Prospectus

Tap2Change – Cashless Transactions with Panhandlers
(Technical Topic)

Consumer Data Collection
(STS Topic)

Presented to the Faculty of the School of Engineering and Applied Science

University of Virginia • Charlottesville, Virginia

In Fulfillment of the Requirements for the Degree

Bachelor of Science, School of Engineering

Rohit Chhatre
Fall 2020
Department of Computer Science

Technical Project Members:
Jeffrey Rhoads, Harish Chandrasekaran, Caden Moses, Cutter Grathwohl

On my honor as a student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Signed: Rohit Chhatre _____

Approved: _____ Date: _____
Prof. S. Travis Elliott, Department of Engineering and Society

Approved: Yuan Tian _____ Date: 12/09/2020
Prof. Yuan Tian, Department of Computer Science

Introduction

Nowadays, so many people have smartphones and other technology devices where they can provide their personal information to these different platforms. Although some data is encrypted and kept secure, like social security numbers and exact credit card numbers, other information is leaked through to the companies and they have access to all of our personal data. My technical project focuses on creating an application to promote cashless transactions between donors and panhandlers. Through figuring out goals for my app, one was to gather information on how much people donate to those causes. This made me think about consumer data collection in general which led me to choose it as my topic for my STS research. A lot of people and the media harp on all the negatives of collecting people's data through app usage, but I believe there are positives that people do not consider. Using the Social Construction of Technology (SCOT) framework, I will analyze the topic of data collection, data analysis, and the selling of data. In choosing the SCOT framework it is important to consider its differences with its counterpart, technological determinism. This framework suggests that technology is the driving factor in shaping society. On contrast, SCOT argues that technological innovation occurs as a response to different societal forces and is shaped by human actions. The main factors of this framework are relevant social groups, flexibility of interpretation, and the technological frame (YOUSEFIKHAH 2017). I will show that SCOT is the best methodology to analyze the collection of information from consumers. By analyzing the different stakeholders involved and by showing that there are positives within all the scrutiny, I will show that perhaps spreading greater knowledge about this technology could foster greater acceptance and could be beneficial for all parties.

Technical Topic – Tap2Change

With the assistance of computer science professor Yuan Tian, economics major Caden Moses, biomedical engineering major Cutter Grathwohl, and computer science majors Jeffrey Rhoads, Harish Chandrasekaran, and myself, my capstone project seeks to help empower panhandlers in an increasingly cashless world by developing an application which will connect donors with homeless individuals. When people are walking down the street or are in their car and encounter people who are asking for money, they are faced with a few dilemmas. In some cases, people just don't donate to homeless people. But if we look at cases where someone is willing to donate, they might not end up giving money for a few reasons. Firstly, they may not have cash, as a lot of people tend not to carry cash with the increased usage of card payments and other cashless transaction methods. Secondly, people may want to donate but they are worried about what their money is going towards. Their money may be spent on satisfying harmful habits alcohol and drug usage, although there was a study conducted in British Colombia which showed that homeless people tend to actually spend money on useful items (Giuliani-Hoffman 2020). The technical project I am going to pursue will focus on fixing these problems with panhandling in the hopes of increasing the amount of revenue these people can generate. Another goal of my project is to collect data on how much people donate to these causes, because currently there is no good way of measuring the amount people actually donate. This is due to the nature of how these transactions are mainly small cash amounts. By developing this application, we will have basically created a platform in which we can track how much money people give to panhandlers. This eventually led me to focus on the topic of consumer data collection for my STS research paper, because I realized the amount of information that all of these technology companies have on all of its consumers.

