



Research Article

A Cross-sectional Study on Improving Privacy Policy Read Rate and Comprehension via Better UX/UI Design

Glorin SEBASTIAN

School of Public Policy, Georgia Institute of Technology, Atlanta, GA, USA
gsebastian6@gatech.edu

Received date: 10 December 2020; Accepted date: 18 January 2021; Published date: 16 February 2021

Academic Editor: Yuping Li

Copyright © 2021. Glorin SEBASTIAN. Distributed under Creative Commons Attribution 4.0 International CC-BY 4.0

Abstract

Data privacy is becoming one of the major risks faced by organizations. With the emergence of new technologies such as Internet of Things (IoT), Automated vehicles, mobile payments, privacy is no longer a compliance issue, but also a user responsibility to ensure users are aware of the risks of sharing personal data with third parties. There is a big effort to research on how Privacy policies can be made easier to read and comprehend. In this paper, we share the results of a cross-sectional study on users who read Privacy statement from an autonomous vehicle provider (Tesla as case study) and compared its effectiveness to that of a video visualization that explained the policy. The study focused on the users' understanding and recall. Our research shows that interfaces with better UX/UI design (User experience/User interface) greatly increase click and read through rates as well as help the user comprehend risks associated with privacy notices better.

Index Terms- Privacy Policy, Autonomous vehicles, UX/UI Design.

Introduction

In the research study by Taeihagh, Araz, and Lim, Hazel Si Min. (2018), a self-driving car, also known as an autonomous vehicle (AV), connected and autonomous vehicle (CAV), driverless car is a vehicle that is capable of discerning the environmental inputs and drive with no human instructions. Self-driving cars combine a variety of sensors to perceive their surroundings, such as radar, lidar, sonar, GPS, odometry and inertial

measurement units. Advanced control systems interpret sensory information to identify appropriate navigation paths, as well as obstacles and relevant signage.

The report published by Allied Market Research, Global Autonomous Vehicle Market by Level of Automation, Component, and Application: Global Opportunity Analysis and Industry Forecast, 2019-2026, notes that the global autonomous vehicle market is valued at \$54.23 billion in 2019,

and is projected to reach \$556.67 billion by 2026, registering a CAGR of 39.47% from 2019 to 2026.

As the autonomous vehicle market grows, the IoT market is expected to grow at a similar pace, if not faster. Security concerns surround autonomous vehicles, given that the vehicle is constantly sharing telematics data. These telematics data include VIN number, software information, present speed and distance information, which are very sensitive as they clearly let someone know of the driver's exact location, and car information details which could be used by a malicious hacker to perform cyber hacks and attacks on connected cars.

Another privacy issue associated with Autonomous vehicles is unsolicited marketing. If the user doesn't read the privacy statement fully, they might miss that the vehicle vendor could have contracts with third parties to enable third party cookies which gather information about the user which in turn can be used by the third parties for unsolicited marketing including spam emails. Privacy policies are used by companies to notify users about data practices and to validate their agreement with privacy requirements. Privacy policies provide information such as what data are collected from the users, with whom they are shared, what steps are taken to protect the collected user data. The research study by F. Schaub, R. Balebako and L. F. Cranor, May-June 2017, notes that the privacy notices are lengthy in general that describe the company's data practices and they typically not only fail to influence privacy decisions but are largely ignored by the users. However, privacy policies in general are ineffective due to the following reasons:

1. *The complexity of language and legal jargons used:*

The jargon related to regulations used in the policies is difficult to understand. Based on McDonald et al., (2008), reading a privacy policy takes on an average 10 Minutes. However, understanding most privacy policies requires higher educational level. As per the 2019 US Census Data, only over one in four adults (27 percent) had attained at least a bachelor's degree. A

study by Schaub et al. (2017) discusses the reasons behind the ineffectiveness of privacy policies and they show that most of these privacy policies are not read by the users or if they are read, the users cannot understand the privacy risks due to the complexity of the privacy policies.[3]

2. *Readers do not understand what is relevant and important:*

One very important reason for the esoteric nature of the Privacy policies is that the policy is laid out without any indication of which part is relevant or important to the end user and which is not. There have been studies by Shayegh, Parvaneh, and Ghanavati, Sepideh (2017) that illustrate how this can be solved, with methods such as:

Dividing the privacy policy into Definition, Facts, Action, Non-Relevant, and Cross References section and then removing the Definition, non-Relevant sections and creating cross-referenced section, graph from annotated action section, generate Notice and Choice.

3. *Channel for showing Notices:*

Selecting the correct channel to display the Privacy notice is yet another challenge. Most websites display the privacy notice on the computer screens, however some IoT devices face challenge here due to the limited screen space. Using the video approach to the Privacy Policy also helps the users to hear the privacy Policy if there are limitations to the IoT display interface.

