Aug 10th, 12:00 AM

Data Sharing in the German Food Industry - Empirical Insights

Hannah Stein
German Research Center for Artificial Intelligence (DFKI), hannah.stein@dfki.de

Calvin Rix
Institute for Industrial Management (FIR) at RWTH Aachen University, calvin.rix@fir.rwth-aachen.de

Anna Effertz
Agricultural Market Information Company (AMI), anna.effertz@ami-informiert.de

Sven Bergau
Agricultural Market Information Company (AMI), sven.bergau@ami-informiert.de

Wolfgang Maass
German Research Center for Artificial Intelligence (DFKI), wolfgang.maass@dfki.de

Data Sharing in the German Food Industry – Empirical Insights

Completed Research

Hannah Stein
German Research Center for Artificial Intelligence (DFKI)
hannah.stein@dfki.de

Calvin Rix
Institute for Industrial Management (FIR) at RWTH Aachen University
calvin.rix@fir.rwth-aachen.de

Anna Effertz
Agricultural Market Information Company
anna.effertz@ami-informiert.de

Sven Bergau
Agricultural Market Information Company
sven.bergau@ami-informiert.de

Wolfgang Maass
German Research Center for Artificial Intelligence (DFKI)
wolfgang.maass@dfki.de

Abstract

Big data are collected along the entire food industry value chain, but remain mostly unused. Data sharing in data ecosystems could lead to efficiency gains and new revenue streams. We investigate data sharing within food industry and derive challenges and opportunities for data sharing in this context. We conducted interviews with ten qualified experts from the German food industry. The results reveal that mainly trust, usefulness and value influence users' attitude towards data sharing. Our results confirm social exchange theory in conjunction with technology acceptance model as relevant underlying IS theories of data sharing.

Keywords

Data Sharing, Food Industry, Data Value, Data Ecosystems.

Introduction

With 15.9 million employees in Europe (8% of total employment) and sales of 7 trillion dollars worldwide, food industry plays a socially and economically important role (Eurostat 2020; Frimpong 2021). Along the food production value chain, big data are collected casually (Tao et al. 2021). Using and analyzing them enables to accelerate productivity and realize business value for companies (Anand et al. 2016). Through digital transformation, data can become products themselves, e.g., through business analytics or artificial intelligence techniques, and form the basis of a data economy (KPMG 2016). A data economy represents a global digital ecosystem where data are collected, organized and exchanged by a network of providers to create value from accumulated data products (European Commission 2017). Data ecosystems consist of a series of networks of autonomous actors. These actors use, generate and make data and software available to other actors (Oliveira and Lóscio 2018). Data ecosystems significantly promote and support the sharing of data products. Data sharing in business-to-business context describes the process where a company makes data products available for free or with compensation to another company (Arnaut et al. 2018). Data products differ from raw data. They create economically usable images and representations of real situations that can be reused, shared with others, networked and traded via data markets. Interested parties are willing to pay a price for them (Shapiro et al. 1998). The value of data products increases through reuse and networking with other data products, so that local data views can also be turned into global data views with additional opportunities for exploitation (Maass 2009; Shapiro et al. 1998).

The following example, developed within a workshop together with food producers, business managers and data scientists, illustrates the need for and potential value added of data sharing through food industry data ecosystems: “Oil mills O_1, O_2, O_3 and grain mills G_1, G_2 provide anonymized, aggregated data products regarding their target production quantities for the next 12 months. Within a data ecosystem the data are aggregated and a data product for production volume planning on oil and grain for the next 12 months is generated. Through the data product, farmer F_1 can optimize the cultivation of his plots regarding the crop choice and volume to be grown and thus prevent over- and underproduction.” In this way, the data product, bought by the farmer, enables targeted planting and prevents monetary losses. The oil and grain mills create further value through the data product as they get paid for the data they provide. In turn, their raw material needs are met through targeted planting: underproduction is prevented. Data products in food industry show potential to (1) create new income sources, (2) prevent over- and underproduction, (3) enhance transparency in food production processes and (4) improve production planning and safety for both farmers and further processing plants (Tao et al. 2021). Data sharing within data ecosystems and the resulting increase in data utilization is expected to create up to 30 percent of the world’s gross domestic product (Lorenz 2018) and in turn increase involved companies’ revenue. Engagement in data ecosystems is regarded as necessity for future economic survival of organizations (Selander et al. 2013).

