

The Importance and Value of Personal Identifying Information: Towards Information Worth

Jeffrey P. Kaleta
kaletajp@appstate.edu
Computer Information Systems
Appalachian State University
Boone, NC 28608, USA

Lakshman Mahadevan
lmahadevan@fgcu.edu
Information Systems and Operations Management
Florida Gulf Coast University
Fort Myers, FL 33965, USA

Russell Thackston
russell@thackston.me
ThackApps, CTO
Statesboro, GA 30458, USA

Abstract

Today's businesses aim to personalize online consumer experiences by making use of people's personal identifying information (PII). Whereas the phenomenon of privacy paradox has established that consumers are willing to reveal their PII in the presence of incentives, little is known about the worth of a consumer's PII. Understanding PII worth can help companies strategize their incentive structures that can motivate consumers to provide their PII and offer an enhanced experience while transacting online. In this study, we employ a survey instrument that answers the question of "What is the worth of PII?" and examine those perceptions across different groups of several socioeconomic indicators: education, income, and age. Our findings suggest that the worth of some PII considered in this study are not equally perceived across all groups, which can impact how researchers evaluate PII and how online businesses evaluate PII worth to improve the consumer experience.

Keywords: personal identifying information (PII), PII importance, PII value, PII worth, privacy paradox.

1. INTRODUCTION

Today's online businesses aim to provide seamless and personalized customer experiences through digital channels such as social media, web, and mobile apps. To be successful in this endeavor, companies are moving toward making use of information about individuals on a personal level instead of generic socioeconomic indicators such as age, income, and education. To

accomplish this, it requires the collection of personally identifiable information (PII) such as an individual's name, phone number, address, social media account, and their preference for various services and affiliations. In contrast, the same PII is also collected by nefarious actors that build doxing databases of individuals, relying mostly on stolen PII available on the dark web for a price. For instance, a hacked Gmail account sells for \$80 in 2021 (Sen, 2021) while for people

with high credit scores, a Social Security number, birth date, and full name can sell for \$60 to \$80 on the digital black market. For hackers, a victim with a good credit score can fetch a premium price (Kan, 2017). Some stolen identity information can go for as little as \$1 per person, or even \$0.10 when bought in bulk, according to a 2017 report from security firm Flashpoint. Such baseless and fluctuating valuations can make it difficult for online businesses to offer a suitable incentive to the consumer in exchange for their PII, leading to a negative experience.

Even though consumers have a protective attitude toward their PII, they have been known to reveal their PII in the presence of incentives and personalization, a phenomenon known as privacy paradox (Acquisti et al., 2015; Awad & Krishnan, 2006; Kokolakis, 2017; Martin, 2020; Norberg, Horne et al., 2007). Although a large amount of research exists regarding consumers' attitudes toward the collection of personally identifying data (Kolotylo-Kulkarni et al., 2021), there is little knowledge about the financial *value* that consumers assign to such data (Fehrenbach and Herrando, 2021) and whether the importance they attach to the data matches its monetary value. To our knowledge there are few examples where researchers have looked at the true value of PII as a financial transaction (e.g. Carrascal et al., 2013; Montes et al., 2019), but a gap still remains in the privacy literature from a socioeconomic perspective regarding the differences between information 'importance' versus its 'value', which we refer to as information 'worth'.

Thus, our research mainly aims to understand the nature of the incentives to be provided by 1) having customers assess information by importance and monetary value, 2) checking for consistency in correspondence between what customers consider as important information with its relative monetary value and 3) conducting a drill down to glean insights on how customers in different demographics -- such as education, income, and age -- ranks the worth of their PII. By considering the importance and value of the PII, companies can understand its worth and implement a strategy to provide incentives aligned with customer expectations, resulting in a higher likelihood of disclosure. Additionally, companies can directly incentivize customers in exchange for PII without having to involve third party providers.

Results of our study illustrate both the perceived importance and value of PII by an individual. We surveyed people asking them to rate the

importance of keeping certain PII private, based on an adapted scale. Additionally, we asked the respondents to attach a monetary value to PII based on the relative cost of a meal. In addition, we subdivided the responses by education, income, and age to determine if there were significant differences between each group.