Some of the related work includes Sepideh et al. (2017) who has suggested the use of AI and NLP (Natural Language Processing) to parse the policies to determine which sections need to be part of the Privacy statement and which do not.

The approach suggested by Shayegh, Parvaneh, and Ghanavati, Sepideh (2017) includes two phases, one is analysis of Privacy Policies via annotation and the second phase is annotating action with its sub-annotations into five categories including:

#D (Definition): defines the key terms including what is data, privacy, and security

#F (Facts): this section includes important information about the business of the company that forces them to collect, use or share user data

#A (Action): Most important part of the Privacy policy for the end user as it describes the actions that can be done by the user including the options to opt-out from data collection

#NR (Non-Relevant): Mainly consists of the part that is added to make the privacy terms more readable, examples for each statement, titles etc.

#CR (Cross References): Includes the links to other pages as cross references

Finally, the suggestion is to remove #D and #NR, create cross-referenced section #CR including a graph from annotated action section.

Our proposed approach builds on the research performed by Shayegh, Parvaneh, and Ghanavati, Sepideh (2017). But our recommendation is to perform the analysis and divide the Privacy Policy into 5 sections by the Policy experts in the company along with legal assistance.

In this study, we have studied the understanding of an autonomous vehicle Privacy statement (Tesla, 2013) in pdf format, against a video created from this Policy statement based on the relevant parts of the pdf. The approach proposed in this paper is to use the emerging

technologies of Analytics and Data visualization as well as animation to ensure Policy statements are better read and that the users are fully aware of the risks and their options before accepting the Privacy policies.

Materials and Methods

This study was conducted among 172 study participants over a period of two months. 172 respondents who were randomly split into two groups of 86 each. Each of the respondents was emailed an online survey. The control group had to read the sample Privacy Policy in pdf and answer the questions, while the intervention group had to watch the video shared and answer the same 5 questions.

Only the main sections that the user needs to be aware of were included in the video including other references and details on the FAQ document section. The main considerations while designing the video format for this Policy included:

1. The user needs to be made aware of which of their personal information is collected in the form of Telematics data as well as the steps taken to ensure that the stored information is safeguarded. The user should be aware of which of his data would be shared with third parties
2. Compliance with State, Federal and international Privacy laws
3. How and for which data can the user opt-out from data sharing

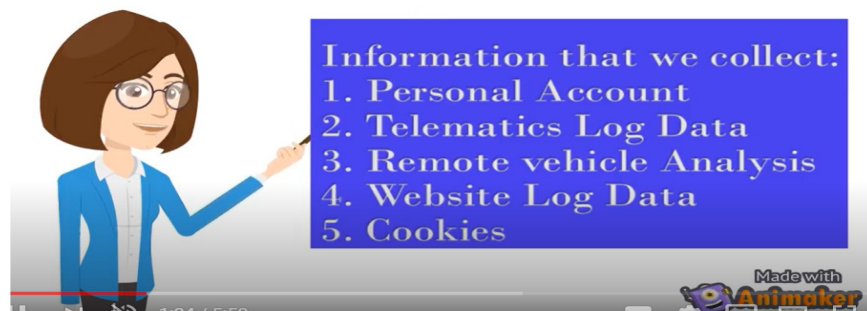


Fig 1. Screenshot from the YouTube video on Privacy Policy (Tesla), Sebastian G (2020)

Data Analysis: The data collected were analyzed and expressed as percentages and proportions for the purpose of this study. Please refer Table 1 for the Socio-demographic data of the study population, Table 2 provides the text statistics for the sample Tesla Policy statement, while Table 3 explains the readability test scores for the sample Tesla Policy statement used in the Study.

Results

The socio demographic data of the study population are shown in Table 1. Although the majority of the study participants of both intervention and control groups were well educated, they were not in the habit of reading privacy policies fully.

Table 1: Socio-demographic data of the study population

Sl.no	Demographics	Control group	Intervention group
1.	<i>Highest educational attainment</i>		
	Bachelor’s degree	5.8%	37.2%
	Master’s degree	94.2%	61.6%
	PhD degree	Nil	1.2%
	Did not complete Bachelor’s degree	Nil	Nil
2.	<i>Place of residence</i>		
	USA	11.6%	64%
	India	68.6%	19.8%
	Others	19.8%	16.3%
3.	<i>Habit of reading privacy policies fully</i>		
	Yes	12.8%	11.6%
	No	86%	73.3%
	Maybe	1.2%	15.1%

With an average reading speed of 300 wpm, the time taken for an average user to read the sample Privacy statement (Tesla) was 12 minutes. The average time taken for watching the video is 5.50 minutes. The text

statistics and readability test scores of the sample privacy policy are given in table 2 and table 3 respectively (Tesla Privacy Policy).

