

**Donor Mining for the Children's Inn**  
**Ethics of Donor Mining to Compete Against Nonprofits**

A Thesis Prospectus  
In STS 4500  
Presented to  
The Faculty of the  
School of Engineering and Applied Science  
University of Virginia  
In Partial Fulfillment of the Requirements for the Degree  
Bachelor of Science in Systems Engineering

By  
Clare Hammonds  
November 24, 2020

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  
Clare Hammonds

Date:  
11/24/2020

Approved:  
William T. Scherer, Department of Engineering Systems and Environment

Date:

Approved:  
Catherine D. Baritaud, Department of Engineering and Society

Date:

Competition in the nonprofit sector for resources, namely donations, has been increasing steadily in recent years. According to the National Center for Charitable Statistics (2020), there was a 4.5 percent growth in the number of nonprofits registered with the Internal Revenue Service from 2006 to 2016, and an almost 16 percent increase in total charitable giving in the United States. Although the increase in giving is larger than that of the number of organizations, the market for nonprofits is becoming increasingly competitive as each organization tries to stand out from similar nonprofits and attract new donors. Some studies have shown that although the average revenue per nonprofit is increasing, the median is decreasing, indicating a potential disproportionate distribution of funds (Beaton & Hwang, 2017). Because of this, each nonprofit organization must find a way to retain their current donors, as well as attract potential new donors. One such method is data mining, a technique that “uses statistical analysis, artificial intelligence, and machine learning technologies to identify patterns that could not be found by manual analysis alone” (Wang et al., 2010, p. 43). With data mining, organizations can more easily and accurately identify individuals with a high propensity to give, making it an approach that has been becoming increasingly popular amongst nonprofits.

One such nonprofit is the Children’s Inn (CIN) at the National Institute of Health (NIH). Serving more than 1,500 children each year, CIN is a place for families with children participating in research at the NIH to stay without cost (<https://childrensinn.org/learn-more/>). Their annual report (2020) explains that each year they have around \$13 million in income and \$12 million in operating expenses (p. 9). To stay competitive in their market, CIN has asked us to provide them with insights about their current donor database using data mining techniques. These insights will be provided as they appear throughout the year, and we hope to also provide them with a dashboard, as the culmination of our work in the spring, using predictive analytics to

retain and identify donors. The STS research, tightly coupled with the technical work, will examine the ethics behind data mining, specifically when used as a competitive mechanism in the nonprofit space.

## **DONOR MINING**

This project will allow my team to gain exposure to and experience in the art of data mining. Led by Professor William T. Scherer, the Associate Chair in the Department of Engineering Systems and Environment, the team also includes Joshua Eiland, Sofia Ponos, and Shawn Weigand, all students of fourth year standing pursuing B.S. degrees in Systems Engineering. We have also enlisted the help of Jerry Montgomery, cofounder of 5W, a company that uses machine learning and data mining to provide insights and models to their clients (<https://5wstrategists.com/about.html>). In working so closely with CIN, Liza Cole, the director of leadership and legacy giving, and Peter Springer, donor relations manager, are both integral parts of the project as well and provide bi-weekly feedback. While we have not required funding thus far, CIN paid a fee to participate in the capstone project, the funds of which can be accessed by the team should it become necessary to acquire more data or purchase licenses for mining and analysis software.

The objective of this project is to provide insights and models that CIN does not currently possess. As a public nonprofit, since CIN receives some funding from the NIH which in turn is funded by the government, CIN is in a sector growing even more quickly than the overall nonprofit sector, since public nonprofits grew by 19.6 percent from 2006 to 2011, as compared to 4.5 percent for all nonprofits (National Center for Charitable Statistics, 2020). In a growing competitive environment, CIN has hired outside consultants and firms in the past to perform analyses for them, but we hope to bring a new perspective to the problem and leave CIN with a

lasting model rather than a one-time analysis. Lenczner and Phillips (2012), describe nonprofit data mining, saying “if properly analyzed, it could give them the ability to identify new funders, better predict their behavior, and make more robust resource plans” (p. 13). Our hope is to create a dashboard that continually analyzes the data of CIN to more accurately retain current donors and identify potential new donors.

