Valuation of Personal Data in the Age of Data Ownership (Paper-a-thon)

Hannah Stein
Saarland University, hannah.stein@iss.uni-saarland.de

René Herbert Reich
Katholieke Universiteit Leuven, rene.reich@kuleuven.be

Lucian Visinescu
Texas State University, llv19@txstate.edu

Valuation of Personal Data in the Age of Data Ownership

Paper-a-Thon

Hannah Stein
Saarland University
Saarbruecken, Germany
hannah.stein@iss.uni-saarland.de

René H. Reich
KU Leuven
Leuven, Belgium
rene.reich@kuleuven.be

Lucian L. Visinescu
Texas State University
San Marcos, Texas, United States
llv19@txstate.edu

Abstract

In order to tackle uncertainties about data ownership and data misuse, more accessible and competitive data markets are proposed, especially concerning the use and access rights of data generated by the Internet of Things (IoT) devices. Legal proposals suggest that companies and individuals become owners of their self-generated data, enabling new ways of data monetization. Still, individuals are often uncertain about the value and price of their own generated data. This research builds on construal level theory to propose influencing factors fostering an understanding of intraindividual data value. The results of a pilot study survey (n = 104), conducted during the ICIS 2022, show that data proximity and data sensitivity factors significantly influence intraindividual data value. Our research extends the knowledge on data value from individual perspectives and builds the foundation for future work on data valuation and pricing in intraindividual data trading.

Keywords: Intraindividual Data Value, Construal Level Theory, Psychological Ownership, Data Markets

Introduction

Individuals lack data control more and more, for example, in social networks – remember the Cambridge Analytica Scandal before and during the presidential election in the US in 2016. The response to this lacking control is a trend toward more accessible and competitive data markets (European Commission 2020). For instance, the Data Act of the European Union proposes that companies and individuals will become owners of their self-generated IoT data, meaning they will obtain more control over their data (Geiregat 2022). Consequently, the control over data includes a reinforced data portability right, i.e., copying or transferring data across different services, where data are generated through intelligent objects, machines, or devices (European Union 2022). For example, an individual owning a connected car can decide what to do with this data – unlike the current situation in which an automotive group selling such connected car also retains the data generated. While individuals could start monetizing their data based on this law, many do not know how much their data are worth. The intraindividual data valuation, which builds the foundation for developing data value and prices for trading and sharing, remains unclear (Karampela et al. 2018).

Existing approaches to the valuation of personal data focus on privacy as as influencing factors for specifying a data value (Cichy et al., 2014; Hubert et al. 2020.; Krasnova et al. 2014; Li et al. 2014). Still, the privacy paradox, describing discrepancies between users' privacy concerns and their disclosing behavior, could also bias its influence on data valuation. Individuals are likely to value being in control of personal data for reasons other than privacy. Furthermore, existing approaches for personal data valuation mostly neglect the individual perspective while focusing on pricing and value interests from the buyer side. Alternatively, they leave aside data-specific attributes in their approaches (Farrelly and Chew 2016; Gkatzelis et al. 2015; Spiekermann et al. 2012; Spiekermann and Korunovska 2017).

We are therefore aiming to answer the following research question:

> *Which factors influence the intraindividual value perception of personal data from a data holder and buyer perspective?*

The following section describes the theoretical background of our research. We draw on the literature on the valuation of personal data combined with Construal-Level Theory and Psychological Ownership as the theoretical foundations. We then present our research design, conceptual model, and method for a quantitative study before describing and discussing preliminary results. Finally, we conclude the research with future work and current and expected contributions.

Theoretical Background

Drawing upon construal level theory and psychological ownership, we theorize how individuals value their personal data. We rely on the term of intraindividual data valuation, i.e., the valuation is occurring within the individual. The term “intraindividual” is often applied when changes or differences within the individual are described. As we will elaborate below, valuation of personal data through individuals differs, e.g., in terms of distance to the data. Therefore, intraindividual data value indicates that the valuation is subject to different assessments within the individual.

