Data Anonymization as Instrument to manage Knowledge Risks in Supply Chains.

Johannes P. Zeiringer
University of Graz
johannes.zeiringer@uni-graz.at

Stefan Thalmann
University of Graz
stefan.thalmann@uni-graz.at

Jürgen Fleiss
University of Graz
juergen.fleiss@uni-graz.at

Abstract

In times of interconnected and digitalized supply chains (SCs), managing knowledge risks is challenging. As sharing data is associated to the risk of unintentional disclosure of competitive knowledge, SC partners must balance knowledge sharing and protection. However, knowledge risks can inhibit knowledge sharing and therefore harm the SC management as well as desired innovation. To address this problem, data anonymization can be a solution. Further, decision support how to use the data anonymization on data sets seems necessary. For this, an already developed data anonymization tool was used as basis for a vignette study with 1.000 participants, to investigate the effect of a decision support, in form of a tradeoff visualization, on knowledge sharing. The results showed that having an anonymization tool in place does increase knowledge sharing if also decision support is provided. This helps in making an individual decision easy and transparent, and, despite a high perception of risk, there is willingness to share data and it is also considered to be beneficial.

Keywords: knowledge risks, knowledge sharing, decision support, supply chain, vignette survey.

1. Introduction

Inter-organizational knowledge sharing requires balancing knowledge sharing and protection to manage knowledge risks (Loebbecke et al., 2016; Zeiringer, 2021). Supply chains (SCs) are one prominent example of interorganizational collaboration in which the data exchange continues to grow and the risk of unintentional knowledge loss increases (North et al., 2019; Zeiringer & Thalmann, 2020). However, despite all risks, knowledge sharing with SC-partners is vital to innovate, stay competitive and improve the SC performance (Zacharia et al., 2019).

Data exchanged in a SC collaboration such as quality data or production data in particular can be used to discover sensitive knowledge and is thus considered a knowledge risk (Kaiser et al., 2020). However, the evaluation of such (big) data sets exchanged within the SC is difficult for humans and suitable decision support seems necessary to balance risks and benefits of knowledge sharing (Ilvonen et al., 2018).

As risks related to decision making can be caused by misapplied knowledge management (Durst & Ferenhof, 2016), it is important to look at how data sharing is handled by individuals, and whether decision support leads to improved knowledge sharing. One approach to deal with knowledge risks in situations about sharing or protecting is data anonymization (Manhart et al., 2015). Data anonymization is common practice in managing privacy (Majeed & Lee, 2021) and can also be used to manage knowledge risks in industry (Kaiser et al., 2020). Data anonymization is the process of modifying the input data in such a way that sensitive attributes and relationships are not traceable for others (Murthy et al., 2019). It is especially useful because it protects sensitive personal, or machine-generated data while it still allows us to use the data for analysis purposes (Chicaiza et al., 2020). However, interpreting large data sets and deciding on a suitable anonymization is challenging. In this case, a visualization can help, as it is easy to process and understand for a human (Islam & Jin, 2019). Therefore, we tested if a data anonymization tool with a visualization is suitable to support decision makers in managing knowledge risks.

We conducted a vignette study to evaluate the impact of a data anonymization tool on the willingness to share data, respectively to manage knowledge risks. The trade-off visualization (showing utility of anonymized data and privacy after anonymization) was added to the data anonymization tools as decision support for users deciding on the anonymization task. We build upon previous research showing that trade-off visualization enhances the interpretation of humans and can lead to balanced decisions (Al-Kassab et al., 2014; Caldeira et al., 2019; Oprean et al., 2019). By this we want to answer the following research question:

What effect does a tradeoff visualization in a data anonymization tool have on knowledge sharing in a SC?

The results show that the proposed data anonymization tool does help employees in balancing

URI: https://hdl.handle.net/10125/107046
978-0-9981331-7-1

the sharing and protecting of knowledge if a tradeoff is visualized.

2.Hypotheses development and measures; data collection & procedure

2.1. Hypotheses development

Organizations engaged in SCs must balance knowledge sharing and protection and thus the risks as well as the benefits of knowledge sharing (Zeiringer & Thalmann, 2021). Knowledge sharing in a SC improves collaboration, performance and risk management (Baihaqi & Sohal, 2013; Zeiringer, Durst, & Thalmann, 2022). Hence, sharing is crucial for organizations to participate in SCs and rigorous knowledge protection can be harmful to innovation (Lee et al., 2017; Manhart et al., 2015; van Oorschot et al., 2018). Thus, decision makers need to carefully evaluate risks and benefits, which is a complex task in SC collaborations and demand tools to support decision-making (Colicchia et al., 2019).

Due to the complexity of data sets exchanged, the frequency of SC partners switching, and the number of decisions which need to be taken towards knowledge sharing, decision makers need guidance in order to take reasonable decisions (Enders et al., 2020; Lee et al., 2017; Pereira et al., 2013; Zeiringer & Thalmann, 2020). Furthermore, if knowledge risk cannot be assessed properly, partners are reluctant to share data and this can limit knowledge risk management as well (Zeiringer & Thalmann, 2021). In a SC, a restriction of data exchange could lead to a breakdown of the entire collaboration, and is not in the sense of a successful SC collaboration (Zeiringer, 2021). Having decision support in general is helpful for managing knowledge risk and can be designed in various forms, like a knowledge risk map, or a process model (Durst & Zieba, 2019; Ilvonen et al., 2015). Furthermore, data anonymization is useful, because it helps in fostering data exchange and simultaneously in preventing unauthorized data mining or data misuse (Hanna et al., 2020; Wu et al., 2022). To tackle data sharing reluctance, decision support tools can be helpful (Kaiser et al., 2020; Manhart et al., 2015) and the combination of a data anonymization tool with additional visualization seems a promising path (Zeiringer, Weber, & Thalmann, 2022). Overall, we refer to the technology acceptance model (TAM), which states that the use of technology improves user performance, as measured by the perceived benefit in our study (Davis et al., 1989).

