A HOLISTIC APPLICATION OF VALUE SENSITIVE DESIGN IN BIG DATA APPLICATIONS: A CASE STUDY OF TELECOM NAMIBIA

Emilia Shikeenga, Rosa Gil, Roberto García
Universitat de Lleida (Spain)
es16@alumnes.udl.cat; rgil@diei.udl.cat; roberto.garcia@udl.cat

ABSTRACT

In order to encourage ethical considerations and integrity in Big Data applications that incorporate Machine Learning techniques, this paper introduces a case as to how we intend to apply Value Sensitive Design (VSD) methodology in the design of a Telecom Customer Churn Prediction model. The VSD approach identifies stakeholders throughout the design process and this assists in steering clear of any biases in the design choices that might compromise any of the stakeholders' values. In this paper, we realize a VSD conceptual investigation of a churn prediction model, including stakeholder identification and the selection of human values to be included in the design.

KEYWORDS: big data, machine learning, telecommunications, human values, value-sensitive design.

1 INTRODUCTION

In recent years, big data technologies have been putting some pressure with regard to what is deemed acceptable or not acceptable from an ethical point of view. A great deal of the literature that focuses on ethical issues related to big data mostly concentrates on the following values: privacy, human dignity, justice or autonomy (La Fors et al., 2019).

Telecom Namibia is facing ever-increasing competition from new entrants such as MTN, Paratus Telecom, and Capricorn Mobile. With these new entrants, all chasing the same pool of customers and declining customer spend, Telecom Namibia needs to be able to retain its customer base in order to protect its revenues and ensure growth.

According to Harvard Business Review (Gallo, 2014), it costs between 5 times and 25 times as much to find a new customer than to retain an existing one. Thus, preventing customer churn is quickly becoming an important business function.

Telecom Namibia currently has a very basic churn model in place, which simply looks at churn on the basis of how many customers have discontinued the use of telecom services but that is too late to win back the customer. Consequently, this churn model is no longer practical nor efficient.

Telecom needs to have a more robust churn model built to predict customer churn with machine learning algorithms. Ideally, telecom can nip the problem of unsatisfied customers in the bud to keep the revenue flowing and ring-fencing its customer base.

During the development of the churn model, the team will also take the opportunity to explore the customer data by determining the different personality types of each customer through accessing their personal social media profiles. This will allow the team to provide proof that it is in fact possible for companies to "use" their customers' personal data in various ways that might be unethical.

This will therefore require that the value aspect be considered because the model is going to utilize data that is sensitive which might have value implications. It is for this reason why the churn model design process will employ the Value Sensitive Design approach.

Value Sensitive Design is a theoretically grounded approach to the design of technology that accounts for human values in a principled and comprehensive manner throughout the design process (Friedman et al., 2013).

Our approach will be to apply Value Sensitive Design to design a churn prediction model for Telecom Namibia. To allow us to proactively make use of the values, we will be engaging the stakeholders throughout the design process including the prototype development. The study will consider the VSD values starting from those listed by Friedman et al. (2013) and focusing on the values for big data technologies listed by La Fors et al. (2019), as shown in Table 1.

Through the implementation of this case, this paper will provide support on incorporating ethics and human values in Big Data applications. The findings will outline and demonstrate how viable the VSD approach is for providing a more comprehensive view and balancing of human values and ethics in Big Data applications.

Table 1 Integrated view of human values from different domains. Source: La Fors et al. (2019)

Technomoral values (Vallor 2016)	**Values from value-sensitive design (VSD)** (Friedman et al. 2006)	**Values from Anticipatory emerging technology ethics** (Brey 2012)	**Values in biomedical ethics** (Beauchamp & Childress 2012)	**Integration: values for big data technologies**
Care	Human Welfare	Well-being and the common good	Beneficence	Human welfare
Autonomy	Autonomy	Autonomy	Autonomy	Autonomy
Humility, self-control	Calmness	Health, (no) bodily and psychological harm	Non-maleficence	Non-maleficence
Justice	Freedom from bias; Universal usability	Justice (distributive)	Justice	Justice (incl. equality, nondiscrimination, digital inclusion)
Perspective	Accountability	N/A	N/A	Accountability (incl. transparency)
Honesty, self-control	Trust	N/A	Veracity	Trustworthiness (including honesty and underpinning also security)
N/A	Privacy; informed consent; ownership and property	Rights and freedoms, including Property	N/A	Privacy
Empathy	Identity	Human dignity	Respect for dignity	Dignity
Empathy, flexibility, courage, civility	Courtesy	N/A	N/A	Solidarity
Courage, empathy Environmental	Sustainability	(No) environmental harm, Animal welfare	N/A	Environmental welfare