Table 2: The text statistics for the sample Tesla Policy statement

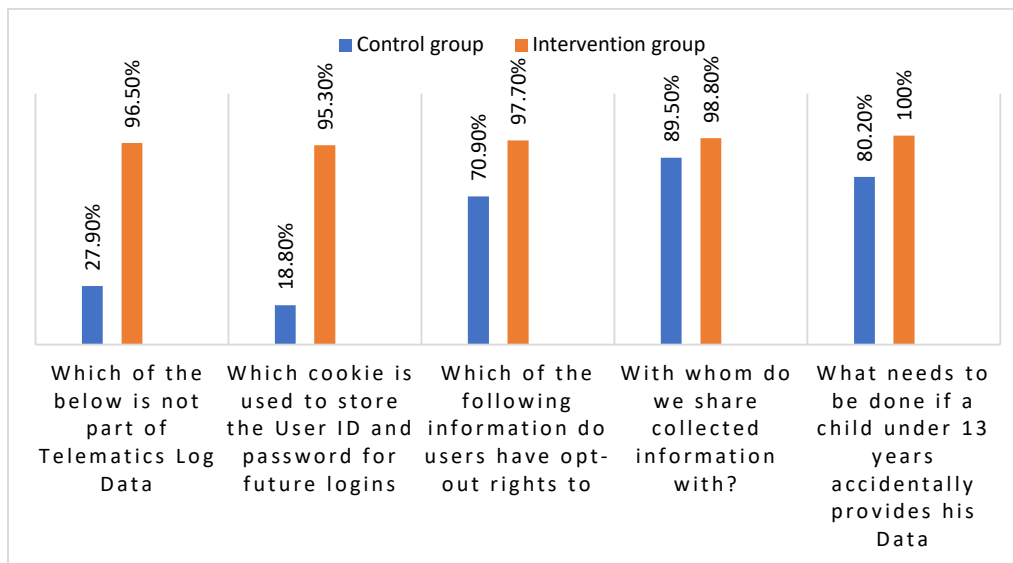
Text Statistics	Details
No. of sentences	176
No. of words	3977
No. of complex words	909
Percent of complex words	22.86%
Average words per sentence	22.60
Average syllables per word	1.80

Table 3: The readability test scores for the sample Tesla Policy statement explained

Readability test	Score	Required Education Level
Flesch Kincaid Reading Ease	31.6	College (Difficult to read)
Flesch Kincaid Grade Level	14.5	Difficult to read
Gunning Fog Score	16.4	College Senior (Difficult)
SMOG Index	13.1	College (Difficult to read)
Coleman Liau Index	14.7	College (Difficult to read)
Automated Readability Index	14.2	College (Difficult to read)

Assessment of the control group and intervention group using the same assessment questions revealed that the

intervention group had a better recall and scored higher compared to the control group. (Fig 2.)

**Fig 2. Comparison of assessment scores between intervention and control groups**

Discussion

Our proposed approach builds on the research performed by Shayegh, Parvaneh, and Ghanavati, Sepideh (2017). Our approach is to read through the Policy statement and divide it into 5 sections:

1. *Data Collected and Security*: The details of what data are collected from the end user, how they are processed and what security is enabled for these Data while transit and storage.
2. *Sharing information Policy*: This details about the third parties with whom the entity shares customer information.

3. *Purpose used for*: Includes the details on the purpose for which the Data collected are used
4. *Permit*: Opt-out options: This lists the opt-out options for the end user, including the rights and responsibilities of the user
5. *Compliance to state, federal and international privacy Laws*: The policy should also mention how it complies to state federal and international privacy laws.

Even though earlier research (Sepideh et al. (2017)) has suggested the use of AI and NLP (Natural Language Processing) to parse the

policies to determine which sections need to be part of the Privacy statement and which does not, based on our research, we determine that every company is different based on the industry it is associated with. Hence, we suggest parsing and determining the best sections to be visualized based on the review of the Company Policy experts, with legal background.

Once the sections required for the privacy Policy are finalized, the company needs to ensure all the decided five components need to be included in the Animation or the visual presentation for the end user.

From the results of the study, the video form of the privacy statement is much more effective for the end users. It takes ~12 minutes for the users to read the privacy policy, however the video cuts that time almost by half to less than 6 minutes.