In this study, we want to investigate the impact of a data anonymization tool on the willingness to share knowledge in a SC. We use three different scenarios and propose three different hypotheses in our research model. We want to show to what extent a tradeoff visualization (between privacy and utility) within a data anonymization tool influences knowledge sharing in a SC, compared to a data anonymization tool without visualization or even no data anonymization at all. The proposed model is shown in *Figure 1*. The dependency of knowledge sharing is evaluated in three different scenarios by the participants. Overall, we propose that knowledge sharing is increasing from S1 to S3.

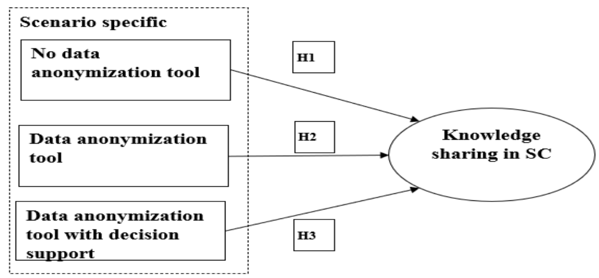

Figure 1: research model.

The first scenario (S1) revolves around decision making with neither data anonymization nor tradeoff visualization and shows plain data, see *Figure 2a*. While this approach may have benefits in terms of transparency and traceability, it could potentially lead to knowledge loss due to data analysis techniques applied by a SC partner (Zeiringer & Thalmann, 2020). Furthermore, lacking anonymization can lead to trust issues, data privacy and security issues, and therefore a high risk perception, which leads to reduced knowledge sharing in a SC (Enders et al., 2020; Goswami & Madan, 2017). We therefore propose:

H1: Willingness to share data is lowest in the scenario having no data anonymization tool.

Next, we pick up a data anonymization tool in the second scenario (S2), which anonymizes the production table, seen in *Figure 2a*. With the help of customized anonymization, it can be useful for protecting sensitive knowledge during the process of data sharing between SC partners, and the protection level depends on the chosen anonymization methods (Terrovitis et al., 2011; W. Yang & Qiao, 2010). The anonymized table of production data is illustrated in *Figure 2b*. By anonymizing certain attributes from datasets before they are shared, organizations can enhance sharing activities and mitigate potential knowledge risks simultaneously (Enders et al., 2020; Kaiser et al., 2020). For this purpose, perturbative, like noise addition or micro aggregation, and non-perturbative, like generalization or suppression, anonymization methods or completely synthetically generated data can be used (Drechsler, 2011; Hundepool, 2012). Concluding, we propose that by providing a tool with customized data anonymization, decision makers will rather engage in knowledge sharing within a SC collaboration than without any data anonymization. The hypothesis reads as follows:

H2: Willingness to share data is higher when a data anonymization tool is available compared to H1.

Last, we propose that an additional tradeoff visualization within a data anonymization tool can serve as decision support in the third scenario (S3). As even anonymized data can be assessed as vulnerable by decision makers, visualization approaches can be used to help make the exchanged data more comprehensible (Gashi et al., 2022; Sedayao et al., 2014). Scholars show that the tradeoff visualization of privacy and utility can influence the decision maker in regard to data sharing (Asikis & Pournaras, 2020). Furthermore, media richness theory argues that performance improves when factors like feedback, which can be in the form of visual decision support, come into play (Daft & Lengel, 1986). We try to give guidance and improve knowledge sharing by providing such a tradeoff visualization, see *Figure 2c*, and therefore propose the following hypothesis:

H3: Willingness to share data is highest when a data anonymization tool and visualization is available.

To test these hypotheses, we asked participants scenario-specific questions about their risk and benefit perceptions and their willingness to share the data, which subsequently leads to knowledge sharing. In addition, SC specific aspects were also asked. Thus, trust in partners, personal relationship with partners, and collaboration perceptions were elicited. Finally, the respondents were also asked about their trust in technology to check whether this has an influence.

2.2. Measures and data collection

A vignette study was used to determine the impact of a data anonymization tool on the willingness to share knowledge in a SC. This approach combines elements of traditional survey research with an experimental design. The method collects people's attitudes towards specific facts and, in the case of this study, helps to evaluate the willingness to share data. The main evaluation methods of the conducted vignette study are statistical tests, in which the vignette dimensions serve as independent variables and the assessment dimension serves as the dependent variable (Frings, 2010; Rost & Arnold, 2018). The vignette study allows to create fictive scenarios that capture the complexity and context of decision-making (Ghiselli & Ismail, 1999). Finally, vignette studies play a prominent an essential part in work on preferences in general (Kahneman & Tversky, 2013).

We build upon previous studies which show that knowledge sharing, e.g., in the SC, affects organizations and decision support is helpful (Zeiringer, Weber, & Thalmann, 2022). The authors therefore designed a vignette study, in which different scenarios are shown to participants. The scenario setting is in a SC that produces goods for wholesale. The modern production plant continuously collects data with sensors, which are exchanged with partners in the SC. There is the base scenario 1 (S1), where there is neither an anonymization tool nor a tradeoff visualization available at all, scenario 2 (S2), in which shared data is automatically previewed and can be anonymized in an anonymization tool, but without any assessment of the tradeoff, and scenario 3 (S3), where data can be anonymized and the tradeoff between privacy and utility can be visualized, before sharing. For each scenario, the respective participant evaluates his willingness to share data, and his risk and benefit of perception of data sharing.