2 APPROACH

Our approach draws on the Value Sensitive Design theory and involves three types of investigations: conceptual, empirical, and technical (Friedman et al., 2013, La Fors et al., 2019). As the goal of our research is to design a churn prediction model for Telecom Namibia using VSD, our research consists of the following investigations:

1. **Conduct conceptual investigations** to find the indirect and direct stakeholders, plus the values that are implicated. To achieve this, we have to identify the different stakeholders including discovering how they are affected and the values that are implicated with regard to the implementation of the application. Applying stakeholder analysis (Friedman & Hendry, 2019) we identify:
    - **Policy Makers:** This includes the government Republic of Namibia, as well the regulators- Communications Regulatory Authority of Namibia (CRAN) which is

mandated to regulate the telecommunication services and networks in Namibia. It also includes the Ministry of ICT**.**

- **Contractors:** Any person or firm that undertakes a contract to provide materials or labor to perform a service or do a job for Telecom Namibia**.**
- **Competitors:** Other companies in Namibia that offer the same products and services offered by Telecom Namibia. The competitors include MTC, MTN, Paratus Telecom, and Capricorn Mobile.
- **Shareholders:** The organizations and individuals that have a stake in Telecom Namibia.
- **Customers:** The people, organizations, businesses, etc. who buy and apply for the products and services that Telecom Namibia offers and makes use of those services.
- **Marketing and Sales representatives**: Responsible for monitoring customer churn and coming up with relevant solutions to retain customers.
- **Lead Data Scientist:** In charge of developing the churn model.

2. **Choosing the ethical values to consider:** the values are selected according to the Telecom Namibia company values and the value considerations for techno-social change in Big Data contexts presented by La Fors et al. (2019). Telecom Namibia's company values are (Telecom Namibia, 2017): Integrity, Care, Commitment, Accountability, Empowerment, Teamwork and Mutual Respect.
   The following VSD values are the values we will consider for the particular case of the design of the churn prediction model using Big Data techniques: human welfare, ownership, and property, autonomy, calmness, universal usability, accountability, trust, privacy, identity, courtesy and sustainability (Friedman et al., 2013).

3 ANALYZE THE TECHNOLOGY AGAINST VALUES

Following the previous approach, in Table 2 we carry out an analysis of the technology against the values. The analysis includes hashtags such as #Risk and #Need4Action to indicate any ethical risk or as an indicator where actions need to be taken to address specific challenges. The technology can be seen as INPUT, MODEL and OUTPUT, as detailed in the following table.

Table 2 Analyzing the technology against values. Source: Open Roboethics Institute. (2019)

	Value Questions	Telecom Namibia
INPUT		
Transparency	Do the relevant stakeholders know how/when the information is collected/changed/used?	No
	Are the data provided by the stakeholders used to collect any secondary sources of information (e.g., connected to social media profiles, external online platforms)? If so, are the stakeholders informed of this?	No, so far telecom Namibia does not use data to access users' social media. A note is that we will be using the data merely to also prove that a company can access user data on their social platforms to for example find out their personalities.