Construal level theory (CLT) explains the ability of humans to think about psychologically distant events (Trope and Liberman 2010). Central to this theory is psychological distances to an object or event from direct experience. According to the CLT, people perceive an event differently depending on the degree to when (temporal distance), where (spatial distance), to whom (social distance), and whether (hypothetical distance) it will happen. In our case, providing or selling data to a third party corresponds to disclosing data (event) from exclusive possession. Hence, it is reasonable to assume that psychological distances influence the perceived data ownership, the willingness to disclose the data, and the intraindividual value of the data.

Likewise, psychological ownership relates to the feeling of possession over a target (such as an object, concept, organization, or another person) that may or may not be supported by formal ownership. Whereas ownership regulations on data are in preparation in the European Union, no formal data ownership currently exists yet. Psychological ownership defines not only the object but also the owner. Individuals often become invested in the ownership target as an expression of who they are and to which they belong (see Dittmar 1992; Pierce et al. 2001). The individual has a personal stake in the performance of the object, as its performance reflects upon their identity (Pierce et al. 2001). Consequently, this leads to a feeling of possessiveness, a desire to retain ownership and a mental attachment to the target (Pierce et al. 2001). Psychological ownership in the context of data disclosure could influence how willing individuals are to remove data, i.e., how willing to disclose they are (Cichy et al. 2014).

Based on literature analysis on data value and price, we derived two intraindividual factors influencing the value perception of personal data through the lenses of construal level theory. These factors (or distances) mainly focus on the individual level but also might be valuable from a data buyer perspective.

First, we draw on the factor of **data proximity,** describing the psychological distance between someone and a data set (Lee et al. 2018; Pu and Grossklags 2015; Trope and Liberman 2010). Data proximity relates to psychological proximity aspects, such as emotional connectedness to something (Lee et al. 2018), and shows how important data are to its owner (Bauer et al. 2012). Studies point to the fact that data about oneself are valued higher than data about friends (Pu and Grossklags 2015), and data identifying a specific individual are valued higher than anonymous data (Li et al. 2015). Second, **data sensitivity** is the potential loss associated with data disclosure (Mothersbaugh et al. 2012) and thereby represents a factor related to data removal. For example, individuals perceive higher risks of psychological loss when dealing with higher personal sensitivity level data (Moon 2000). Studies suggest that the degree of sensitivity influences the value perception of personal data (Benndorf and Normann 2018; Cichy et al. 2014; Schomakers et al. 2019).

We, therefore, propose data proximity and sensitivity as insightful antecedents of the intraindividual value of personal data, as represented in the conceptual model below (Figure 1). Building on these main factors, we derive the following hypotheses for this work:

H_{VP}: High data proximity is associated with a high intraindividual data value.

H_{VS}: High data sensitivity is associated with a high intraindividual data value.

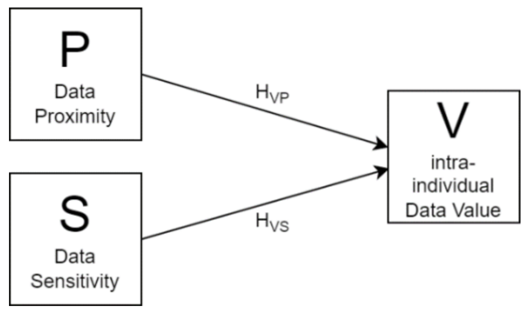

Figure 1. Price-Influencing Factors for Personal Data from Individual Perspective

Research Design and Method

Within the paper-a-thon timeframe of twenty-four hours, we derived a first conceptual model on how individuals develop a value estimation of their personal data. Figure 2 depicts the research process followed within our work. We focused on data proximity and sensitivity since the literature thereon is sparse (see Theoretical Background). Further, we wanted to preliminary test our theory through a pilot study during ICIS 2022. We omitted other influencing factors in this study also to keep the duration for participants in the pilot study at a low level (max. 10 minutes) while still acquiring the essential information, represented by the influence of data proximity and sensitivity to the intraindividual data value (Thabane et al. 2010).