Measures. For the questionnaire, established scales were consulted and adapted. The questionnaire consists of ten sets of questions. Regarding the questionnaire scales, it can be shown that the reliability is given, all necessary Cronbach’s alpha values are between 0.72 and 0.92 (see *supplementary data file, Table 2 & 3*).

Each participant was randomly assigned one of the three vignette scenarios, resulting in a between-subject design. We presented participants with only a single vignette to focus on that particular scenario in order to explore participants' decision-making processes in depth and in a controlled manner. In addition, presenting participants with multiple scenarios can generate order effects. First, the assigned scenario is presented. The participants were asked to think themselves in the randomly allocated vignette scenarios. The scenario was about data exchange with a SC partner (supplier, customer, contractor, etc.), and the decision-making process regarding sharing without and with data anonymization, or visualization. All questions, except for the demographics, are provided with a seven-point Likert scale. For each scenario, the participants evaluate their willingness to share data, their risk, and benefit perception, from which the overall knowledge sharing perception is deduced. The survey questions are described in the following and can be found in the *supplementary data file, Tables 2 & 3*.

Managing knowledge risks in SC interactions entails assessing risks and benefits and finding a suitable tradeoff (Ilvonen et al., 2018; Zeiringer & Thalmann, 2020). We propose that the sharing of knowledge in SC settings is closely tied to the willingness to share data, and thereby dependent on the perceived risk and perceived benefit that can be influenced by a data anonymization tool. We use the willingness to share data or knowledge as a dependent variable and define several independent variables in the following.

First, risk perception is used as an independent variable, which is defined as an uncertainty that results from a potential for a negative outcome (Havlena & DeSarbo, 1991). Furthermore, rapid technological change influences the individual risk perception (Acar & Göç, 2011) and knowledge risk perception might lack due to the incomprehensibility of these technological changes (Durst & Zieba, 2019). In the context of this study, the authors target the perception of risk in sharing data within a SC, which may arise from uncertainties related to cybercrime, espionage, or outdated technologies (Temel & Durst, 2020). If the perception of these possible risks is high, the authors assume that the willingness to share data and further knowledge is low.

However, collaborating with SC partners increases SC performance, enhances innovation capabilities and proper risk-benefit evaluation drives the disclosure of data, information, respectively knowledge (Barth & Jong, 2017; Heckmann et al., 2015; Un & Asakawa, 2015). Studies show that, e.g., efficiency is increased by data sharing and risk elimination prospers, when it is done jointly (Kumar & Pugazhendhi, 2012; Zeiringer, Durst, & Thalmann, 2022). The authors therefore propose that the willingness to share data is high if the benefit perception, as another independent variable, is high. Concluding, questions about *risk perception* and *willingness to share data*, based on the study from (Malhotra et al., 2004), were embedded. Additionally, a question of *benefit of the data anonymization tool*, *respectively visualization* was included, which was based on the study from (Holsapple et al., 2019).

Besides the effects of risk and benefit perception on knowledge sharing, we also investigate the effect of SC collaboration related independent variables (trust in partners, personal relationship and collaboration perception). Communication is a major component of personal relationships and, in addition, trust and commitment are the main drivers of collaboration outcomes (Qian et al., 2021; Sambasivan et al., 2011). Frequent communication and data sharing have a positive impact on the perception of collaboration within the SC (Qian et al., 2021). Trust increases the level of commitment in SC collaborations and decreases uncertainty, e.g., in data sharing, respectively knowledge sharing (Cai et al., 2013; Fawcett et al., 2017; Kwon & Suh, 2005). Furthermore, a trusted SC collaboration has a positive effect on innovation performance, e.g., for the joint development of new products (Patrucco et al., 2017; Zeiringer, Durst, & Thalmann, 2022). In data-centric collaborations, where tacit knowledge can be created by sensor-collected data, new trust issues emerge, because trust building may lack due to the fluctuating SC partners (Kaiser et al., 2020; Zeiringer & Thalmann, 2021). Having a close personal relationship and frequent activities within collaborations in the SC leads to more commitment and trust and to more willingness to share data (Du et al., 2012; Nyaga et al., 2010; J. Yang et al., 2012). Moreover, a trusting personal relationship leads to better knowledge sharing (Levin & Cross, 2004). Based on these findings, the authors used survey items from (Levin & Cross, 2004) for trust in partners and personal relationships. Furthermore, the collaboration perception was adapted from (Qian et al., 2021).

Another important independent variable to look at is trust in technology, which has a strong impact on the successful use of an information system, like a data anonymization tool in our case (Jian et al., 2000; Kivijärvi et al., 2013; Öksüz et al., 2016). We adopt the twelve-item scale, developed by (Jian et al., 2000), which measures trust in automated systems and is the most commonly used scale for this self-assessment. For the German questionnaire, the modification was made according to (Pöhler et al., 2016). The model applied measures the two factors trust, consisting of seven items, and distrust, which consists of five items.

Following the assigned vignette scenario, the participants had to answer general questions to assess their position towards trust in technology, trust in partners, personal relationships and collaboration perception. Finally, the participants were questioned about their company size, working experience in years, and their highest level of education. Overall, we use the results of these questions to gain insights into our three hypotheses. The whole questionnaire is shown in the *supplementary data file, Tables 2 & 3*.

Data collection & analysis. The vignette survey was conducted within an online questionnaire using *Limesurvey*. Pre-tests were conducted within the academic and professional environment to ensure participants would understand the vignettes correctly. A quota-representative sample, regarding education and working experience was recruited by the survey firm *Norstat* using an ISO 26362 certified online access panel in February 2023 in Austria and Germany. The target group for the survey were employees from the areas of, e.g., SC management, marketing & sales, accounting, R&D, respectively positions, where data sharing is needed.