Ms. Cole and Mr. Springer initially provided us with many datasets to get us started with our analysis. This included lists of constituents of the Inn and donations made since CIN’s opening. The list of constituents includes both donors and non-donors who have been solicited by CIN. We also received appeals-specific data, giving us better insight into the results of specific efforts by CIN. In addition, we collected census data about the constituents with the help of the UVA library system, and a large data set from Mr. Montgomery, which is similar to the census data but which contains more fields of data pulled from various sources. Finally, by the end of this semester, we will have the results of a survey run by CIN polling their donors about the donor-nonprofit relationship. As Waters (2011) says, nonprofits must devote time and resources to create and maintain these relationships, as he proves in his study that there is a correlation between donation amounts and donors’ perceptions of their relationships with the nonprofit (p. 465). Knowles and Gomes (2009) develop this idea further, creating a model called AID-TIM, which stands for awareness and understanding, interest and involvement, desire to help, trial gift, information about what and how to give, and major gift action (p. 388). The survey that CIN has sent out indicates that they are making efforts to understand their donors, and much of the rest of the AID-TIM model is recorded already in their database. While our team will likely not use the AID-TIM model extensively, it provides a base for us on how to interpret the survey results and combine those results with the rest of our analysis.

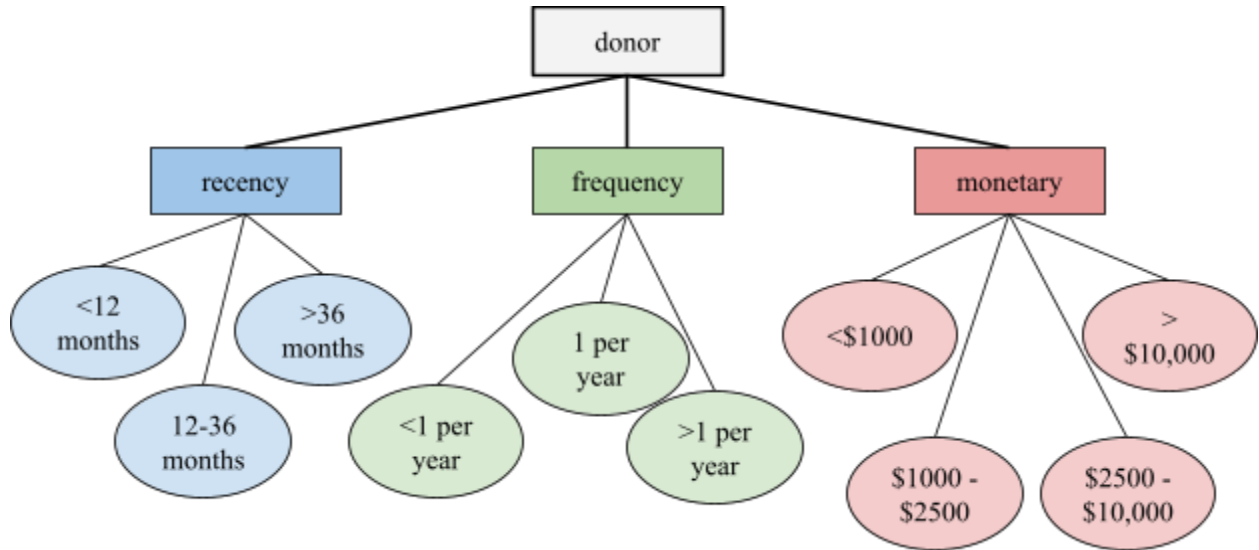


Figure 1: RFM depiction of a CIH donor. Each donor is classified by certain buckets within each of the recency, frequency, and monetary categories. (Hammonds, 2020).