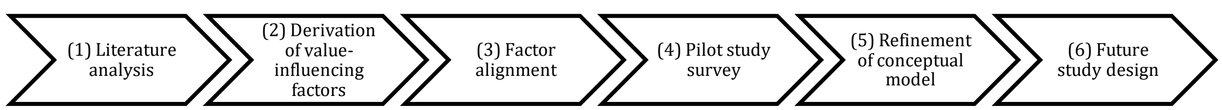

Figure 2. Research Process

For the pilot study, we developed an online survey to test the influence of the two factors on the intraindividual data value. We implemented the survey using the online software *Qualtrics Research Core*. First, we asked the participants to rank different data types, such as health, location, or education data, according to their perceived degree of proximity and sensitivity (see the survey design in the Appendix). Next, we asked the respondents to assign percentages of their monthly income reflecting the intraindividual value of high, medium, and low sensitivity, respectively high, medium, and low proximity data (see excerpt in Figure 3). If a participant wanted to indicate that they do not want to disclose their data or their data is priceless, they would assign a value of *999* for the percentage of the monthly income. In addition, we gathered demographic data as well as the risk propensity of participants.

For collecting research data, we opted for convenience sampling, an established approach for pilot studies (Bornstein et al. 2013; Patton 2015; Robinson 2014). We distributed the survey link and QR code at ICIS 2022 and within our research network via social media. The survey was accessible for approximately 72 hours (11 – 13 December 2022). As our survey responses contain ordinal data and our respondents' answers were measured repeatedly, we used the Friedman repeated measures analysis followed by Wilcoxon rank tests for data sensitivity and proximity to test the two hypotheses H_{VP} and H_{VS} (Pereira et al. 2015). The following section describes and discusses our preliminary results.

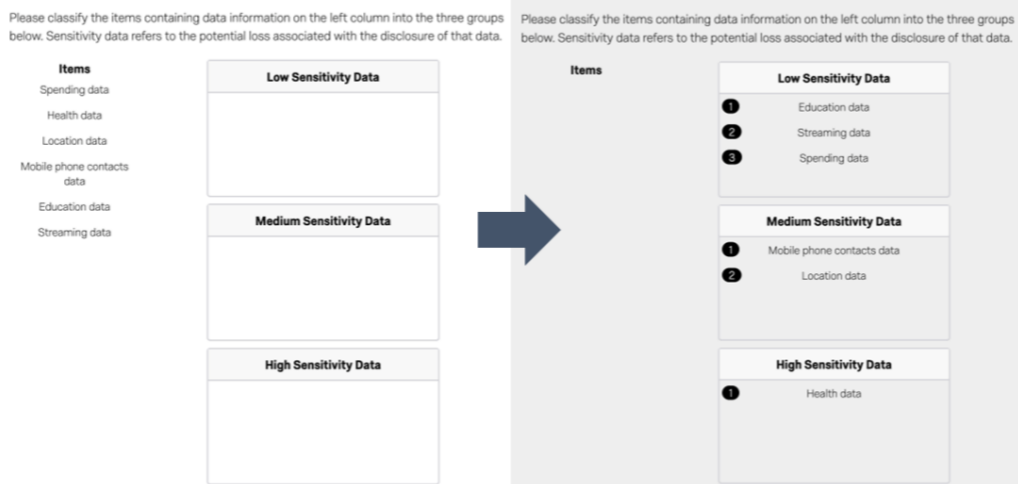

Figure 3. The blank survey (left) and a filled-in example (right) of question S1.