The animated video helps the end users with better recall as well as helps them connect with in a much personal level with the content of the presentation. The animated version of Policy would greatly help change the perception towards compliance from dry and boring to interesting and relevant. If the company can create an animated avatar, that avatar can be used to represent the company in several other ads as well, thereby increasing the relatability of the company with the end users.

It increases the brand image and also makes the end users feel that the company cares about their Data and personal well-being, than including a long poorly constructed multi-page Privacy Policy as a check the box item. We believe that increased customer trust in the company brand would lead to better and longer relationships with the end users. It also reduces the compliance cost of companies, since the users would be better aware of the data privacy risks and follow the procedures which lead to lesser privacy Data issues.

The results show that based on the survey questions, the users in intervention group answered the questions in the same way as that of the control group, but after watching the video had a better recall than the users in the control group who read the Policy

document. The survey results show that using more Video and visualization cannot just reduce the time for the end users but also increases the effectiveness of information recall, thus helping the end users better understand the Privacy Policy. Also, from the comments' section of the survey, when asked if the video was better than the usual pdf format, the end users responded overwhelmingly positively saying that it greatly helped them concentrate and that the normal data format could be boring thus causing the users to not concentrate enough while reading through.

Future studies can include details on how the video visualizations can further be improved for example by making use of Chatbots, Artificial intelligence and other cutting-edge computer technologies.

Conclusion

Data Privacy concern is a growing issue for every entity that handles user personal information. With more US States proposing privacy laws, also a possible federal Privacy statute on the horizon, the number of Data privacy issues is bound to increase. Even though most Privacy laws require mandatory Privacy notices be displayed on websites and interfaces where customer private information is collected, as discussed in this study, most end users do not read these notices, and even if they do, given the esoteric nature of the Policy statement, they do not fully comprehend the content.

Through our study, we were able to demonstrate that the use of visualization can greatly impact the readability in a positive manner. Video visualizations increase the curiosity of the user and keep them engaged, while also effectively educating the user about the Policy details. This study applies to Policy documents in general as noted in the review article by Sebastian, G (2020), Risk and Controls including adequate policy documentation is becoming very crucial in most areas including system implementation. It is not only important for the law to mandate that privacy notices be displayed, or that the end users are notified of any changes to the

Privacy notices, but it is also required to make sure that the Privacy notices are readable. Based on our research and analysis, we found that using video visualizations greatly improves the readability of Privacy Policies.

References

- Autonomous Vehicle Market by Level of Automation (Level 3, Level 4, and Level 5) and Component (Hardware, Software, and Service) and Application (Civil, Robo Taxi, Self-driving Bus, Ride Share, Self-driving Truck, and Ride Hail) - Global Opportunity Analysis and Industry Forecast, 2019-2026
- M. McDonald and L. F. Cranor. (2008), 'The cost of reading privacy policies.', *ISJLP*, 4:543.
- F. Schaub, R. Balebako and L. F. Cranor, May-June 2017, "Designing Effective Privacy Notices and Controls," in *IEEE Internet Computing*, vol. 21, no. 3, pp. 70-77, doi: 10.1109/MIC.2017.75
- Sebastian, G (2020): "Automated vehicle-Privacy Policy,' *Youtube*. [Online], [Retrieved September 05, 2020], [//www.youtube.com/watch?v=h_mAuXGVQnI&t=281s](https://www.youtube.com/watch?v=h_mAuXGVQnI&t=281s)
- Shayegh, Parvaneh, and Ghanavati, Sepideh. (2017) "Toward an Approach to Privacy Notices in IoT.": 104-10. Web.
- Sebastian, G (2020), Evolution of the role of risk and controls team in an ERP Implementation, *IJMPERD*, ISSN (P): 2249-6890; ISSN (E): 2249-8001 Vol. 10, Issue 3, Jun 2020, 15529-15532
- Sebastian, G., 2021. An Exploratory Survey on The Perceptions Regarding the Inclusion Of Security and Privacy By Design (SBD, Pbd) Principles During Software Development Lifecycle (SDLC) Requirements Gathering Phase Among Business Analysts. *Zenodo*. Available at: <https://doi.org/10.5281/zenodo.4427206>.
-
- Taeihagh, Araz, and Lim, Hazel Si Min. (2018) "Governing Autonomous Vehicles: Emerging Responses for Safety, Liability, Privacy, Cybersecurity, and Industry Risks." *Transport Reviews* 39.1: 103-28.
- Tesla. (6/14/2013) "Privacy Statement. [Online],' [Retrieved July 31, 2020] https://www.tesla.com/sites/default/files/pdfs/tmi_privacy_statement_external_6-14-2013_v2.pdf
- US government. Retrieved from www.census.gov/data/tables/2019/demo/educational-attainment/cps-detailed-tables.html