The team began with trying to form some initial understandings of the datasets, using excel, R, minitab, and tableau to perform analysis and visualizations. Moving forward, trying to combine “analytical skills, information construction, and professional knowledge” (Wang et al., 2012, p. 42), we will be trying to create recency, frequency, and monetary (RFM) as well as lifetime value calculations for each donor. To do this, we have started analysis as seen in Figure 1 above, by separating each of the RFM categories into buckets, where each constituent will fall into one bucket in each category, giving them an overall RFM “score”. To create lifetime value statistics, this value will need to be combined with the potential legacy gift in order to figure out what type of person is expected to have the highest lifetime value for CIN. This is very important to CIN, as they have expressed to us already that retention is hugely important for them, noting the existence of the “Northern Stars”, a group of donors who, according to the CIN annual report (2020), have been donors for more than twenty consecutive years. After identifying characteristics in the current donors with the highest expected lifetime giving according to our

model, those characteristics will be applied to the dataset from Jerry to try to identify potential new donors. While CIN would have to pay for that specific dataset, and we therefore cannot share specific names from our findings with them, we can at least generate an idea of whether or not our model works, and how much can be expected to be brought in by each additional donor.

A large goal for the team is to identify types of people that have not previously been solicited to by CIN but who would perform well in our model, which is difficult because this information is most likely not directly in the data given to us by CIN. The team will potentially combine our model with social network analysis, which gives, as Johnson et al. (2010) describe, “a precise, quantitative process through which social structures and its constituent relationship patterns can be operationalized, mapped, and measured” (496). It is unclear at this point in time which networks we would look at, but we have considered social media platforms such as LinkedIn. Finally, the team will create a dashboard that will continually analyze the transactional and constituent data from CIN to determine which types of donors are most likely to be most beneficial to CIN, and what methods should be taken to target those donors, as well as retain current donors. The work will culminate with a technical thesis, written for submission to the Systems and Information Engineering Design Symposium (SIEDS). Our team hopes to make a meaningful contribution to CIN and leave them with a lasting system that will allow them to continue to utilize data mining techniques without having to perform the technical work.

## **ETHICS OF DONOR MINING TO COMPETE AGAINST NONPROFITS**

While data collection of individuals is not new, and neither is data mining, mining has only recently become a phenomenon and is now gaining interest very quickly as more companies and organizations learn the value of this new analysis approach (Cook & Cook, 2003, p. 396).

Because this method has developed and taken hold so quickly, it is largely unregulated. Frank Pasquale, a professor at the University of Maryland School of Law, suggests that “[r]ather than saying, ‘Everything is permitted, and we’ll try to legislate against certain things,’ it goes in the opposite direction” (as cited in Fowler, 2020). However, this is not yet the case, and so the fact that many people may not even be aware of data being collected in the first place, much less being used in other organizations’ marketing efforts, is disconcerting (Fowler, 2020). This brings to light one of the largest ethical concerns surrounding data mining, the general lack of consent being given for data to be used in the ways that it is. van Wel & Royakkers (2004) mention the benefits of data mining, saying “s single fact can take on a new dimension when it is combined with another fact, or when it is compared with similar facts” (p. 131) but also acknowledge that “people prefer to focus on the advantages of web-data mining instead of discussing the possible dangers” (p. 133).

The advantages of data mining certainly exist, or it would not be a current debate in the analytics world. But do the disadvantages outweigh the benefits, especially in the nonprofit sector? Bopp et al. (2017) certainly think so, arguing that there is increasing disempowerment within nonprofits when data is used to drive decisions, and they often drift away from their missions and goals in the process of using data to move forward (p. 3609). van Wel & Royakkers (2004) also bring up the issue of de-individualism, which can occur “[w]hen the judgement and treatment of people is based on patterns resulting from web-data mining” (p. 131). Some of these major ethical issues are outlined in Figure 2 below, which shows the related problems for social groups as related to CIN, although this model could certainly be applied to any nonprofit organization.