Until today, though data sharing and utilization could lead to increases in efficiency through internal and external data usage and generate new sources of income through the sale of data, the food industry is often reluctant towards data sharing (Durrant et al. 2022). Therefore, data value remains largely untapped. This research is guided by the following research questions: *Why is food industry reluctant towards data sharing? Under which conditions would food industry become more involved in data sharing?*

The paper is structured as follows: After this introductory section, we present foundations, especially on challenges of data sharing in context of data ecosystems. Next, the qualitative study design is described. The study results show data sharing issues and opportunities from provider and utilization perspective. We derive crucial theoretical concepts for implementing data sharing within food industry data ecosystems and discuss the results against the backdrop of IS theories before closing with a conclusion and future work.

Data Sharing in Food Industry Data Ecosystems

Data ecosystems are composed of complex organization networks or individuals that utilize and share data as main resource (Oliveira et al. 2019; Zuiderwijk et al. 2014). They provide an environment for creating, managing and sustaining data sharing initiatives (Oliveira et al. 2018) through technical platforms and infrastructures. Data sharing represents one of the central activities within data ecosystems. This means, efficient, secure data sharing is crucial for emergence and sustaining of data ecosystems. Data sharing in this context not only includes sharing of raw data, but also sharing of data products which have been prepared or present data analytics results. A systematic mapping study revealed further key benefits for data sharing in data ecosystems: Improvements in political, social and economic aspects, ease of data consumption and production, communication and interaction between actors, and improvements in data quality and services or new service and business model development (Oliveira et al. 2019).

Despite some successful cases of data sharing in food industry (Cappiello et al. 2020; Tao et al. 2021), large parts of food industry are reluctant towards data sharing and participation in data ecosystems (Durrant et al. 2022). This could arise from the rather young topic of data sharing in data ecosystems as well as from a vast of challenges, open questions and uncertainties in different aspects of data sharing in ecosystems through lacking in conceptual and theoretical basis (Oliveira et al. 2018). Mostly, organizations and their representatives fear unauthorized data access, missing legal certainty or regulatory lacks (Cappiello et al. 2020; Gasco-Hernandez et al. 2018; Lindner et al. 2021; Röhl et al. 2021). Furthermore, they fear absence of privacy, confidentiality or liability (Oliveira et al. 2019). Especially the legal perspective poses further issues: developing contracts for data sharing is more difficult than contracts for sharing physical assets (Zech 2017), as well as agreeing on legal and governance contracts (Gelhaar and Otto 2020). Unauthorized access and privacy considerations are often subject to fears of lacking data sovereignty or data security. These aspects must be guaranteed to enable data sharing (Gelhaar and Otto 2020) and can be supported in context of decentralized data ecosystems (Gelhaar et al. 2021; Maass 2022). Further challenges include high complexity of tasks such as data discovery and data consumption, lack of actor participation and interaction or organizational structure (Gasco-Hernandez et al. 2018; Oliveira et al. 2019), , as well as interoperability through standards (Cappiello et al. 2020; Gelhaar and Otto 2020). This interaction is often thwarted through challenges such as knowledge protection and jointly establishing standards with competitors (Gelhaar et al. 2021). Beneath legal challenges as mentioned above, often trust, representing a key factor for data sharing success, is missing or insufficient between actors (Cappiello et al. 2020; Gelhaar and Otto 2020). Furthermore, interaction of actors in terms of data sharing requires the demonstration of benefits for all actors (Gelhaar and Otto 2020). A lack of data monetization competency (Lindner et al. 2021; Röhl et al. 2021) and issues in data discovery or utilization (Oliveira et al. 2019) lead to restraint towards data sharing. Despite a larger number of papers on computational pricing and pricing mechanisms compared to topics such as security and privacy (Abbas et al. 2021), the actual value of data from both utilization and provider perspective in business context seems to be unclear to potential ecosystem participants.

To extend existing knowledge, we designed a qualitative expert interview study, presented in the following section. The foundations presented provide a starting point for further investigation of data sharing in data ecosystems and build the basis for our questionnaire development.

Study Design

We adopted the method of expert interviews using a semi-structured questionnaire to benefit from interviewees' diverse backgrounds and to understand their point of view as well as possible (Myers and Newman 2007). The semi-structured design allowed to deviate from the questionnaire, pose in-depth questions and adapt question wording to keep the discussion flowing (Döring and Bortz 2016). The final questionnaire comprised twelve questions (Gläser and Laudel 2009) including introductory questions, questions on data sharing from the utilization and provider perspective, and questions on future trends. We built on literature presented above, research of the author team and informal discussions with practitioners from food industry and manufacturing. Nine questions were phrased as open questions. Within three questions, the interviewees rated, e.g. the importance of data product features on a 5-point-Likert scale (Likert 1932). Still, they explained their choices and open follow-up questions investigated further aspects.