Not only are panhandlers spending money on unnecessary items that potentially propagate the cycle of homelessness, but the rise of e-commerce and cashless transactions, like I mentioned, is part of the reason that people carry less cash. According to a survey conducted by US Bank in 2017, "50 percent of respondents reported carrying cash less than half of the time" ("Digital Payment Platforms Primed to Topple Cash", 2017, para. 4) . Another finding from the same survey was that when people do carry cash, 76% carry less than $50. Income and age also play a major factor in determining whether people carry cash or utilize cashless options. The Pew Research Center surveyed over 13,000 adults in 2018 and found that "adults with an annual household income of over $75,000 were more than twice as likely as those making less than $30,000 to say they do not make any purchases using cash in a typical week" (Dickler, 2019, para. 6) . People who are more capable of making donations to the less fortunate are using cash at a decreasing rate, which leads to less cash being donated. As for the age disparity, Experian reported that "more than 1 in 10 millennials use their digital wallet for every purchase", although there are other trends that suggest older generations are slowly adopting the move to cashless as well (Dickler, 2019, para. 4) . With all this being said, there are also signs that the move to cashless will not necessarily have much of an impact on panhandlers' income. Two professors, Gwendolyn Dordick and Brendan O'Flaherty, conducted a study of panhandlers in New York City and concluded that most people will continue to carry small amounts of cash, which is exactly what panhandlers need. Dordick and O'Flaherty then go on to claim that panhandlers should try to collect cashless donations because "the transfer of electronic funds could benefit both panhandlers and potential donors if having an app [could] provide some kind of credential [which] would signal to potential donors that this panhandler was not 'scamming' them" (Kiger, 2017, para. 8) . Even though both professors claim that cash will not entirely disappear from

people's pockets in the near future, they agree that having an online platform would be beneficial for both potential donors along with panhandlers.

My project aims to create an application where people can donate to the homeless without having to use cash. For panhandlers that don't have smartphones to receive transactions, they can use a QR code/card reader so that people can donate money to that person's "account". This aspect takes care of the issue when people who are willing to donate don't have cash available. When they use this application, it will be the same effect of using Venmo or Apple Pay. The second aspect of the project is to make sure these donations go to the right cause. We are seeking to partner with different entities that panhandlers currently use to get their goods. Some of these include grocery stores, gas stations, clothing stores, pharmacies, convenience stores, homeless shelters, educational and vocational training institutions. In talking with a coordinator at the Haven, a low-barrier day shelter and social resource center based in Charlottesville, we also learned that our platform can connect with existing infrastructures like smartphone drop-off's and that our platform also should help balance overdrawn accounts that these people usually have (R. White, personal communication, November 22, 2020). These places can then offer whatever essential items they have. In most cases, this would include buying most things except alcohol or drugs that are not for medicinal purposes. When building this application, it is important to consider different factors that would make this similar to a regular cash donation. This means it should include anonymity, security, and ease of use. Panhandlers or donors may not want their name or information shared between each other when making the transaction, so keeping everything anonymous could be important. The transactions between the individuals should be secure as well so personal data like bank account information isn't compromised. In order for this application to be reasonable to use and for it to be a valid

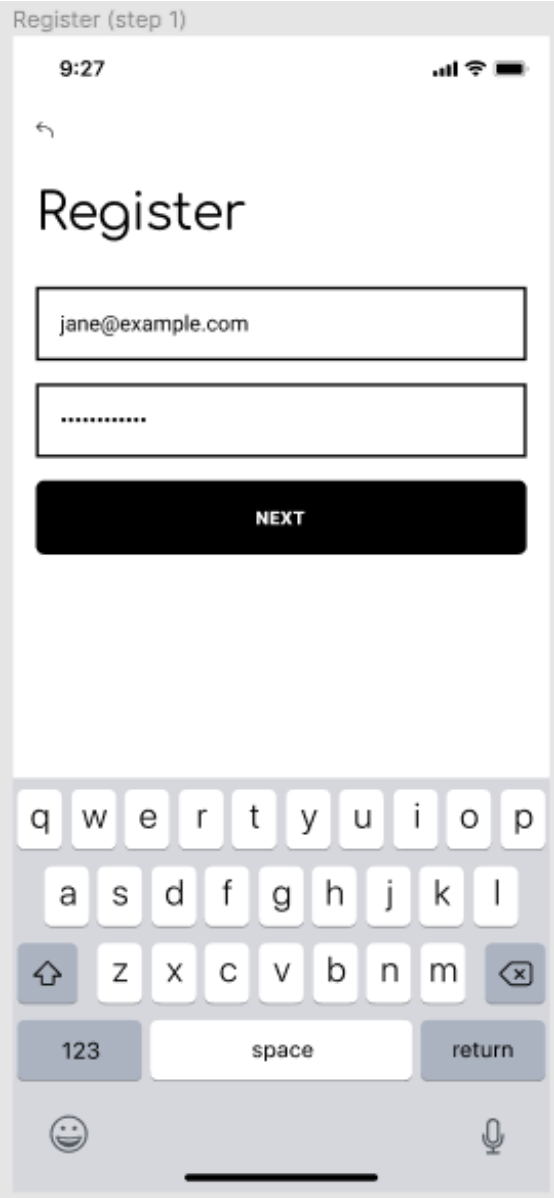
replacement for the current donation style, it is important for transactions to be super quick and easy, similarly to a simple cash donation. I am pursuing this project in hopes of providing a helpful service to people in need.