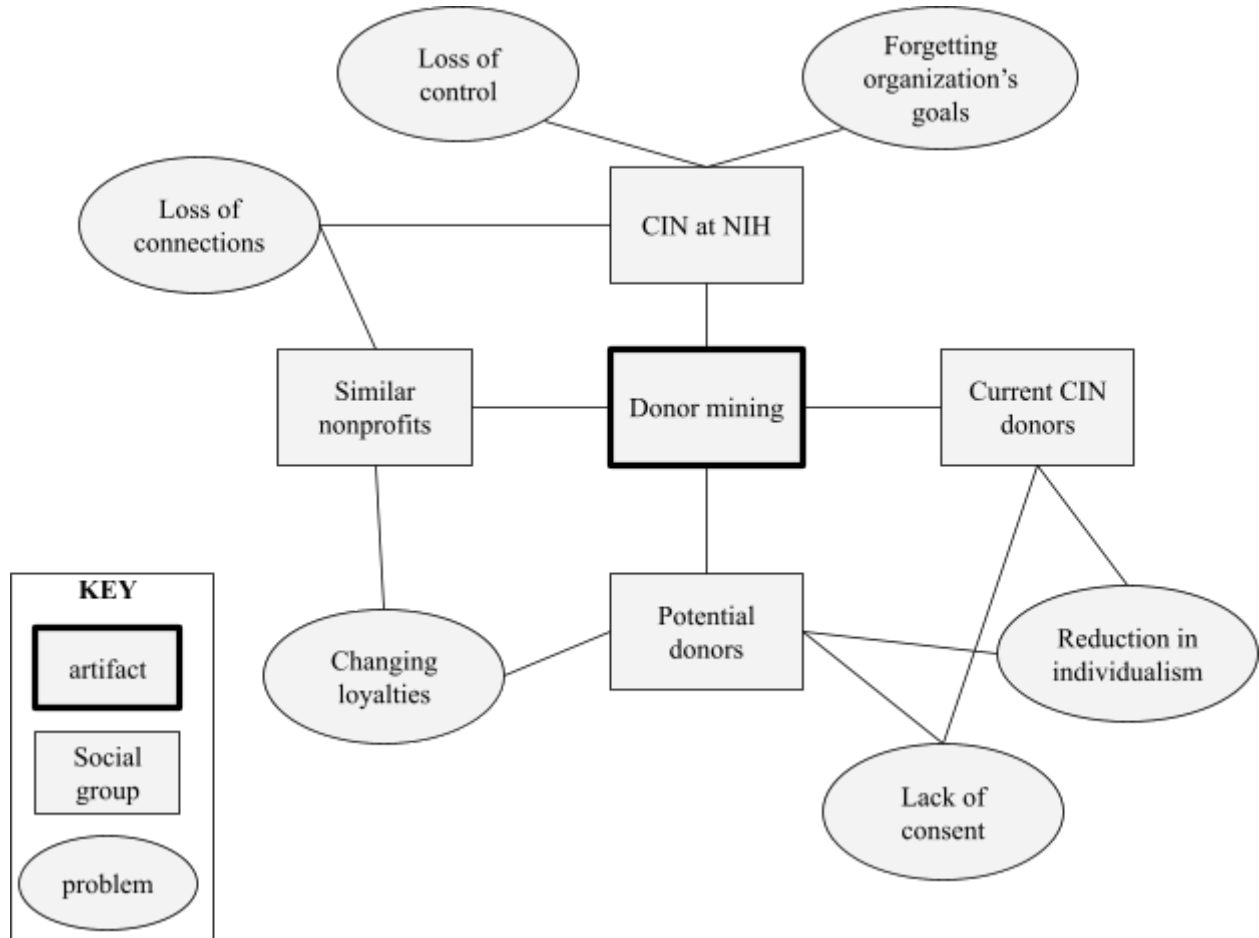


Figure 2: The major problems presented by donor mining technology. Donor mining involves multiple social groups which each have stakes in certain problems associated with the technology. (Adapted by Hammonds from Bijker et al., 1970).

Although the ethical dilemmas surrounding data mining in general will be a part of my STS research, the focus will be on the more specific topic of using data mining in order to compete in the nonprofit market. The nonprofit market is getting increasingly crowded, and as Beaton and Hwang (2017) describe it, a larger pie is being cut into more and more pieces, and solicitation is the way to increase one's piece of the pie (p. 216). Barman (2002) explains that nonprofits will try to differentiate themselves to make themselves stand out in the nonprofit market (p. 1192). As nonprofits often work with similar organizations to collect data, it may be a



concern that trying to identify donors that would potentially be strong donors for CIN would pull those donors from similar organizations. As seen in Figure 3 below, a SCOT model can be applied for CIN to see which stakeholders are involved in donor mining, and what each of them values related to mining. CIN must be acutely aware of these things in order to preserve each relationship. Another concern is whether or not donor mining gives some nonprofits an unfair competitive advantage in a delicate market. Tuckman (1998) might argue that because nonprofits often must also compete with for-profit organizations, donor mining is, in a way, leveling the playing field (p. 186).