For expert selection, we considered two main criteria. 1) Diversity of backgrounds, i.e., representatives of different value chain stages in food industry and 2) depth of individual experience to maximize insights (Tamm et al. 2020). We required at least five years of experience in food industry and prior knowledge on data use and value in the respective business contexts. Ten experts from ten different companies in German food industry were acquired. These included four experts in food production, four from food wholesalers, one from food machinery and plant engineering, and one from a consulting and service company in food production. All of them possess basic or good knowledge on data utilization, data value, and data trading or sharing. The experts' responsibilities span from data science, sales, and management to food science and food services. They worked in food industry for an average of 10 years - all of the experts have excellent knowledge in their domain of food production. The experts and companies wish to stay anonymous.

From April 13 to May 28, 2021, ten individual expert interviews were conducted in German via Microsoft Teams and phone with an average duration of 55 minutes. The interviews were guided by two co-authors. They acted as co-experts for exchanging at higher professional level than would be possible with a layperson (Bogner et al. 2014). One author posed the questions, interacted with the expert and took a few notes on most important points. The other one took extensive notes. After each interview, the notes were expanded in an iterative process to capture all relevant information and to create a verbatim transcript. The raw, transcribed, anonymized interview data were transferred to a third author of this work for coding. To this end, we chose a semiotics approach, more specifically a content analysis approach. This mode of analysis deals with the meaning of signs and symbols in language (Krippendorff 2018). Words are assigned to primary conceptual categories. In this content analysis approach, references were made from transcripts to their contexts plus structures and regularities were searched for. In our application of content analysis, one author adopted thematic coding (Gibbs 2007), using a mixture of a priori and emergent coding. The codes emerged from raw, transcribed interview data. To ensure covering the full data set in the synthesis of findings, data from all transcripts were combined. The following steps were adopted: 1) *descriptive coding*: summarizing or labeling relevant sentences or phrases in a few words, 2) *categorical coding*: combining descriptive codes having things in common to categories, 3) *analytic coding*: examining connections between categories, 4) *alignment with IS theories*: comparing coding results with IS theories with similar concepts and refining coding results with respect to according theories, 5) *discussion process*: three workshop sessions with the author team for iterative results refinement. We took countermeasures to avoid the five threats to validity (Maxwell 2012). Descriptive validity relates to missing important interview information. Two authors participated in the interviews and created the transcript to avoid this. Interpretation validity concerns misunderstandings during the interview. These were conducted in interviewers' and respondents' mother tongue and the study goal was explained to participants prior to the interview. Definitions of special terms were prepared and shared with the interviewees during interview. The interviewers properly prepared by deeply studying the interviewees' backgrounds. Researcher bias and theory validity refer to the possibility that researchers interpret interview results in ways that serve their goals or confirm their original theory. At the time of the interviews, we were interested in a general overview of the food industry's attitude towards data sharing, and were much open to any outcome. Reactivity means that interviewees behave differently because of the interviewer's presence. Due to our research design, it is impossible to avoid reactivity. Still, we are aware of it and how it might influence what is being observed.

Study Results

Most experts see potential benefits in sharing or trading data through data ecosystems. Five main benefits emerged from the analysis: (1) support or increased simplicity of data sharing in a data ecosystem, e.g. in terms of customer acquisition, pricing, valuation or handling of sales; (2) new income sources arising through extension or development of new business models, e.g., through selling of data; (3) better data: with data ecosystems, more extensive data sharing becomes possible and network effects, such as the combination and aggregation of numerous data across food industry or further industries, leads to better insights and increased knowledge; (4) service availability is increased: services can be acquired in data ecosystem app stores to support data analysis. Thereby, data sharing network effects can improve services and algorithms; (5) process optimization: through better data, service availability and simplicity, experts expect that processes can be optimized, e.g., in food supply chains or production processes. Still, numerous requirements and issues towards data sharing exist, which will be presented in the following section. We present the results of the study in three parts, starting with the utilization and provider perspectives and then deriving the concepts for implantation of data sharing within food industry data ecosystems.

Data Sharing: Utilization Perspective

Eight of ten experts use free or chargeable third-party data products. These originate from market research, address data, promotion evaluations or sales, supply chain or controlling data. They primarily use them for local optimization of internal processes. The two not using or buying data products indicate that relevant data products are unavailable. The industry's heterogeneity aggravates valuable data product creation.