Autonomy & Consent	Is there an informed consent process in place for the data collection that outlines the fact that the data can be used for this use case?	No and it is not required.
	Can the stakeholders decide not to have their data used for the algorithmic system?	Not yet #Need4Action
	Are there any elements in the data collection process (e.g., user interface used for inputting data) that could result in unintended outcomes?	No
Fairness	If people are involved in directly collecting data from someone/something, how diverse are these people in terms of race, gender, age, class, and other socioeconomic factors? Teams of people who are similar to one another can lead to similarly biased observations and data entries.	No data needs to be collected, all data to be used is already in the company's databases.
	Are certain groups of stakeholders' information collected disproportionally more than others? If so, does this fact support or conflict with the societal and stakeholder values?	No
Human Rights	Does the input data include sensitive/identifying information (e.g., gender, race/ethnicity, religion, location of work/residence, education, social and professional associations/groups)?	Yes #Risk
	Can the stakeholders opt not to enter the sensitive/identifying information?	Yes
MODEL		
Transparency	Will the model and its performance be understandable to and monitored by those training it?	We will include this premise in our design.
	If there is a questionable/erroneous outcome or an incident in the future, is it possible to explain to a third party what aspects of the model led to the outcome/incident?	Yes
Accountability	How often is the model updated/re-trained and is the frequency adequate for the use case?	Firstly we will train it every evening and ensure that the frequency is adequate for the use case.
	Who oversees the model training/updating process and are they the right people who can	The Lead data scientist will be overseeing the process and will be able to detect any new

	detect new problems and act upon them?	problems and act upon them.
Fairness	Are there sources of bias that could lead to unfairly discriminating against individuals/groups, especially against specific gender, race/ethnicity, religion, social class or otherwise marginalised groups?	No
	Are there any parameters or technical aspects of the system that can contribute to biases in the output against specific gender, race/ethnicity, religion, social class or otherwise marginalised groups?	No
Human Rights	Is the model designed to reveal or predict an individual's identity (e.g., sexual orientation), potential (e.g., a child's probability of success in life), such that it contradicts with stakeholder and societal values, including human rights?	No
OUTPUT		
Transparency	Will the output from the algorithm presented in such a way that is understandable to its audience?	The team will be working to ensure it is understandable to its audience
	Is the output presented to the stakeholders in a way that allows them to understand how/why the system has produced the specific output? Is it important for them to understand this?	It will be important for Telecom Namibia to know how the system produces a particular output when the output is surprising.
Trust	Will the output from the algorithm translated from a probability score to a categorization (e.g., 90% probability of being X is presented as being X)? Is the translation of the probability to categorization appropriate for the use case and trustworthy?	It will be important to have the translation of the probability to categorization appropriate for the use case and trustworthy
	Will the technology and its output have the potential to lead to a destructive cycle of behaviours or operations (e.g., reinforcing gender bias of those who are the primary source of input data)?	There will be periodic supervision and monitoring of the design process.
	If someone were to take the outputs from the system and generalise it to other use cases, is it	There is a possibility that could happen. Any bias that will be perpetuated by the algorithm

	reasonable to foresee problematic interpretations or increase in distrust among stakeholders?	may be wrongly interpreted as facts.
Accountability	Who will be responsible for acting on the output, and does this stakeholder group have ways to remedy or override erroneous or questionable output?	Telecom Namibia will be responsible for using the output appropriately. Will have to design a way for Telecom Namibia to handle erroneous recommendations
	Is there a communicated and unobstructed means for different stakeholder groups to raise an alarm on possibly dangerous usage of the technology?	No #Need4Action
	For cases where sensitive findings arise from the outcome, is there a clear means for different stakeholder groups to deal with the potentially uncomfortable truths (burden of knowledge)?	No #Need4Action
	What are the implications of false positives? What are the implications of false negatives? Are the appropriate decision makers aware of the balancing of risks between the two?	Yes, false positives can have severe impacts on any initiatives that will be undertaken to address churn since that would increase the cost of retaining a customer. For e.g if the model assigns as someone to be more likely to churn but is not the case then the organization would essentially be spending money trying to retain customers that were never really at risk of leaving the company because of the false positive predictions of high risk. On the other hand, false negatives can cost the company more than false positives. In this case, the model predicts customers as not churning, while customers actually will churn. The company will therefore lose profits by making them leave without doing any action for them not to churn.
Autonomy/ Consent	Is the output connected to another process or technology without human intervention being necessary? If so, are the risks from worst case scenarios minimal and acceptable?	No, it will not
	Will the technology be designed to replace or assist human decisions? If it is meant to replace	Yes, it will be designed to assist human decisions and techniques from HCI will be used.