Pilot study Results and Analysis

We have collected 175 responses out of which 104 were complete and included in our analysis. In Table 1, we summarize the demographics of the respondents. Only 20% of them indicated to be female, and the rest identified as male. Our sample can be considered geographically biased, with many respondents living in Austria. The distribution over the age groups is more diverse than the other demographics variables, although skewed towards the youngest groups. In Tables 2, 3, and 4, we show the number of responses assigning the different levels of data sensitivity and proximity to data types and median data values (in percent of the personal monthly income) depending on the data sensitivity and proximity.

Gender	Responses	**Age group**	Responses	**Residence**	Responses
Male	83	20-25 years	25	Austria	41
Female	21	26-31 years	35	Germany	26
Non-binary	0	32-37 years	16	Belgium	20
Prefer not to say	0	38-43 years	11	USA	10
		44-49 years	5	Denmark	1
		50-55 years	4	Italy	1
		≥ 56 years	8	Netherlands	1

Table 1. The number of responses to the question of the respondent's gender, age, and country of residence.

The Friedman test for data sensitivity was conducted to determine whether individuals assign different percentage values of their salary when considering low-sensitivity, medium-sensitivity, and high-sensitivity data (Table 2). The results show a significant difference $\chi^2(2)= 133.161$, p=0.000. Thus, we reject the null hypothesis and conclude that there is a difference in how individuals value low-sensitivity, medium-sensitivity, and high-sensitivity data. Still, the Friedman test only tells us if there is a difference in how individuals perceive the value of low-, medium-, and high-sensitivity data. Therefore, we conducted two Wilcoxon rank tests to know how these sensitivity levels differ: one between low and medium sensitivity assigned values and another between medium and high sensitivity assigned values. A Wilcoxon signed-rank test showed that individuals assigned higher values for medium sensitivity than for low sensitivity data (*Z*

= -6.616, p = 0.000), the median for low sensitivity being 10 compared with the median sensitivity being 50 for medium sensitivity data. Another Wilcoxon signed-rank test showed that individuals assigned higher values for high sensitivity than for medium sensitivity data (Z = -6.811, p = 0.000), median for medium sensitivity being 50 in comparison with the median being 999 for high sensitivity data (Table 4).

Therefore, our H_{VS} hypothesis is supported, appearing that there is enough evidence to conclude that high data sensitivity is associated with a high intraindividual data value.

	Sensitivity		
Data type	Low	Medium	High
Health data	2	6	96
Location data	14	41	49
Spending data	18	38	48
Mobile phone contacts data	11	46	47
Education data	49	44	11
Streaming data	59	33	12

Table 2. The number of respondents assigning the data type to low, medium, or high sensitivity.

The Friedman test for data proximity was conducted to determine whether individuals assign different percentage values of their salary when considering low proximity, medium proximity, and high proximity data (Table 3). The results show a significant difference $\chi^2(2)$= 94.516, p=0.000. Thus, we reject the null hypothesis and conclude that there is a difference in how individuals value low-proximity, medium-proximity, and high-proximity data. As the Friedman test only tells us if there is a difference in how individuals perceive the value of low, medium, and high proximity data, we conducted two Wilcoxon rank tests to know how exactly these levels of proximity are different: one between low proximity and medium proximity assigned values and another between medium proximity and high proximity assigned values. A Wilcoxon signed-rank test showed that individuals assigned higher values for medium proximity than for low proximity data (Z = -5.101, p = 0.000), the median for low proximity being 20 compared to the median being 80 for medium proximity data. Another Wilcoxon signed-rank test showed that individuals assigned higher values for high proximity than for medium proximity data (Z = -6.045, p = 0.000), the median for medium proximity being 80 in comparison with the median being 999 for high proximity data (Table 4).

Therefore, our H_{PS} hypothesis is supported, appearing that there is enough evidence to conclude that high data proximity is associated with a high intraindividual data value.