1050 data sets were completed, which means that the participants were employed in the mentioned work areas and have not aborted the questionnaire in the meantime. Within this 1050 participants, outliers were detected and removed, which means very fast participants (mean duration of 67,6 seconds for the fastest 2,5%) or very slow participants (mean duration of 1343,8 seconds for the slowest 2,5%) were excluded. This leads to a total of 1000 subjects who on average took 366,54 seconds to complete the survey (SD = 205,11 seconds). All results presented are based on those 1000 subjects from Austria and Germany (S1: n = 362, S2: n = 329, S3: n = 309).

Likert scales that were negatively framed were recoded so that higher values consistently indicate agreement. Further additional independent control variables (cv) are sociodemographic nature and are measured in five-point scales from 0 to 4, in which 0 means “no entry”. For the statistical analyses we assume that the participant’s scoring can be treated as an interval scale. All participants were informed that they were presented predefined scenarios that are hypothetical, and the shown data was for illustration purposes, to stimulate their decision making.

2.3 Procedure

To explore the impact of a data anonymization tool on the willingness to share knowledge, an overarching scenario was developed first. The scenario setting is in a SC that produces goods for wholesale. The modern production plant continuously collects sensor data, like timestamps or temperatures, which are exchanged with partners in the SC. The exchange of data is agreed between the companies, as this drives the improvement of jointly developed goods and thus creates added value on both sides but can also lead to knowledge loss.

In this setting, we present one out of three tool scenarios to the participants and ask them to do their individual assessment towards knowledge sharing. In all three vignette scenarios (S1, S2, S3), the participants had an overview of the data to be shared, see *Figure 2a*. The all-inclusive scenario was S3, in which the data is shown previewed in an anonymization tool, with the option to customize the tradeoff between privacy and utility with the help of a visualization (*Figure 2c*). Depending on the chosen level of anonymization, the red circle moves in the direction of utility or privacy on the graph. Stepped down from this, in S2, the data is shown previewed in an anonymization tool, with the option of custom anonymization, but without any tradeoff visualization (*Figure 2b*). Lastly, in S1, only the table with production data was shown that cannot be anonymized, and no anonymization tool or tradeoff visualization is available (*Figure 2a*). The participants from S2&S3 get to see the difference between the anonymized and non-anonymized data set. All scenarios can be found in the *supplementary data file.*

Time	Charge	Injection pressure [bar]	Flow time [ms]	Dough temperature [°C]	Humidity [%]	Sugar content [%]	Fat content [%]
11:26:44	20230281	1117	1833	58,5127	59,7543	25,939	16,622
11:26:46	20230282	1316	1357	53,4091	56,5906	26,977	16,841
11:26:48	20230283	1496	1390	53,5949	50,8888	25,896	16,666
11:26:50	20230284	1035	1334	56,8474	60,1866	26,133	14,915
11:26:52	20230285	1433	1441	60,3385	51,3612	25,001	14,216
11:26:54	20230286	1125	1207	50,1308	60,9032	26,928	15,667
11:26:56	20230287	1256	1387	59,4220	56,0160	26,038	14,216
11:26:58	20230288	1458	1001	53,5143	57,1087	26,929	15,209

Figure 2a: production data in S1-3.

Data preview

Time	Charge	Injection pressure [bar]	Flow time [ms]	Dough temperature [°C]	Humidity [%]	Sugar content [%]	Fat content [%]
*	20230281	1000-1500	1000-2000	53	46	25	14
*	20230282	1000-1500	1000-2000	59	46	25	14
*	20230283	1000-1500	1000-2000	53	47	25	15
*	20230284	1000-1500	1000-2000	55	47	26	16
*	20230285	1000-1500	1000-2000	60	47	26	14
*	20230286	1000-1500	1000-2000	51	47	25	16
*	20230287	1000-1500	1000-2000	59	46	26	14
*	20230288	1000-1500	1000-2000	60	47	25	15

Edit Anonymize Reset

Figure 2b: data anonymization in S2 & S3.

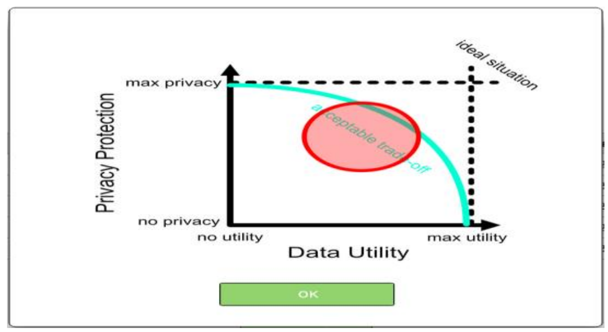

Figure 2c: tradeoff visualization in S3.

3. Results

A one-way ANOVA, conducted to test for differences in the willingness to share data between the three scenarios, indicates the presence of a significant effect of our scenarios on the willingness to share data ($F(2, 997) = 3.430$, $p = 0.033$). Surprisingly, we find the lowest willingness to share data in S2 with a mean value of 3.68 instead of S1 which has a mean value of 3.84. In line with our assumptions, we find the highest willingness to share data in S3 with a mean value of 4.01. This indicates that the presence of a data anonymization tool reduces the willingness to share data as compared to our baseline where the tool is absent. However, when combining the data anonymization tool with a tradeoff visualization, the willingness to share is increased.