From here, our paper is organized as follows. First, we introduce our methodology of collecting data on people's perception of PII. We follow by summarizing our analysis and results. We then provide a discussion of our findings. Lastly, we conclude with implications and steps forward with this research.

2. METHODOLOGY

Aligned with our research goals, we developed a survey instrument aimed at collecting people's perceptions of importance and value toward personal identifying information (PII). With this goal in mind, we adapted an instrument used by Fehrenbach and Herrando (2021) that identified the types of PII people find vital. Our adaptation asked people how important each type of information was to keep private. Expanding on this scale, we constructed an instrument requesting a person's perceived value of each PII with respect to their average cost of lunch. There is some research that links the cost of lunch to a family's socioeconomic situation (e.g. Domina et al., 2018), thus we felt this offered a normalized monetary reference to people at all income levels. People's budgets for lunch offers a baseline for people to evaluate their perceived value of PII. The measurement items can be found in the Appendix.

To collect the data, we administered the survey to a sample group through Amazon's Mechanical Turk. All participants who fully completed the survey received monetary compensation.

Once the data was collected, incomplete and/or erroneous data was removed from the final dataset. The dataset consisted of 222 valid responses with 33.3 % Females and 66.7 % Males. There were predominantly 3 levels of education: High School (or equivalent) (11.7%), College Degree (undergraduate) (64.0%), and Master's Degree (24.3%). Using information found at Beresford Research, (Brunjes, n.d.) we transposed each year of birth to their respective generation. Lastly, the participants had a normal distribution of income levels with the mean occurring at \$40 – \$59.9k (33.3%). Table 1 below provides a summary of the sample statistics.

Sample Size:	N = 222
<u>Gender</u>	
Female	33.3
Male	66.7
No response or other	0
<u>Level of Education</u>	
High School or Equivalent	11.7
College Degree	64.0
Master's Degree	24.3
<u>Generation (based on YOB)</u>	
Boomer I	3.6
Gen X	23.4
Gen Y	62.2
Gen Z	10.8
<u>Income Level</u>	
< \$20k	10.4
\$20k – 39.9k	20.3
\$40k – 59.9k	33.3
\$60k – 79.9k	20.7
\$80k – 99.9k	10.4
\$100k +	5.0

Table 1: Demographic Sample Statistics

3. ANALYSIS & RESULTS

We examined the data for each individual PII and their overall mean scores. Regarding information importance, the highest ranked PII was 'Current Bank Account Balance, followed by 'Social Media Account Access'. The lowest ranked PII was 'Political Preference, closely followed by 'First and Last Name'. With regards to information value, the highest value PII was also 'Current Bank Account Balance', followed closely by 'Social Media Account Access'. Like information importance, the least valued was 'Political Preference', and secondarily 'First and Last Name'. All PII and respective values are summarized in Table 2.

To evaluate the effect of education, income, and age, as potential socioeconomic indicators, we conducted a multivariate analysis of variance (MANOVA) with the information importance and value as dependent variables as it relates to each respective PII. Using Wilkes Lambda test for significance, the results showed that differences in education was significant on the dependent variables for all PIIs. Overall, this suggests that education level can impact different perceptions on the worth (importance and value) of PII.

With respect to income, there were no significant differences found. This suggests that regardless

of income, people have an equal view of the worth of PII.

Personal Identifying Information	Importance (1 to 5)		Value (-5 to +5)	
	Mean	Std Err	Mean	Std Err
First and Last Name	3.33	1.22	1.75	2.52
Personal Phone Number	3.76	1.04	2.39	2.22
Personal Street Address	3.91	1.04	2.35	2.28
Mother's Maiden Name	3.66	1.24	1.99	2.51
Political Preference	3.23	1.30	1.18	2.71
Current Bank Account Balance	4.14	1.00	2.75	2.09
Browser History	3.92	1.01	2.49	2.20
Social Media Account Access	3.97	1.04	2.62	2.13

Note: N = 222

Table 2: Descriptive Statistics

Lastly, in terms of generational (age) difference, significance was found based on 'First and Last Name' will all other PIIs showing no significance. This suggests there are generational gaps regarding perceptions of the worth of PII that is offered in online contexts. A complete breakdown of significant differences among all groups is offered in Table 3.