	Proximity		
Data type	Low	Medium	High
Data about yourself	14	5	85
Data about your close family	12	9	83
Data about close friends	10	34	60
Data about relatives	14	46	44
Data about friends	10	70	24
Data about acquaintances	43	48	13
Data about people you don't know	85	7	12

Table 3. The number of respondents assigning the data type to a low, medium, or high proximity.

Since our statistical test results support our hypotheses, we suggest that individuals place more value on high-sensitivity and proximity data when compared with low-sensitivity and proximity data. A consequence can be that when commercializing their data from IoT devices, individuals would ask the buyers for a price depending on their data's perceived sensitivity and proximity level.

	Median data value (%)		
Antecedents	Low	Medium	High
Data Sensitivity	10	50	999
Data Proximity	20	80	999

Table 4. Median data value (in percent of monthly income) of all 104 responses depending on the level of data sensitivity and proximity.

Discussion

Limitations and Future Work

Our data and analysis support our proposition that data sensitivity and proximity influence the intraindividual value of data. Higher data sensitivity and proximity lead to higher data values perceived by individuals.

As with any research study, this one is not free of limitations. Our data set was composed of non-normally distributed data. We initially attempted to perform a repeated measure ANOVA on our dataset; because Mauchly's test of sphericity was significant, we resorted to a non-parametric repeated measures approach such as the Friedman test. Furthermore, we excluded the results of the risk propensity scale as they were insignificant. We included this, as some literature pointed out that risk propensity could influence willingness-to-disclose data and, consequently, intraindividual value perceptions (Fast and Schnurr 2020).

We tested only a part of an initially larger created model due to time constraints and access to convenient respondents. Our literature analysis suggests that these two concepts, data sensitivity and proximity, might not explain intraindividual data value entirely. Based on our findings, a prospective study should elaborate on how to include further influence factors on intraindividual data value beyond data sensitivity and proximity. First, the timeliness of data is expected to influence data value (Stein and Maass 2022). Timeliness represents the temporal distance since data generation and relates to when data are disclosed by the individual in construal level theory (Trope and Liberman 2010). From a data quality perspective, timeliness means that data are current or regularly updated (Batini et al. 2009; Pipino et al. 2002). Hence, more timely data might increase the intraindividual data value. Second, psychological ownership of data can affect how willing individuals are to share data. In the context of willingness-to-disclose data, high psychological ownership can relate to the fact that even low-sensitivity data might not be shared by individuals, e.g., in the context of IoT data (Cichy et al. 2014). Therefore, a future study must investigate how psychological ownership influences intraindividual data value.

Third, we focused on the monetary aspect of data transactions, i.e., financial incentives for disclosing data. Still, depending on the context of the data transaction, other incentives, such as getting back services for providing data or donating data, could play a role. Therefore, the type of incentive for data disclosure and altruism of individuals could, in addition, influence intraindividual data value perceptions (Gefen et al. 2020). Therefore, future work should include different data disclosure contexts and types of incentives. In addition, the data act coming into law in 2023 in the European Union, providing ownership rights for self-generated IoT data, could influence intraindividual data value.

Contributions

We contribute by conceptualizing intraindividual data value based on construal level theory. We elicited the factors of data proximity and data sensitivity as influencing factors. The statistical analysis of our study results indicates that data proximity and sensitivity influence intraindividual data value. Furthermore, we elaborated on additional influence factors to be included in future studies such as: psychological ownership, the timeliness of data, non-monetary incentives for individual data transactions, and the context of personal data disclosure. We aim to further develop a conceptual model including the aforementioned factors and extend the common knowledge on the intraindividual data value of IoT generated data.

Acknowledgments

All three authors thank Prof. Wai Fong Boh (NTU Singapore) for the great mentoring during the ICIS 2022 paper-a-thon and for guidance while developing this study. This work is part-funded by the German Federal Ministry for Economics and Climate Action (BMWK) within Smart Data Economy under grant number 01MD19010C (Project: Future Data Assets). In addition, author René H. Reich is grateful for the internal C2-funding of KU Leuven (Grant No: 3E210013), enabling the ICIS 2022 and paper-a-thon participation.