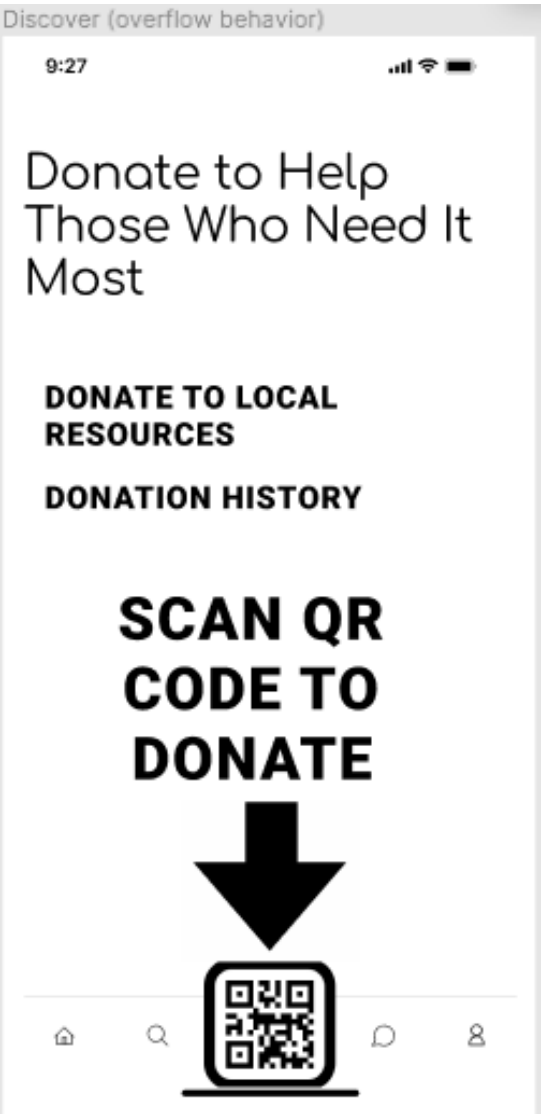
From a technological standpoint, we plan to implement many features to allow for the cashless transaction considering all the critical requirements we have laid out. We will have multiple Single Sign-On (SSO) features including Google, Facebook, and PayPal so we can appeal to the widest audience. In order to conduct our transactions, we will use PayPal and Venmo API along with OpenID Connect (Y. Tian, personal communication, November 5, 2020) to ensure maximum security. For our Minimum Viable Product, we plan to build a web application with a clean, easy-to-follow mobile view which will be deployed on Heroku. If time permits, our ideal option (based on the decision matrix) would be to create a web and mobile interface because this would allow for the most people to be able to use the application. With an increase in capital, we could transition to deploying to a cloud platform like AWS in order to maximize overall availability. My contribution to this project will be mainly from a technological standpoint. I will focus on implementing a little of the frontend and backend of the application, as I have some experience in both.