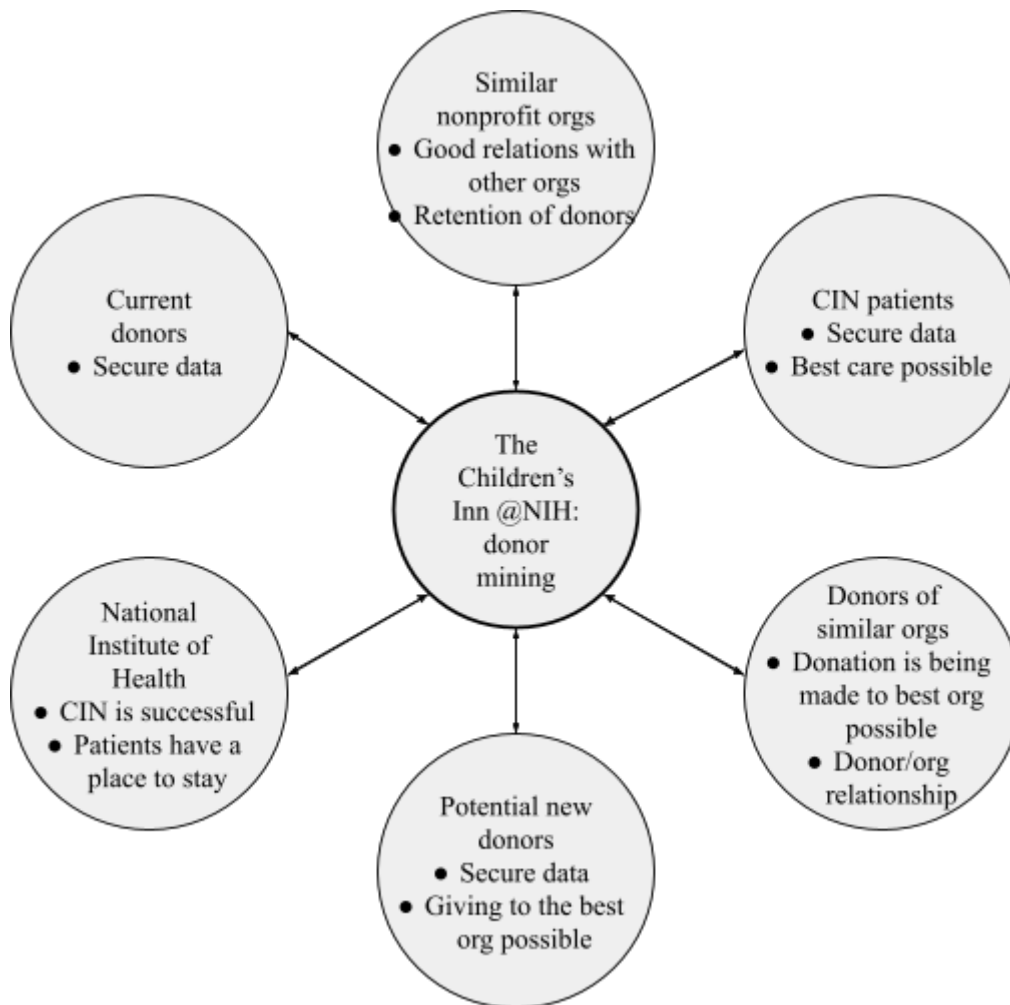


Figure 3: SCOT depiction of the Children's Inn at NIH using donor mining. The Children's Inn must understand what each stakeholder values and what each could gain and/or lose from CIN employing donor mining. (Adapted by Hammonds from Carlson, 2009).

My hope with my STS research is to use the SCOT framework to further examine how each stakeholder views donor mining and would respond to it being effectively used. I also would like to combine research on the ethics of data mining with the research on nonprofit competition in order to further examine the question of whether certain organizations deserve donors more than others and should be able to use donor mining to accomplish that. In examining the ever-growing debate surrounding data mining, I hope to write an article, without bias, clearly outlining the benefits and drawbacks of a nonprofit using donor mining to compete in its market.