References

Batini, C., Cappiello, C., Francalanci, C., and Maurino, A. 2009. "Methodologies for Data Quality Assessment and Improvement," in *ACM Computing Surveys (CSUR)* (41:3), ACM New York, NY, USA, pp. 1–52.

Bauer, C., Korunovska, J., and Spiekermann, S. 2012. "On the Value of Information-What Facebook Users Are Willing to Pay," in *European Conference on Information Systems*.

Benndorf, V., and Normann, H.-T. 2018. "The Willingness to Sell Personal Data," *The Scandinavian Journal of Economics* (120:4), Wiley Online Library, pp. 1260–1278. (https://doi.org/10.1111/sjoe.12247).

Bornstein, M. H., Jager, J., and Putnick, D. L. 2013. "Sampling in Developmental Science: Situations, Shortcomings, Solutions, and Standards," *Developmental Review* (33:4), pp. 357–370. (https://doi.org/10.1016/j.dr.2013.08.003).

Cichy, P., Salge, T.-O., and Kohli, R. 2014. "Extending the Privacy Calculus: The Role of Psychological Ownership," in *International Conference on Information Systems*.

Dittmar, H. 1992. *The Social Psychology of Material Possessions: To Have Is to Be*, Palgrave MacMillan.

European Commission. 2020. "A European Strategy for Data. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions," No. COM/2020/66, , February 19. (https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52020DC0066).

European Union. 2022. *Proposal for a Regulation of the European Parliament and of the Council on Harmonised Rules on Fair Access to and Use of Data (Data Act)*, European Union. (https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022PC0068).

Farrelly, R., and Chew, E. 2016. "Who's in to Win?: Participation Rate in a Primary Personal Information Market," in *Australasian Conference on Information Systems*.

Fast, V., and Schnurr, D. 2020. "The Value of Personal Data: An Experimental Analysis of Data Types and Personal Antecedents," in *International Conference on Information Systems*.

Gefen, G., Ben-Porat, O., Tennenholtz, M., and Yom-Tov, E. 2020. "Privacy, Altruism, and Experience: Estimating the Perceived Value of Internet Data for Medical Uses," in *Companion Proceedings of the Web Conference 2020*, pp. 552–556. (https://doi.org/10.1145/3366424.3383414).

Geiregat, S. 2022. "The Data Act: Start of a New Era for Data Ownership?," *Available at SSRN*.

Gkatzelis, V., Aperjis, C., and Huberman, B. A. 2015. "Pricing Private Data," *Electronic Markets*, pp. 109–123.

Hubert, M., Carugati, A., Brock, C., and Obel, B. 2020. "Take It Personally–The Role of Consumers' Perceived Value of Personalization on Cross-Category Use in a Smart Home Ecosystem," in *Hawaii International Conference on System Sciences*.

Karampela, M., Grundstrom, C., and Isomursu, M. 2018. "Personal Health Data: Accessibility and Value in a Danish Context," in *International Conference on Information Systems Development*.

Krasnova, H., Eling, N., Abramova, O., and Buxmann, P. 2014. "Dangers of'Facebook Login'for Mobile Apps: Is There a Price Tag for Social Information?," in *International Conference on Information Systems*.

Lee, A. R., Hon, L., and Won, J. 2018. "Psychological Proximity as a Predictor of Participation in a Social Media Issue Campaign," *Computers in Human Behavior* (85), pp. 245–254. (https://doi.org/10.1016/j.chb.2018.04.006).

Li, C., Li, D. Y., Miklau, G., and Suciu, D. 2014. "A Theory of Pricing Private Data," in *ACM TODS*, pp. 1–28.

Li, X.-B., Liu, X., and Motiwalla, L. 2015. "Valuation of Personal Information," in *International Conference on Information Systems*.