To better understand the drivers of those differences, we also look at the effect of risk perception of data sharing. Interestingly, in S2, risk perception is again higher with a mean value of 4.749 as compared to compared to a mean value of 4.622 in S1. This indicates that participants perceive a higher risk for sharing data, when a data anonymization tool is available. We again find the lowest mean of 4.620 in S3, where the decision support in the form of a visualization is available. Thus, only if the visualization is available together with an anonymization tool, the willingness to share data increases and allows for the benefits of data sharing to be realized. Otherwise, the presence of only an anonymization tool increases risk perception and decreases the willingness to share data. As for the effect of risk perception on the willingness to share data, we assumed an increased risk perception will lead to a decreased willingness to share data. A one-way ANOVA indicates only weak evidence of an effect of our different scenarios on the willingness to share data ($F(2, 996) = 2.5845$, $p = 0.076$).

3.1. Multiple Regression Analysis

To test our hypotheses, stating that that the willingness to share increases from the base scenario S1 to the scenario including the anonymization tool S2 and again increases when combining the anonymization tool with a tradeoff visualization in S3, we estimate an OLS regression model including two scenario dummies for S2 and S3, respectively. S1 serves as our base category. Next, our SC related variables are included. We also add our control variables (company size, working experience in years, and their highest level of education), and found out that there is no significant influence on the willingness to share data. Further, we add interaction terms between both risk and benefit perceptions and the scenario dummies to investigate whether risk perception and benefit perception have different effects when an anonymization tool only (S2) and both the tool and tradeoff visualization are present (S3). A robustness check adding participants with the top and bottom 2,5% completion times leaves our results qualitatively unchanged.

An R2 of 0.37 together with a significant F-test indicates a good overall model fit, see *Table 1*. However, for the two scenario dummies, we find no significant effects, thus we cannot show that the willingness to share data increases in S2 and S3 compared to S1. Pairwise comparisons of all scenarios show significant differences between scenarios S2 and S3 (Tukey corrected p-value = 0.046). This supports the findings above that anonymization tools need to be complemented by a trade-off visualization.

For risk perception, we observe a negative regression coefficient of -0.575 ($p < 0.001$), indicating that as risk perception increases, the willingness to share data decreases. We find a significant and positive regression coefficient of 0.174 ($p = 0.046$) for the interaction term between risk perception and the S3 dummy variable only. Therefore, the net effect of risk perception is still negative in S3, but higher risk perception yields a lower reduction in the willingness to share data as compared to S1.

Benefit perception has a positive regression coefficient of 0.453 ($p < 0.001$). Thus, as the perceived benefits of sharing data increase, so does the willingness to share data. The interaction effect between benefit perception and S2 of -0.252 ($p = 0.013$) reveals that effect of benefit perception, while still positive, is significantly dampened in S2 when the anonymization tool only is present.

For our remaining variables, we find no significant effects for collaboration perception or the personal relation. For the trust in the supplier as well as trust in technology in general, we find significant positive effects with regression coefficients of 0.177 ($p = 0.001$) and 0.127 ($p = 0.037$), respectively.

	Model 1
Scenario = 2	0.901 (0.704)
Scenario = 3	-0.154 (0.695)
Risk Perception	-0.575** (0.060)
Risk Perception * S2	0.022 (0.087)
Risk Perception * S3	0.174* (0.087)
Benefit Perception	0.453** (0.077)
Benefit Perception * S2	-0.252* (0.101)
Benefit Perception * S3	-0.120 (0.104)
Personal Relationship	-0.003 (0.081)
Trust to Supplier	0.176** (0.050)
Collaboration perception	0.041 (0.051)
Trust in Technology	0.153* (0.073)
Intercept	2.994** (0.593)
Adj. R2	0.37
N	1000
F	49.46**

Table 1: OLS Regression Results.

4. Discussion

Our main contribution is that we deliver first evidence that a data anonymization tool with a tradeoff visualization increases the willingness to share data in a SC setting. This is interesting as a data anonymization tool without the visualization does not increase the willingness to share data, compared to a setting without data anonymization at all. Thus, it can be assumed that the visualization helps to balance risks and benefits of knowledge sharing and provides decision support.

Looking at the risk perception in S3 compared to S1, the willingness to share data rises, although the risk perception in data sharing also rises. This leads to the assumption that a tradeoff visualization offered to the participants does positively influence the willingness to share data. The benefit perception of data sharing, seen in the comparison of S2 with S1, is decreasing the willingness to share data. We interpret this in the way that having data anonymization without any tradeoff visualization is making decision makers rather more cautious regarding the willingness to share data. In addition, protective technologies, like an anonymization tool, can lead to less perceived ease of use, or less attributed competence to the tool as even anonymized data can be vulnerable (Dinev & Hu, 2007; Sedayao et al., 2014). We therefore propose that multiple concerns regarding anonymized data exchange can appear, like distorting the data leads to uselessness for decision-making or lacking reliability, and this in turn can finally cause knowledge risks.

In summary, it can be stated for *H1: Willingness to share data is lowest in the scenario having no data anonymization tool*, that the results do not confirm this. We assumed that willingness to share increases consistently from S1 to S3, however we find that in S2 willingness is lower than in S1. Based on this, we must reject the first and the second hypothesis.