We also conducted a post hoc analysis of the results to see which dependent variable (information importance vs. information value) had significance among the different factors of education, income, and age. For 'First and Last Name' there was differing significance found for information importance with levels of education and information value for difference of age. Information value also had shown a significance between levels of education for 'Mothers Maiden Name'. Only 'Political Preference' shown significance for both information importance and value, whereas all other PIIs only shown significance for information value at differing levels of education. The post hoc summary is

indicated in Table 3.

Personal Identifying Information	Educ	Income	Gen
First and Last Name	<0.001* ⁱ	0.080	0.033* ^v
Personal Phone Number	0.001* ^v	0.187	0.322
Personal Street Address	<0.001* ^v	0.822	0.361
Mother's Maiden Name	0.019* ⁱ	0.515	0.146
Political Preference	0.004* ^{iv}	0.063	0.266
Current Bank Account Balance	0.006* ^v	0.720	0.478
Browser History	0.002* ^v	0.324	0.665
Social Media Account Access	0.022* ^v	0.925	0.296

note:

*All PII's satisfied assumptions test of normality
 Dependent Variables: Information Importance
 and Information Value*

*p values shown, * significant $p < .05$, N=222*

*ⁱ post-hoc analysis (Tukey) indicates
 significance between groups for Information
 Importance*

*^v post-hoc analysis (Tukey) indicates
 significance between groups for Information
 Value*

**Table 3: Multivariate Analysis of Variance
 Wilkes Lambda test for Significance
 (post hoc summary noted)**

4. DISCUSSION

The set of data items were chosen to reflect a holistic overview of the various data types that previous research has deemed as personally identifying and important to consumers (Carrascal et al., 2013; Huberman et al., 2005; Tsai et al., 2011). Table 2 shows that overall, consumers rank 'Current Bank Account Balance' and 'Social Media Account Access' as the most *important* and the most *valuable* PII. Also, our analysis shows that 'First and Last Name' and 'Political Preference' rank the least in terms of both information *importance* and *value*. We believe this suggests that consumers do have some consistency in the worth of PII. This finding

suggests that individuals are aware that their financial wellbeing, as shown in their current bank account balances is important to maintain as private over other PII's. Furthermore, we can see that there is high worth placed on a person's social media information, which can also indicate that people seek to maintain a degree of privacy as it pertains to their social lives when asked to divulge related information. This can lead to future research in these areas to determine as to the reasons people find high value in such types of PII's.

There are significant differences in the worth perceptions of PII with respect to their degree of importance and value based on the education level of the individual. Our findings show that differences exist in all PII variables as it relates to education. Additionally, our findings also suggest that education may be a unique factor for further research to understand incentive adjustments for PII requests, since a majority of the PII are significant for information value.

Examining differences in income on PII, we find that there was almost no difference in information importance or value. Overall, our findings suggest that regardless of income, the view of PII are equal between groups as identified in the MANOVA analysis in Table 3. This finding is important as it indicates that the same incentives can be given as income has minimal influence on PII worth.

Investigating generational (age) differences, the data suggests differences as it relates to the importance and value of PII. More specifically, some PII shows a significant difference based on its value but not importance as it comes to requests for 'First and Last Name'. However no other PII was found to show significance. When reviewing our sample demographics, there were fewer later generation participants (e.g., 3.6% of Boomer Generation), who had adapted using information technology later in life. Whereas the majority of other generations were born at time with information technology introduced earlier in their lifetime (e.g., Digital Natives). Although our findings may show minimal impact in this area, there is a need to explore this further to include a larger sampling of later generations.