## REFERENCES

- Barman, E.A. (2002). Asserting difference: the strategic response of nonprofit organizations to competition. *Social Forces*, 80(4), 1191-1222. doi:10.1353/sof.2002.0020.
- Beaton, E., & Hwang, H. (2017). Increasing the Size of the Pie: The impact of crowding on nonprofit sector resources, *Nonprofit Policy Forum*, 8(3), 211-235. doi-org.proxy01.its.virginia.edu/10.1515/npf-2016-0012
- Bijke, V.W., Bonig, J., & van Oost, E. (1970). The social construction of technological artifacts. In J. Woodforde, *The story of the bicycle* (pp. 39-51). London: Routledge & Kegan Paul.
- Bopp, C., Harmon, E., & Volda, A. (2017). Disempowered by data: nonprofits, social enterprises, and the consequences of data-driven work. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems: CHI 2017*. Retrieved from <https://doi.org/10.1145/3025453.3025694>
- The Children's Inn at NIH. (2020). *2019 Annual report*. Retrieved from <https://childrensinn.org/learn-more/performance-and-impact/>
- Cook, J.S., Cook, L.L. (2003). Social, ethical and legal issues of data mining. In J. Wang (Eds.), *Data Mining: Opportunities and Challenges* (pp. 395-420). Hershey, PA: Idea Group Publishing.
- Fowler, G.A. (2020, June 18). Nobody reads privacy policies. This senator wants lawmakers to stop pretending we do. *The Washington Post*. Retrieved from <https://www.washingtonpost.com/technology/2020/06/18/data-privacy-law-sherrod-brown/>
- Hammonds, C. (2020). *RFM depiction of a CIH donor*. [Figure 1]. *Prospectus* (Unpublished undergraduate thesis). School of Engineering and Applied Science, University of Virginia. Charlottesville, VA.
- Hammonds, C. (2020). *The major problems presented by donor mining technology* [Figure 2]. *Prospectus* (Unpublished undergraduate thesis). School of Engineering and Applied Science, University of Virginia. Charlottesville, VA.
- Hammonds, C. (2020). *SCOT depiction of the Children's Inn at NIH using donor mining*. [Figure 3]. *Prospectus* (Unpublished undergraduate thesis). School of Engineering and Applied Science, University of Virginia. Charlottesville, VA.
- Johnson, J., Honnold, J., & Stevens, F. P. (2010). Using social network analysis to enhance nonprofit organizational research capacity: a case study. *Journal of Community Practice*, 18(4), 493–512. doi.org/10.1080/10705422.2010.519683

- Knowles, P., & Gomes, R. (2009). Building relationships with major-gift donors: a major-gift decision-making, relationship-building model. *Journal of Nonprofit & Public Sector Marketing*, 21(4), 384–406. doi.org/10.1080/10495140802662580
- Lenczner, M., & Phillips, S. (2012). From stories to evidence: how mining data can promote innovation in the nonprofit sector. *Technology Innovation Management Review*, 2(7): 10-15. doi.org/10.22215/timreview/575
- Tuckman, H.P. 1998. Competition, commercialization, and the evolution of nonprofit organizational structures. *Journal of Policy Analysis and Management*, 17: 175-194. doi:10.1002/(SICI)1520-6688(199821)17:2<175::AID-PAM4>3.0.CO;2-E
- Urban Institute National Center for Charitable Statistics. (2020). In *The nonprofit sector in brief 2019*. Retrieved from <https://nccs.urban.org/publication/nonprofit-sector-brief-2019>
- van Wel, L., & Royakkers, L. (2004). Ethical issues in web data mining. *Ethics and Information Technology* 6, 129–140 (2004). doi:10.1023/B:ETIN.0000047476.05912.3d
- Wang, Z., Yan, R., Chen, Q., & Xing, R. (2010). Data mining in nonprofit organizations, government agencies, and other institutions. *IJISSS*, 2(3), 42-52. doi:10.4018/jisss.2010070104
- Waters, R. D. (2011). Increasing fundraising efficiency through evaluation: applying communication theory to the nonprofit organization— donor relationship. *Nonprofit and Voluntary Sector Quarterly*, 40(3), 458–475. doi-org.proxy01.its.virginia.edu/10.1177/0899764009354322