Coming back to the different scenarios, there is a need for a comparison of having a tradeoff visualization and the pure data anonymization tool without visualization. There is a statistically significant difference between S2 and S3. Interestingly, S2 and S3 only differ in the additional tradeoff visualization, to enable customized balancing of privacy and utility, respectively protecting and sharing in the data anonymization tool. This indicates that this additional decision support, i.e., by providing a trade-off visualization leads to a better guidance of the users and enhances decision-making. In our concrete scenario setting, this means that production generated data like timestamps or temperature measurements, pose a challenge for people making decisions even if the data is in anonymized form. As decision support supports the decision maker, we propose that our tradeoff visualization is a possible way to give users a better, respectively more comprehensible decision-making basis. Especially because organizations that provide data perceive a risk of losing control over the shared data, including privacy concerns and a possible negative impact on the organization's competitive knowledge (Enders et al., 2020). Furthermore, people need to see that there is reciprocity of utility and privacy, and by sharing their data, both sides can benefit (Cropanzano et al., 2017). Having a tradeoff visualized as decision support improves human perception and provides more control over the knowledge to be shared (Al-Kassab et al., 2014). To conclude, we state that the results support H3: *Willingness to share data is highest when a data anonymization tool and visualization is available.*

Overall, it is observed that the more complex the exchanged data sets become, the more important visual preparation becomes for decision makers. To manage knowledge risks in the future, more research into tool design will be needed, as decision support is necessary for increasing knowledge sharing (Manhart & Thalmann, 2015). Deducing from this case, we show that data anonymization must be as transparent, respectively comprehensible as possible. Visualization can practically help to enhance knowledge sharing by converting complex data into understandable figures, and it can reduce knowledge risks by supporting informed decision making. Users need easy to understand guidance to help them in decision making, because, as we can see, a data anonymization tool in a SC does have a positive impact on knowledge sharing, but only if decision support in the form of a tradeoff visualization is provided.

5. Conclusion

In this study, we investigate the impact of a data anonymization tool on the willingness to share data in a SC. We focus on the risk of unintentional disclosure of competitive knowledge by sharing data. First results already indicated that decision makers need guidance and decision support (Zeiringer, Weber, & Thalmann, 2022). To investigate the phenomenon in a larger sample, a vignette study with 1000 participants was conducted to investigate if willingness to share data increases, if a data anonymization tool in combination with additional visualization is available. The results showed that if an anonymization tool is made available, additional decision support features such as a trade-off visualization are essential.

A practical implication of our study is that data anonymization tools, especially with decision support components, can have positive impacts on data and knowledge sharing. Organizations should investigate how they can support their workers to enhance their SC performance. Referring to the knowledge management literature, it can be stated that assessing data sharing upfront to evaluate possible knowledge risks is a highly relevant topic in which more research needs to be conducted.

This study comes with limitations. First, the survey participants are all from GER and AUT, nevertheless the variety of branches and the data set were comprehensive. Secondly, the trade-off visualization was fictive, and we only tested one way of decision support, other technical options can certainly lead to other conclusions and call for more research. Likewise, different types of data anonymization can lead to different results, but this was not the main investigation of this study since we focused on visual decision support. Thirdly, problems related to a vignette survey could always rise due to interpretation variability or missing contextual information beyond the scope of the scenarios, although we tried to communicate them as comprehensible as possible. The results can be evaluated in future studies and furthermore, the trade-off visualization is something that would be interesting to be developed and tested. This also highlights the importance of well-developed user guidance in decision making, closely linked to knowledge management. For knowledge management, protection strategies with a visualization tool perspective are calling for more research, as more data is generated, and more decision makers will apply low code to no code data science tools in the future.

References

Acar, E., & Göç, Y. (2011). Prediction of risk perception by owners' psychological traits in small building contractors. *Construction Management and Economics*, *29*(8), 841–852.

Al-Kassab, J., Ouertani, Z. M., Schiuma, G., & Neely, A. (2014). Information visualization to support management decisions. *International Journal of IT & Decision Making*, *13*(02), 407–428.

Asikis, T., & Pournaras, E. (2020). Optimization of privacy-utility trade-offs under informational self-determination. *Future Generation Computer Systems*, *109*, 488–499.

Baihaqi, I., & Sohal, A. S. (2013). The impact of information sharing in supply chains on organisational performance: an empirical study. *Production Planning & Control*, *24*(8-9), 743–758.

Barth, S., & Jong, M. D. de (2017). The privacy paradox – Investigating discrepancies between expressed privacy concerns and actual online behavior – A systematic literature review. *Telematics and Informatics*, *34*(7), 1038–1058.

Cai, S., Goh, M., Souza, R. de, & Li, G. (2013). Knowledge sharing in collaborative supply chains: twin effects of trust and power. *International Journal of Production Research*, *51*(7), 2060–2076.

Caldeira, C., Freire, F., Olivetti, E. A., Kirchain, R., & Dias, L. C. (2019). Analysis of cost-environmental trade-offs in biodiesel production incorporating waste feedstocks. *Journal of Cleaner Production*, *216*, 64–73.

Chicaiza, J., Cabrera-Loayza, M. C., Elizalde, R., & Piedra, N. (2020). Application of data anonymization in Learning Analytics. In *Proceedings of the 3rd International Conference on Applications of Intelligent Systems* (1–6). ACM.

Colicchia, C., Creazza, A., Noè, C., & Strozzi, F. (2019). Information sharing in supply chains: a review of risks and opportunities using the systematic literature network analysis. *SCM: An International Journal*, *24*(1), 5–21.

Cropanzano, R., Anthony, E. L., Daniels, S. R., & Hall, A. V. (2017). Social Exchange Theory: A Critical Review with Theoretical Remedies. *Academy of Management Annals*, *11*(1), 479–516.

Daft, R. L., & Lengel, R. H. (1986). Organizational Information Requirements, Media Richness and Structural Design. *Management Science*, *32*(5), 554–571.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. *Management Science*, *35*(8), 982–1003.

Dinev, T., & Hu, Q. (2007). The Centrality of Awareness in the Formation of User Behavioral Intention toward Protective Information Technologies. *Journal of the AIS*, *8*(7), 386–408.

Drechsler, J. (2011). Synthetic datasets for statistical disclosure control: Theory and implementation. *Lecture notes in statistics: Vol. 201*. Springer.