Overall, our analysis shows that there is consistency in the ranking of PII in terms of information importance and value. The highest ranked PII was a person's 'Bank Account Balance' and 'Social Media Account Access'. However, upon deeper inspection, we find that differences of education have a larger impact on the worth of

information with a greater degree of significance found in terms of information value.

5. IMPLICATIONS

Our analysis shows that consumers bestow different worth for the PII considered in this study. This may be a critical component for researchers that study the privacy paradox, the idea people believe it important to maintain information private yet offer it for a benefit. The findings in the paper suggest that people are willing to offer their PII for a price that is consistent with its perceived worth.

Our study also reveals that there is a consistency in information worth across income and generation since the valuation and importance of PII across income and generation do not have significant changes. However, from an education viewpoint, information worth is perceived differently since information value and importance are perceived differently as seen in how importance remains the same for most PII, but value differs significantly between education levels.

Additionally, our study offers a unique measurement of information value providing a standardized monetary basis to evaluate PII. This is directly usable for researchers that seek to explore not just the importance of PII, but what value people place on PII.

We believe this research offers companies a perspective that some PII's are worth more than others and that the worth of PII's differ across socioeconomic variables, especially when considering the education level of the consumer. Thus, companies should pay attention to each PII and carefully design their monetary incentive for each PII to increase the likelihood of its disclosure.

6. LIMITATIONS

As with all research, limitations are present, and we recognize that there are gaps in this study that may limit the suggested findings. For example, we recognize there is a smaller distribution of Boomer generations, which when applied to our analysis, may cause some concern due to its imbalance across the dataset. This can be easily remedied by collecting data targeting this age group to increase the sample size in this dimension.

Also, we recognize that a survey instrument alone has its limitation and can introduce some

inconsistencies of people's true perceptions. To improve upon this, other studies may be able to add some qualitative analysis (e.g., via open ended questions) by targeting people's perceptions.

7. FUTURE WORK

With this study we can see there a need to look deeper into the educational and generational difference in the perceptions of the worth of PII. For example, our study illustrates there are some differences based on education, however, the question as to the degree of difference and cause of differences needs further investigation. This is also the same with differences found based on the different generations. Furthermore, it would be valuable to see what PII each individual group finds of highest worth. For example, what PII is of highest worth to Gen Z, or those with a High School level of education. There is need for future studies to understand these differences and create a ranking of PII worth between groups based on education and/or generation (age).

Future studies could consider other non-socioeconomic variables such as ethnicity and gender to understand the variation if any in the worth of different PII. Furthermore, since our study analyzed only a subset of consumer's PII, future studies can consider the worth of additional PII relative to modern consumers such as their health-related information.

Additionally, we can see the need for a information worth construct from a privacy paradox perspective since our results show that though importance and value are consistent over most of the PII's considered, there are perceived differently in the case of education levels possibly across demographic variables that can be considered in future research. This may require the construction and validation of a formative construct with indicators given that will provide a single construct of information worth.

8. CONCLUSION

With the aim of smoothing the consumer experience, we investigated the perception of PII worth and whether the incentives offered can remain consistent across the demographic groups considered in this study. To this end, we conducted a survey of individuals asking questions regarding different types of PII and their importance and relative value. Our findings show that there are significant differences on the perceptions of PII worth based on an individual's education level and age. Also, we show that

perceptions of PII worth are consistent across gender. This study opens new perspectives to both the research community examining the privacy paradox and businesses seeking to collect PII to benefit the consumer experience.