Moon, Y. 2000. "Intimate Exchanges: Using Computers to Elicit Self-Disclosure from Consumers," *Journal of Consumer Research* (26:4), The University of Chicago Press, pp. 323–339.

Mothersbaugh, D. L., Foxx, W. K., Beatty, S. E., and Wang, S. 2012. "Disclosure Antecedents in an Online Service Context: The Role of Sensitivity of Information," *Journal of Service Research* (15:1), pp. 76–98. (https://doi.org/10.1177/1094670511424924).

Patton, M. Q. 2015. *Qualitative Research & Evaluation Methods: Integrating Theory and Practice*, (Fourth edition.), Thousand Oaks, California: SAGE Publications, Inc.

Pereira, D. G., Afonso, A., and Medeiros, F. M. 2015. "Overview of Friedman's Test and Post-Hoc Analysis," *Communications in Statistics - Simulation and Computation* (44:10), pp. 2636–2653. (https://doi.org/10.1080/03610918.2014.931971).

Pierce, J. L., Kostova, T., and Dirks, K. T. 2001. "Toward a Theory of Psychological Ownership in Organizations," *The Academy of Management Review* (26:2), pp. 298–310. (https://doi.org/10.2307/259124).

Pipino, L. L., Lee, Y. W., and Wang, R. Y. 2002. "Data Quality Assessment," *Communications of the ACM* (45:4), ACM New York, NY, USA, pp. 211–218.

Pu, Y., and Grossklags, J. 2015. "Using Conjoint Analysis to Investigate the Value of Interdependent Privacy in Social App Adoption Scenarios," in *International Conference on Information Systems*.

Robinson, O. C. 2014. "Sampling in Interview-Based Qualitative Research: A Theoretical and Practical Guide," *Qualitative Research in Psychology* (11:1), pp. 25–41. (https://doi.org/10.1080/14780887.2013.801543).

Schomakers, E.-M., Lidynia, C., Müllmann, D., and Ziefle, M. 2019. "Internet Users' Perceptions of Information Sensitivity – Insights from Germany," *International Journal of Information Management* (46), pp. 142–150. (https://doi.org/10.1016/j.ijinfomgt.2018.11.018).

Spiekermann, S., and Korunovska, J. 2017. "Towards a Value Theory for Personal Data," *Journal of Information Technology*, pp. 62–84.

Spiekermann, S., Korunovska, J., and Bauer, C. 2012. "Psychology of Ownership and Asset Defense: Why People Value Their Personal Information beyond Privacy," in *International Conference on Information Systems*.

Stein, H., and Maass, W. 2022. "Requirements for Data Valuation Methods," in *Hawaii International Conference on System Sciences*. (https://doi.org/10.24251/HICSS.2022.746).

Thabane, L., Ma, J., Chu, R., Cheng, J., Ismaila, A., Rios, L. P., Robson, R., Thabane, M., Giangregorio, L., and Goldsmith, C. H. 2010. "A Tutorial on Pilot Studies: The What, Why and How," *BMC Medical Research Methodology* (10:1), p. 1. (https://doi.org/10.1186/1471-2288-10-1).

Trope, Y., and Liberman, N. 2010. "Construal-Level Theory of Psychological Distance," *Psychological Review* (117:2), pp. 440–463. (https://doi.org/10.1037/a0018963).

