy and Practice*, 1-11.