9. REFERENCES

- Acquisti, A., Brandimarte, L., & Loewenstein, G. (2015). Privacy and human behavior in the age of information. *Science*, *347*(6221), 509–514. <https://doi.org/10.1126/science.aaa1465>.
- Awad, N. F., & Krishnan, M. S. (2006). The personalization privacy paradox: An empirical evaluation of information transparency and the willingness to be profiled online for personalization. *MIS Quarterly*, *30*(1), 13–28. <https://doi.org/10.2307/25148715>.
- Brunjes, K. (n.d.). Age Range by Generation. Beresford Research. Retrieved July 12, 2022, from <https://www.beresfordresearch.com/age-range-by-generation/>
- Carrascal, J. P., Riederer, C., Erramilli, V., Cherubini, M., & de Oliveira, R. (2013). Your browsing behavior for a Big Mac: Economics of personal information online. In *Proceedings of the 22nd International Conference on World Wide Web* (pp. 189–200). ACM. <https://doi.org/10.1145/2488388.2488406>
- Domina, T., Pharris-Ciurej, N., Penner, A. M., Penner, E. K., Brummet, Q., Porter, S. R., & Sanabria, T. (2018). Is Free and Reduced-Price Lunch a Valid Measure of Educational Disadvantage? *Educational Researcher*, *47*(9), 539–555. <https://doi.org/10.3102/0013189X18797609>
- Fehrenbach, D., & Herrando, C. (2021). The effect of customer-perceived value when paying for a product with personal data: A real-life experimental study. *Journal of Business Research*, *137*, 222–232.
- Flashpoint (2017) <https://flashpoint.io/blog/pricing-analysis-deep-dark-web/>
- Huberman, B. A., Adar, E., & Fine, L. R. (2005). Valuating privacy. *IEEE Security & Privacy*, *3*, 22–25. <https://doi.org/10.1109/MSP.2005.137>
- Kan (2017) <https://www.pcmag.com/news/heres-how-much-your-identity-goes-for-on-the-dark-web>
- Kokolakis, S. (2017). Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon. *Computers & Security*, *64*, 122–134. <https://doi.org/10.1016/j.cose.2015.07.002>.
- Kolotylo-Kulkarni, M., Xia, W., & Dhillon, G. (2021). Information disclosure in ecommerce: A systematic review and agenda for future research. *Journal of Business Research*, *126*, 221–238. <https://doi.org/10.1016/j.jbusres.2020.12.006>.
- Martin, K. (2020). Breaking the privacy paradox: The value of privacy and associated duty of firms. *Business Ethics Quarterly*, *30*(1), 65–96. <https://doi.org/10.1017/beq.2019.24>.
- Montes, R., Sand-Zantman, W., & Valletti, T. (2019). The Value of Personal Information in Online Markets with Endogenous Privacy. *Management Science*, *65*(3), 1342–1362.
- Norberg, P. A., Horne, D. R., & Horne, D. A. (2007). The privacy paradox: Personal information disclosure intentions versus behaviors. *Journal of Consumer Affairs*, *41*, 100–126. <https://doi.org/10.1111/j.1745-6606.2006.00070.x>.
- Sen (2021) <https://www.pbs.org/newshour/science/here-s-how-much-your-personal-information-is-worth-to-cybercriminals-and-what-they-do-with-it>
- Tsai, J. Y., Egelman, S., Cranor, L., & Acquisti, A. (2011). The effect of online privacy information on purchasing behavior: An experimental study. *Information Systems Research*, *22*(2), 254–268. <https://doi.org/10.1287/isre.1090.0260>

APPENDIX A Survey Instrument

Information Importance (adapted from Fehrenbach and Herrando, 2021)

Question 1: For the following personal information, what **degree of importance** do you feel it is to maintain as private? *5 point likert scale: Not important to keep private, Slightly important to keep private, Moderately important to keep private, Important to keep private, Very important to keep private.*

1. First and Last Name
2. Personal Phone Number
3. Personal Street Address
4. Mother's Maiden Name
5. Nationality
6. Political Preference
7. Current Bank Account Balance
8. Browser History
9. Social Media Account Access
10. Preferred Brand of Underwear

Information Value (new)

Question 2: For the following personal information, what **amount of value** would you expect to receive if shared? *11 point polar scale: -5 Less than the cost of lunch, 0 Average cost of lunch, +5 More than the cost of lunch*

1. First and Last Name
2. Personal Phone Number
3. Personal Street Address
4. Mother's Maiden Name
5. Nationality
6. Political Preference
7. Current Bank Account Balance
8. Browser History
9. Social Media Account Access
10. Preferred Brand of Underwear