In order to promote the empowerment of the homeless and encourage more people to donate to their causes, we decided to include the limitation of goods and services that can be bought using the funds received from this application. This will encourage receivers to spend on useful items and will give a vote of confidence in terms of the giver to donate. We also have found that we should partner with restaurants so that they can offer discounts for receivers in order to show that the community cares about the population and wants to offer their services where possible. If we plan to have a monetization strategy we must either sell the data we will

collect on receiver spending and the amount that donors are giving or take a cut from partners that are helping us provide the services for the homeless.

Another feature we hope to include are partnerships with treatment center resources to aid recovery for possible drug abuse, mental health, and other medical issues. Furthermore, we aim to use easily understandable symbols and color schemes to cater to less-educated users and adhere to WCAG standards to make it accessible disabled users as well. Some of the designs we plan to implement in our application are shown below:





STS Topic – Consumer Data Collection

Consumer data collection has been a hot topic recently, as more and more of our world is transitioning online. Everyone has the chance to post content to the web and provide information about themselves. This leads companies who manage these platforms, like social media, email, and other online applications, to have control over the personal information of so many customers. It has become a common trend for companies to track and analyze the data that comes through their sites. There has been a lot of negative backlash on these practices, like we

have seen with Facebook, but I believe that there are also a lot of positives to tracking online data. For my research, I could look at this topic through many viewpoints. SCOT tells us that we need to analyze all the different social groups that play a role in developing this technology.

From an economic standpoint, these companies that collect data make so much money from selling it to other companies to use. We have even seen products like the Amazon Alexa go down in price because it is clear that these companies aren't selling the actual product, rather they are providing a platform as a means for collecting user information. Some of these uses could be malicious but there are a lot of positive uses for the data as well. Many complex technical aspects go into actually tracking user data in an application and some of this data is actually useful, like when predicting interests. Sometimes these sites know what you want even before you, as a consumer, knows what you want. So many of my friends have bought different items off of the "recommended" list or catered advertisements that show up on their feed. Due to the many intricacies of retrieving the data, this data is filtered pretty well and is also great to feed into different Machine Learning algorithms that we use for other technologies. The three main reasons why online learning is crucial is because of its volume, velocity, and variety (Hunt 2017). There is just so much data out there online. This volume can never be replicated in terms of a survey or physically collecting the data. Also, filtered data from the application is usually better than data collected from surveys because it is most "raw". People tend to not include inherent biases in a physical survey but they can't hide these biases online. The velocity of data is also important. Online data is constantly updated due to new data always coming in. This means that our algorithms can be constantly changing to fit the new data which will be better in the long run, in terms of adapting to a changing society. The variety of online data may be the most important, though. When it comes to physical collection of data, there may be human biases

that are unaccounted for, like lack of representation of race or gender for example. This bias would be counteracted by the usage of online learning. These online learning models are the foundation of technologies like self-driving cars, online shopping, and voice recognition technology. From this standpoint, consumers could be pleased that their information is used in other technologies that could in turn benefit themselves.

So why is it such a big deal for companies to take this data? Privacy has always been a big talking point when it comes to consumer data collection. For instance, from a political standpoint, people are concerned with companies being able to "spy" due to all the tracking that happens behind the scenes. As an example, there has been a scandal with TikTok having access to location data and other inner phone workings when maybe they shouldn't be authorized to that information. They have gotten into a $650 million lawsuit after companies found out they had access to the Notes app in iOS (Allyn 2020). With access to this very personal information, it could be very harmful if it gets into the wrong hands. TikTok has been labeled as a "Chinese Spyware" company (Doffman 2020) because all of this data that they have collected has allegedly been sent to China or the Chinese Communist Party. Because of this backlash, United States has forced government employees and their families to delete the app. There has also been a push for American companies to buy the app, like Microsoft, so "data on American citizens would remain in U.S. borders" and data backed-up or outside the country would be deleted after being transferred (Allyn 2020). Some people say that the data collection fraud shown in TikTok doesn't warrant the deletion of the app because other domestic technology companies also follow these malicious practices. Facebook has faced millions of dollars in lawsuits because of "unethical practices" (Solon & Farivar 2019) in terms of recording and selling data. They had a huge data breach in which 50 million accounts got leaked, which equated to almost a third of the

