Jason Yu

**DSP Project Report: Amazon**

The aim for this study was to gain an understanding of how people perceive the data privacy policies of Amazon and offer suggestions on how to address any concerns identified via survey.

**Are users comfortable with the service’s privacy defaults? Do Amazon users complain about the privacy terms?**

Users are overall comfortable with Amazon privacy defaults, but this may be due to a lack of information concerning the specifics of those policies rather than actual comfortability. The majority of responders (38/53) categorized themselves as knowing little or only vague details about the privacy terms surrounding their Amazon accounts. There were very few participants (2/53) who responded that they had or wanted to complain about Amazon privacy policies. A comparison between the responses to inquiries about user knowledge and user concern with privacy yields a negative correlation coefficient (Pearson) of -0.214. This suggests a negative relationship between familiarity with privacy policies and comfortability. Notably, many of the participants who indicated they have a ‘not at all’ familiarity with Amazon privacy policies also do not use Amazon frequently (compare graph 1,4). Examining only users who use Amazon frequently, the correlation coefficient (Pearson) between user privacy policy awareness and concern is -0.278. This suggests that infrequent usage of Amazon is not enough to explain why there is a negative relationship between awareness and concern.

**Are users aware of the data Amazon collects about them?**

Users are overall aware that Amazon does collect data about them, with only two participants responding that they were unaware of Amazon’s data collection. Both of these participants also indicated that they used Amazon ‘Not at All,’ indicating their lack of knowledge may have come from little interaction with Amazon. The majority of participants also responded that they knew how to access their Amazon data, but indicated only knowing how to access the basic account information provided on the Amazon website. The participants who did know about the option to access more detailed data from Amazon via ‘Data Subject Access Request’ also indicated that they were familiar with Amazon’s privacy policies and thought they were insufficient to protect user data.

In terms of categorical data that participants were aware Amazon collected on them, most indicated knowing that Amazon tracked their addresses (49/53), Order/Return History (48/53), and Search History (46/53). In sum, users are aware that Amazon collects their data, but are unaware of the extent of the data that Amazon collects.

**Supposing users were unaware of the data Amazon collects about them. Would their privacy concerns be more serious after explaining to them that Amazon does, indeed, collect data about them?**

There were only two cases of people who ‘Don’t know Amazon Collects their Data,’ and both indicated an increase in concern with Amazon’s privacy policies. After learning that Amazon does collect their data, one user went from feeling that their data was safe with Amazon to being unsure, while the other went from unconcerned to worried, but not enough to take action.

**Will users’ privacy concerns grow if they see the data Amazon collects about each of them?**

Overall participants indicated an increase in concern after being shown the example of the Kindle dataset meant to demonstrate the extent of data Amazon collects. Of random bootstrapped samples, 90% of mean differences in the pre and post examples answers to the question about user comfortability fall in the range [-1.15, -0.95] suggesting close to an average decrease in comfortability by at least 1 level. Comparing graphs 2 and 3, we can see that the responses to Question 2 appear to be normally distributed, but responses to the same question after the example are skewed towards the right. This indicates a left shift in responses and overall decrease in user comfortability.

**Recommendations**

**Increase Transparency:**

Participants of the study were not privy to the list of categories and were only asked to list what they thought answered the question. Responses were assigned to categories in real time because category names may be prone to misinterpretation. For example, a mention of recording what items are clicked on would be categorized as search history and not as website interaction, while a mention of recording time spent browsing the Amazon webpage would be categorized as website interaction.

The data gathered from the survey indicates a lack in participant understanding of the extent of the data which Amazon collects on users.

The data has been shown to indicate that there is a negative relationship between user awareness of the privacy policies Amazon has in place, and their concern with the protection of their personal information. This suggests that people are in general dissatisfied with the already existing policies or at least what policies they know about. Furthermore, the results from

Two suggestions:

Raising awareness of the Data Subject Access Request, perhaps through email. Many companies provide an end of the year overview of statistics about the user. Amazon could provide one with basic statistics, and include a notification that more data can be accessed via the Data Subject Access Request.

Tagging recommendations or services with the information used to increase transparency with regards to how user information is being used. This will allow the user to understand why all the categories mentioned in this survey (and perhaps those not mentioned) are necessary to improve the user experience. Furthermore, it would provide an opportunity for users to identify any data collected on them that they might want changed or removed, allowing them to more actively exercise the right to rectification and erasure.

**Ease of Concern Expression:**

Most users indicate even afterwards that they are worried after the example, but not enough to take action. This suggests a barrier to sharing privacy concerns, and the data also shows that this might be the case. Of the users who initially answered that they were worried about their personal information on Amazon (12/53), only 3 indicated that they would submit a complaint after the example was shown. The other 9 remained at worried but unwilling to act, despite many of them previously having little to no knowledge of their Amazon data. Providing methods to enable users to express their worries will lessen the burden of responsibility on the company and also allow the company to work towards addressing worries rather than not hearing about them. Some methods of deployment could involve incentivized user surveys on the website after a purchase or a review page (similar to product review pages) for Amazon informatics.

Graph 1 A picture containing text, screenshot, diagram, font

Description automatically generated A picture containing text, screenshot, diagram, font

Description automatically generated A picture containing text, screenshot, software, multimedia software

Description automatically generated