Du, T. C., Lai, V. S., Cheung, W., & Cui, X. (2012). Willingness to share information in a supply chain: A partnership-data-process perspective. *Information & Management*, *49*(2), 89–98.

Durst, S., & Ferenhof, H. A. (2016). Knowledge Risk Management in Turbulent Times. In K. North & G. Varvakis (Eds.), *Competitive Strategies for SMEs: Increasing Crisis Resilience, Agility and Innovation in Turbulent Times* (1st ed.,195–209). Springer.

Durst, S., & Zieba, M. (2019). Mapping knowledge risks: towards a better understanding of knowledge management. *KMRP*,*17*(1), 1–13.

Enders, T., Wolff, C., & Satzger, G. (Eds.) (2020). *Knowing What to Share: Selective Revealing in Open Data: ECIS, Marrakesch, Marokko*.

Fawcett, S. E., Jin, Y. H., Fawcett, A. M., & Magnan, G. (2017). I know it when I see it: the nature of trust, trustworthiness signals, and strategic trust construction. *The International*

Journal of Logistics Management, *28*(4), 914–938.
Frings, C. (2010). Das Messinstrument faktorieller Survey. In C. Frings (Ed.), *Soziales Vertrauen* (193–224).
Gashi, M., Mutlu, B., Lindstaedt, S., & Thalmann, S. (2022). Decision support for multi-component systems: visualizing interdependencies for predictive maintenance. In T. Bui (Ed.), *55th HICSS.*
Ghiselli, R., & Ismail, J. A. (1999). Promoting organizational effectiveness by defining managerial conduct. *International Journal of Contemporary Hospitality Management*.
Goswami, P., & Madan, S. (2017). Privacy preserving data publishing and data anonymization approaches: A review. In *2017 ICCCA* (139–142). IEEE.
Hanna, H., Haroun, M., & Gohar, N. (2020). Developing a framework for block chain adoption using TOE model. *Journal of Business & Retail Management Research*, *15*(01).
Havlena, W. J., & DeSarbo, W. S. (1991). On the Measurement of Perceived Consumer Risk. *Decision Sciences*, *22*(4), 927–939.
Heckmann, I., Comes, T., & Nickel, S. (2015). A critical review on supply chain risk – Definition, measure and modeling. *Omega*, *52*, 119–132.
Holsapple, C., Sena, M., & Wagner, W. (2019). The perceived success of ERP systems for decision support. *Information Technology and Management*, *20*(1), 1–7.
Hundepool, A. (2012). *Statistical disclosure control*. *Wiley series in survey methodology*. Wiley.
Ilvonen, I., Jussila, J. J., & Kärkkäinen, H. (2015). Towards a Business-Driven Process Model for Knowledge Security Risk Management. *International Journal of KM*, *11*(4), 1–18.
Ilvonen, I., Thalmann, S., Manhart, M., & Sillaber, C. (2018). Reconciling digital transformation and knowledge protection: a research agenda. *KMRP*, *16*(2), 235–244.
Islam, M., & Jin, S. (2019). An Overview of Data Visualization. In *2019 ICISCT* (1–7). IEEE.
Jian, J.-Y., Bisantz, A. M., & Drury, C. G. (2000). Foundations for an Empirically Determined Scale of Trust in Automated Systems. *International Journal of Cognitive Ergonomics*, *4*(1), 53–71.
Kahneman, D., & Tversky, A. (2013). Prospect Theory: An Analysis of Decision Under Risk. In L. C. MacLean & W. T. Ziemba (Eds.), *World Scientific Handbook in Financial Economics Series.* (Vol. 4, 99–127).
Kaiser, R., Thalmann, S., & Pammer-Schindler, V. (2020). An Investigation of Knowledge Protection Practices in Inter-organisational Collaboration. Protecting Specialised Engineering Knowledge with a Practice Based on Grey-box Modelling. *VINE.*
Kivijärvi, H., Leppänen, A., & Hallikainen, P. (2013). Antecedents of Information Technology Trust and the Effect of Trust on Perceived Performance Improvement. *International Journal of Social and Organizational Dynamics in IT*, *3*(3), 17–32.
Kumar, R. S., & Pugazhendhi, S. (2012). Information Sharing in Supply Chains: An Overview. *Procedia Engineering*, *38*, 2147–2154.
Kwon, I.-W. G., & Suh, T. (2005). Trust, commitment and relationships in supply chain management: a path analysis. *SCM: An International Journal*, *10*(1), 26–33.
Lee, J., Min, J., & Lee, H. (2017). Setting a knowledge boundary across teams: knowledge protection regulation for inter-team coordination and team performance. *Journal of KM*, *21*(2), 254–274.
Levin, D. Z., & Cross, R. (2004). The Strength of Weak Ties You Can Trust: The Mediating Role of Trust in Effective Knowledge Transfer. *Management Science*, *50*(11), 1477–1490.
Loebbecke, C., van Fenema, P., & Powell, P. (2016). Managing inter-organizational knowledge sharing. *The Journal of Strategic Information Systems*, *25*(1), 4–14.
Majeed, A., & Lee, S. (2021). Anonymization Techniques for Privacy Preserving Data Publishing: A Comprehensive Survey. *IEEE Access*, *9*, 8512–8545.
Malhotra, N. K., Kim, S. S., & Agarwal, J. (2004). Internet Users' Information Privacy Concerns (IUIPC): The Construct, the Scale, and a Causal Model. *Information Systems Research*, *15*(4), 336–355.
Manhart, M., & Thalmann, S. (2015). Protecting organizational knowledge: a structured literature review. *JKM*, *19*(2), 190–211.
Manhart, M., Thalmann, S., & Maier, R. (2015). The Ends of Knowledge Sharing in Networks: Using Information Technology to Start Knowledge Protection. In *Proceedings of the 23rd ECIS, Münster, Germany.*
Murthy, S., Abu Bakar, A., Abdul Rahim, F., & Ramli, R. (2019). A Comparative Study of Data Anonymization Techniques. In *5th International Conference on Big Data Security Proceedings* (306–309). IEEE.
North, K., Carvalho, A. de, Braccini, A., Durst, S., Carvalho, J., Gräslund, K., & Thalmann, S.