	them, is it meant to support the overall function of the stakeholders whose decisions are being replaced?	
Fairness	Will the primary users of the technology be aware of the potential biases that may have contributed to the output?	We will work towards that #Need4Action
	Will the stakeholders who are subjected to the technology be given a means of remedy?	We will work towards that #Need4Action
	Will the output produce the same result for all users? Does it lead to unfairness or discrimination?	There will be some supervision and reviewing to ensure that the output yields the same results for all users.
	Will the output lead to fair distribution of wealth, opportunity, or other positive outcomes?	There will be some supervision and reviewing to ensure that the output leads to positive outcomes.
Human Rights	Will the technology suppress or protect fundamental human rights, such as right to life, liberty, security, freedom of movement and of expression, among others?	We will follow international standards focused on technology to ensure that they are currently being developed as an IEEE Ethically Aligned Design. (IEEE, 2019)

4 CUSTOMER SERVICE SURVEY

To test the hypothesis that is based on user experience, which is: One of the reasons why customers churn is because of poor customer service (Retention Science, 2019), we carried out a customer service survey. The survey was conducted online using google forms and shared among Namibians, about 25 participants took part. We present our findings below:

- About 80% of the participants agree that they value staff at the call center/customer service are friendly and helpful
- About 68% of the participants agree that they value staff who provide them with good feedback and solutions to any issues/problems they might encounter with the products/services.
- About 68% of the participants agree that they value more products/services that fully meet their needs.
- About 64% of the participants agree that they value more products/services that are innovative.
- About 52% of the participants agree that they value more products/services are better compared to competitors' products/services (quality)
- About 44% of the participants agree that they value more the pricing of Products/services is reasonable.

Overall to sum up everything, the survey indeed revealed that customers value good customer service and should they be at the receiving end of poor service, they would most likely churn.

5 CONCLUSIONS AND FUTURE WORK

The increase in providing Ethical considerations in Big Data has become a concern and the values are also indicated in the ACM Principles for Algorithmic Transparency and Accountability (ACM, 2017). This paper introduced the application of VSD in telecom customer churn models construction. We have identified the direct and indirect stakeholders of Telecom Namibia and identified the associated human values. VSD has proven to be a promising approach in promoting ethical considerations in Big Data applications.

Our future work for this study is to clearly outline in detail how we applied VSD through the design process of the Telecom Namibia Churn Model. We will be researching and analyzing any laws or norms around the chosen values and we will define the design requirements. Another step will be to consider how we will verify/evaluate whether the designed model embodies the chosen human values.

REFERENCES

ACM. (2017). Statement on Algorithmic Transparency and Accountability. Available from: https://www.acm.org/binaries/content/assets/public-policy/2017_joint_statement_algorithms.pdf

Customer Survey. Available from: https://forms.gle/KfhVhiBBzxYfUAC49

Friedman, B. & Hendry, D. G. (2019). Value Sensitive Design: Shaping Technology with Moral Imagination. Cambridge, MA: MIT Press.

Friedman, B., Kahn, P. H., Borning, A., & Huldtgren, A. (2013). Value sensitive design and information systems. In: Early engagement and new technologies: Opening up the laboratory (pp. 55-95). Springer, Dordrecht.

Gallo, A. (2014). The Value of Keeping the Right Customers. Harvard Business Review. Available from: https://hbr.org/2014/10/the-value-of-keeping-the-right-customers

IEEE. (2019). Ethically Aligned Design. Available from: https://standards.ieee.org/industry-connections/ec/autonomous-systems.html

La Fors, K., Custers, B., & Keymolen, E. (2019). Reassessing values for emerging big data technologies: integrating design-based and application-based approaches. Ethics and Information Technology, 1-18.

Open Roboethics Institute. (2019). Foresight into AI Ethics (FAIE): A toolkit for creating an ethics roadmap for your AI project. Available from: https://openroboethics.org//ai-toolkit/

Retention Science. (2019) The Data-Driven Marketer's Guide to Predicting Customer Churn. Available from: https://go.retentionscience.com/retention-marketing

Telecom Namibia. (2017). 2016/2017 Annual Report. Available from: https://www.telecom.na/downloads/reports/2016-17/Annual%20Report%202016-2017.pdf