users in North America (Graham-Harrison & Cadwalladr 2018). For such a big company with information on so many people to leak all this data is definitely alarming. In these cases, it is valid for people to have concern for their privacy and companies should also be kept in check on how the data they collect is stored and managed.

Another factor of SCOT that can be analyzed with this topic is interpretive flexibility. This states people can appropriate technologies, or artifacts, differently in order to find the most uses for it (YOUSEFIKHAH 2017). In this case, there are so many different possible uses for data collection and data analysis. Today, there are many things we take for granted which were built on preexisting data sets. There are many examples such as language translation tools, mobile traffic services, digital mapping technologies, instant spell checkers, and spam and fraud detection tools (Thierer 2013). The list just keeps growing. This data can provide a deeper understanding of the overall market, because it can pick up on people's wants and tendencies. This is due to the improvement of the consumer database as a result of increasing consumer information. With data analysis, companies can also find better marketing strategies in order to get people to buy their items because they now have more insight on their market. Data analysis can also help in identifying and preventing racist, threatening, or other malicious content that is on the web. These strategies have been used by companies and government agencies to track down potential terrorist attacks or other potential harmful scenarios. On the consumer end, it provides greater personalization of content (Truyo n.d.). Firstly, there are catered advertisements which encourage people to spend money on things they most likely enjoy. Furthermore, on social media sites, different user recognition technologies help to display only relevant content to that specific person based on their usual interests in general social media content. Lastly, as

mentioned before, the data that is collected from consumers is ideal for feeding into different online learning algorithms because of its optimal volume, velocity, and variety.

The third factor of SCOT that is important is technological frame. There are three domains within technological frame: the nature of technology, technology strategy, and technology in use (YOUSEFIKHAH 2017). One part that is currently missing is a bit of the first domain. Users know some of the functionality of the technology but may not fully understand the full capabilities and functions of the technology. If all the different stakeholders in this issue consider the capabilities of the technology, it can be optimized so that it can satisfy all parties. The second domain is super important in helping combat the negative light that consumer data collection is in right now. A lot of the stakeholders like the technology companies may not necessarily have the right motivation for implementing these technologies. Due to the capitalist nature of the United States, these companies may have set up data collection so that they make the most profit and users may also not know the full extent to which their data is being collected. This may not bode well for all the users of these different platforms. For companies to make data collection a beneficial technology for all, their strategy would need to be modified. Another factor which may not be as prevalent is that developers of other technologies in other companies should also conscious on how data collection works through the usage of existing technological platforms. Facebook invented their signature "Graph API" (Solon & Farivar 2019) which is an Application Programming Interface (API) that developers use in order to utilize in order for different programs to interact with one another. Facebook has been able to collect data through these services as well which means that they can not only collect data from users on their own platform but are enabled to collect data on any platform that uses their API. This scrutiny came about from other companies, like Spotify and Yahoo, who were complaining that their data was

being stolen by Facebook through the usage of these APIs ("Facebook's data-sharing deals exposed" 2018). This means that there is also a conflict between different companies between the technology collected, not just between a user and a company. Bigger companies with more data are at an advantage in this scenario because they are the ones controlling the vast majority of user information. Lastly, people know how to use these technologies and generally are also contributing to the problem. People willingly put all of their personal information on these sites, so they should expect their data to go somewhere behind the scenes. Facebook many times during their lawsuits has claimed that people have accepted the privacy agreement that shows up when you first download the app and agree to sharing your data that is collected on the app. It is important for users to recognize this and post only the content in which they are okay being public.