(2019). Information and knowledge risks in supply chain interactions of SMEs: Proceedings of the 10th International Conference on Practical Knowledge Management, Potsdam, Germany. *Lecture notes on Informatics.*

Nyaga, G. N., Whipple, J. M., & Lynch, D. F. (2010). Examining supply chain relationships: Do buyer and supplier perspectives on collaborative relationships differ? *Journal of Operations Management*, *28*(2), 101–114.

Öksüz, A., Walter, N., Distel, B., Räckers, M., & Becker, J. (2016). Trust in the Information Systems Discipline. In B. Blöbaum (Ed.), *Progress in IS. Trust and Communication in a Digitized World* (205–223). Springer International Publishing.

Oprean, D., Spence, C., Simpson, M., Miller Jr, R., Bansal, S., Keller, K., & Klippel, A. (2019). Human Interpretation of Trade-Off Diagrams in Multi-Objective Problems: Implications for Developing Interactive Decision Support Systems. In T. Bui (Ed.), *52nd HICSS*.

Patrucco, A. S., Luzzini, D., & Ronchi, S. (2017). Achieving innovation through supplier collaboration: the role of the purchasing interface. *Business Process Management Journal*, *23*(6), 1270–1289.

Pereira, L., Tenera, A., & Wemans, J. (2013). Insights on Individual's Risk Perception for Risk Assessment in Web-based Risk Management Tools. *Procedia Technology*, *9*, 886–892.

Pöhler, G., Heine, T., & Deml, B. (2016). Itemanalyse und Faktorstruktur eines Fragebogens zur Messung von Vertrauen im Umgang mit automatischen Systemen. *Zeitschrift Für Arbeitswissenschaft*, *70*(3), 151–160.

Qian, C., Seuring, S., Wagner, R., & Dion, P. A. (2021). Personal and organizational level relationships in relational exchanges in supply chains – a bottom-up model. *SCM: An International Journal*, *26*(1), 32–47.

Rost, K., & Arnold, N. (2018). *Die Vignettenanalyse in den Sozialwissenschaften: eine anwendungsorientierte Einführung*.

Sambasivan, M., Siew-Phaik, L., Abidin Mohamed, Z., & Choy Leong, Y. (2011). Impact of interdependence between supply chain partners on strategic alliance outcomes. *Management Decision*, *49*(4), 548–569.

Sedayao, J., Bhardwaj, R., & Gorade, N. (2014). Making Big Data, Privacy, and Anonymization Work Together in the Enterprise: Experiences and Issues. In *2014 International Congress on Big Data*, IEEE.

Temel, S., & Durst, S. (2020). Knowledge risk prevention strategies for handling new technological innovations in small businesses. *VINE*(Vol. 51 No. 4), 655–673.

Terrovitis, M., Mamoulis, N., & Kalnis, P. (2011). Local and global recoding methods for anonymizing set-valued data. *The VLDB Journal*, *20*(1), 83–106.

Un, C. A., & Asakawa, K. (2015). Types of R&D Collaborations and Process Innovation: The Benefit of Collaborating Upstream in the Knowledge Chain. *Journal of Product Innovation Management*, *32*(1), 138–153.

van Oorschot, K. E., Solli Sæther, H., & Karlsen, J. T. (2018). The knowledge protection paradox: imitation and innovation through knowledge sharing. *International Journal of Technology Management*, *78*(4), Article 10016826, 310.

Wu, C., Liu, J., & Zhang, H. (2022). Data Ecology and Accurate Portrait: Optimization of Credit Risk System for SMEs in Supply Chain Finance Based on Big Data Technology. *Journal of Risk Analysis and Crisis Response*, *11*(4).

Yang, J., Ko, I., Jang, H., & Cho, J. (2012). The Effects of Relationship Characteristics on Task Performance in Collaborative Environments Mediated by Information Systems. *45th HICSS.*

Yang, W., & Qiao, S. (2010). A novel anonymization algorithm: Privacy protection and knowledge preservation. *Expert Systems with Applications*, *37*(1), 756–766.

Zacharia, Z., Plasch, M., Mohan, U., & Gerschberger, M. (2019). The emerging role of coopetition within inter-firm relationships. *The International Journal of Logistics Management*, *30*(2), 414–437.

Zeiringer, J. P. (2021). Tackling Knowledge Risks in Data-centric Collaborations: A Literature Review. In *Proceedings of PACIS.*

Zeiringer, J. P., Durst, S., & Thalmann, S. (2022). Show Me What You Do and I Will Tell You Who You Are: A Cluster Typology of Supply Chain Risk Management in SMEs. *JTAECR*, *17*(1), 345–359.

Zeiringer, J. P., & Thalmann, S. (2020). Knowledge Risks in Digital Supply Chains: A Literature Review. *WI2020* (370–385).

Zeiringer, J. P., & Thalmann, S. (2021). Knowledge sharing and protection in data-centric collaborations: An exploratory study. *KMRP*, 1–13.

Zeiringer, J. P., Weber, M., & Thalmann, S. (2022). Design of a Data Anonymization Tool to Enhance Sharing on an Open Data Platform. In *35 th Bled eConference* (685–693).