Appendix

<table>
<tr><th>Survey block</th><th>Q#</th><th>Survey text, questions, and answers</th></tr>
<tr><td>Introduction</td><td></td><td>Thank you for participating in this survey at ICIS 2022! We are Hannah, René and Lucian and take part in the Paper-a-thon. We are performing a pilot study to gain information about individual value perceptions about personal data. You will need a maximum of 5 minutes to complete the survey. Feel free to share the survey with your peers!</td></tr>
<tr><td rowspan="2">Data sensitivity</td><td>S1</td><td>Please classify the items containing data information on the left column into the three groups below. Sensitivity data refers to the potential loss associated with the disclosure of that data.
<table>
<tr><td>Data items</td></tr>
<tr><td>Spending data</td></tr>
<tr><td>Health data</td></tr>
<tr><td>Location data</td></tr>
<tr><td>Mobile phone contacts data</td></tr>
<tr><td>Education data</td></tr>
<tr><td>Streaming data</td></tr>
</table>
<table>
<tr><td>Groups</td></tr>
<tr><td>Low Sensitivity Data</td></tr>
<tr><td>Medium Sensitivity Data</td></tr>
<tr><td>High Sensitivity Data</td></tr>
</table>
</td></tr>
<tr><td>S2</td><td>What percentage of your monthly income would you accept to sell the data classified in the previous question with a company willing to pay for your data? Please write 999 for not being willing to sell your data at all.
<table>
<tr><td></td><td>Percentage of your monthly income</td></tr>
<tr><td>Low Sensitivity Data</td><td></td></tr>
<tr><td>Medium Sensitivity Data</td><td></td></tr>
<tr><td>High Sensitivity Data</td><td></td></tr>
</table>
</td></tr>
<tr><td rowspan="2">Data proximity</td><td>P1</td><td>Please classify the items containing data information on the left column into the three groups below. Proximity is defined as the psychological distance between an individual and a dataset.
<table>
<tr><td>Data items</td></tr>
<tr><td>Data about yourself</td></tr>
<tr><td>Data about your close family</td></tr>
<tr><td>Data about relatives</td></tr>
<tr><td>Data about close friends</td></tr>
<tr><td>Data about acquaintances</td></tr>
<tr><td>Data about people you don't know</td></tr>
</table>
<table>
<tr><td>Groups</td></tr>
<tr><td>Low Proximity</td></tr>
<tr><td>Medium Proximity</td></tr>
<tr><td>High Proximity</td></tr>
</table>
</td></tr>
<tr><td>P2</td><td>What percentage of your monthly income would you accept to trade the information classified in the previous question with a company willing to pay for your data? Please write 999 for not being willing to sell your data at all.
<table>
<tr><td></td><td>Precentage of your monthly income</td></tr>
<tr><td>Low Proximity</td><td></td></tr>
<tr><td>Medium Proximity</td><td></td></tr>
<tr><td>High Proximity</td><td></td></tr>
</table>
</td></tr>
<tr><td></td><td>D1</td><td>What group of age are you?
○ 20 – 25 years old</td></tr>
</table>

Demographics		○ 26 – 31 years old ○ [...] ○ 62 – 67 years old ○ 68 and more years old
	D2	What is your gender? ○ Male ○ Female ○ Non-binary / third gender ○ Prefer not to say
	D3	What is your monthly salary range? ○ Under 1000$ ○ 1000 – 1999 $ ○ [...] ○ 14000 – 14999 $ ○ more than 15000$
	D4	In which country do you currently reside? ○ Afghanistan ○ [...] ○ Zimbabwe
	D5	Please indicate the extent to which you agree or disagree with the following statement by clicking the option you prefer. Please do not think too long before answering; usually your first inclination is also the best one. *7-point Likert scale of 'Strongly disagree', 'Disagree', 'Somewhat disagree', 'Neither agree nor disagree', 'Somewhat agree', 'Agree', and 'Strongly agree' of the following items.*
	D5.1	Safety first
	D5.2	I do not take risks with my health
	D5.3	I prefer to avoid risks
	D5.4	I take risks regularly
	D5.5	I really dislike not knowing what is gong to happen
	D5.6	I usually view risk as a challenge
	D5.7	I view myself as a risk avoider
End of survey		If you have any comments about our survey please write them here. We thank you a lot. In case of any questions please contact Hannah (hannah.stein@dfki.de). Please click on the arrow to submit the survey.

Table A. Online survey design