Overall, using SCOT we can see that there are many social factors that influence the technology of consumer information collection. The different players need to know how the technology works to build greater acceptance. By laying out all of the societal factors that go into the development of this technology, there may be a way for this technology to be more universally accepted and used in a positive way so that some of the downsides that people harp on can be overcome.

Conclusion

As I continue to research my topic further, my main focus is to examine the positives and negatives of consumer data collection. With inspiration from my technical project of collecting information on how much people donate to panhandlers, I was interested in all the other positive uses of collecting user data. In my case, the government has little to no data on how much money

homeless people actually receives and how much money people are spending to these causes so the data that my application will retrieve will actually be helpful. Nowadays, society is consumed on all the negatives, focusing on breach of privacy and malicious intent on the usage of this information, even though there are a lot of positive uses of this technology. Using SCOT, I plan to analyze all the different societal factors that play into this technology in order to find out how it is accepted and how it could be better improved to fit the needs of all the different stakeholders.

References

Allyn, B. (2020, August 04). Class-Action Lawsuit Claims TikTok Steals Kids' Data And Sends It To China. Retrieved December 9, 2020, from https://www.npr.org/2020/08/04/898836158/class-action-lawsuit-claims-tiktok-steals-kids-data-and-sends-it-to-china

Dickler, J. (2019, January 15). More Americans say they don't carry cash. CNBC. https://www.cnbc.com/2019/01/15/more-americans-say-they-dont-carry-cash.html

Digital payment platforms primed to topple cash. (2017, August 16). https://www.usbank.com/newsroom/news/digital-payment-platforms-primed-to-topple-cash.html

Doffman, Z. (2020, August 11). Is TikTok Seriously Dangerous-Do You Need To Delete It? Retrieved December 9, 2020, from https://www.forbes.com/sites/zakdoffman/2020/07/11/tiktok-seriously-dangerous-warning-delete-app-trump-ban/?sh=15a54f252b0e

Facebook's data-sharing deals exposed. (2018, December 19). Retrieved December 9, 2020, from https://www.bbc.com/news/technology-46618582

Giuliani-Hoffman, F. (2020, October 10). Researchers gave thousands of dollars to homeless people. The results defied stereotypes. Retrieved October 17, 2020, from https://www.cnn.com/2020/10/09/americas/direct-giving-homeless-people-vancouver-trnd/index.html

Graham-Harrison, E., & Cadwalladr, C. (2018, March 17). Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. Retrieved December 9, 2020, from https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election

Hunt, X. (2017, October 17). Online learning: Machine learning's secret for big data. Retrieved December 9, 2020, from https://blogs.sas.com/content/subconsciousmusings/2017/10/17/online-learning-machine-learnings-secret-big-data/

Kiger, P. (2017, March 10). How Will Panhandlers Fare as Our Economy Goes Cashless? HowStuffWorks. https://money.howstuffworks.com/cashless-economy-affect panhandlers.html

Solon, O., & Farivar, C. (2019, April 18). Mark Zuckerberg leveraged Facebook user data to fight rivals and help friends, leaked documents show. Retrieved December 10, 2020, from https://www.nbcnews.com/tech/social-media/mark-zuckerberg-leveraged-facebook-user-data-fight-rivals-help-friends-n994706

Thierer, A. (2013, September 16). Relax and Learn to Love Big Data. Retrieved December 9, 2020, from https://www.usnews.com/opinion/blogs/economic-intelligence/2013/09/16/big-data-collection-has-many-benefits-for-internet-users

Tian, Y. (2020, November 5). Personal Interview [Online Video Interview].

Truyo. (n.d.). The Importance of Data: The Top Benefits of Collecting Customer Data. Retrieved October 20, 2020, from https://insights.truyo.com/consumer-data

White, R. (2020, November 22). Personal Interview [Online Video Interview].

YOUSEFIKHAH, S. (2017). Sociology of Innovation: Social Construction of Technology Perspective. *Ad-Minister*, *30*, 31–43. https://doi.org/10.17230/ad-minister.30.2