Feature	**Mean**	**std. dev.**
Currency of data product	4.3	0.82
Support in decision making	4.2	1.03
Customer service	4.0	0.47
Degree of accuracy	4.0	0.67
Usage rights	4.0	1.0
Value enhancement for my company	3.9	1.29
Continuity	3.9	0.99
Additional services, e.g., visualization	3.7	1.16
Reputation of provider	3.4	1.17
Exclusivity of data product	2.0	1.25

Table 1. Aggregated answers to the question "Which features of data products are important for you from a buyer perspective?" (1 = not important, 5 = very important)

We asked the experts about how important they consider specific data product features are from buyer perspective a 5-point Likert scale. Table 1 shows the features from most to least important as well as mean and standard deviation (std. dev.). The experts rate currency of data products as most important. One expert said that "we don't buy outdated data (...). For example: models that are based on pre-Covid-19 are, at least in most cases, no longer meaningful". Data products should improve and support decision making – "moving away from gut decisions to data-driven decisions", as one expert states. While other experts rank this factor as relevant, they point out that one should not rely entirely on the data obtained when making decisions. The feature "customer service" is synonymous with expertise and experience of the provider and was rated as important, just like the degree of accuracy of data products. Further crucial features are value enhancement for the company and continuity. The experts rated exclusivity, i.e., uniqueness of data products, as rather unimportant. They give mainly two reasons: First, having only one or a few actors who are in a position to buy a relevant data product contradicts the idea of digitization in food industry – data should be available along the full supply chain. Second, they assume that exclusive data products are so expensive, that costs exceed benefits. The experts were asked to point out which additional data product features would be very important to them. Their answers mainly target to the realm of standards, reputation and transparency: they require standards for data product formats to ensure usability and integratability to their existing systems, or to be better prepared for using obtained data products. Surprisingly, provider reputation is only rated as slightly important, which is contradictory with the importance they ascribe to customer service and the additional features mentioned. In their view, reputation of data products themselves is highly important. They distinguish between reputation of data products and its provider. High reputation data products contain high data quality, are relevant and valid to the individual buyer business. In addition, they need to be credible and, in consequence, trustworthy. Transparency of data products means that the general composition of data products should be traceable: data products should provide information on their history and a description of the raw data from which they are composed.

Beneath specific features, we distinguish four types of data products: data product as report, diagnosis, prediction and performance (Rix et al. 2021). We asked the experts how important these data product types are from a buyer's perspective. Table 2 shows the results giving mean, standard deviation (std. dev.) as well as a definition for each data product type. During the interview, we displayed the definition of each data product type and answered queries from the interviewees especially on data product as performance.

Data Product Type	Definition	Mean	std. dev.
Data product as report	Descriptive data product offering information on the present and past; statistical methods for data product	3.4	1.08
Data product as diagnosis	Diagnostic data product; describes information about the past and explains why certain events occurred.	3.6	1.35
Data product as prediction	Predictive data product; correlations and patterns are determined based on aggregated historical data; on this basis, future developments are predicted.	4.6	0.70
Data product as performance	Prescriptive data product; AI methods for predictions and decision making proposals; Value and price of data product determined ex-post based on increased sales.	3.8	1.16

Table 2. Aggregated answers to the question "How important are the following data product types to you?" (1 = not important, 5 = very important)

While data product as report is already implemented at local level by several of the experts, it becomes clear that it constitutes a basis for other data product types. This applies to data product as diagnosis as well. The experts require more valuable information when buying a data product. Still, potential network effects of these data product types are neglected. The interviewees rate data product as prediction as very important offer. These types generate most value with acceptable cost-benefit ratios according to the experts. They need data products that improve market development understanding, future customer needs, and their companies' current and future market positions, including explanations of data product results. The importance of data products as performance was rated as less important. This is due to two facts: First, the pricing approach, as they fear that pricing of data products after trading will lead to an imbalance if the buyer does not generate the expected value and the provider receives less payment. Second, the decision-making proposals. They argue that decisions need to be made company internally, e.g., based on a data product as prediction. With a data product as performance, which provides decision proposals, they fear that deciders might fully rely on such automated decision aids that could be incorrect, e.g., due to missing domain knowledge. Therefore, the experts currently would prefer to prepare their decisions based on data products as prediction to keep the risk with the decision makers in companies. The following section presents the interview results in the area of data sharing from the provider perspective.

Data Sharing: Provider Perspective

Five of ten experts share or sell data products to third parties, e.g., on sales and inventories, production, or trend monitoring. The five experts whose companies do not sell or share data products state several reasons. One main reason is that data considered valuable for sharing is often sensitive. Due to this and data security reservations, they refrain from data sharing. They also fear loss of knowledge-based competitive advantage or unique selling points. The experts worry that their internal data would be too small or irrelevant and thus of little for others. They indicate that higher value could be achieved by combining several data products. One mentions, since it is not part of their business model, the company did not consider sharing data. Still, nine of ten interviewees hold data products suitable for sharing. These include for example historical price, trend- or performance-monitoring, production, customer, or supply chain data products. We asked the experts to rate how important they consider issues in sharing data products on a 5-point Likert scale. Table 3 presents initial results. “Feature” represents the respective issue, mean and standard deviation are given.

Feature	**Mean**	**std. dev.**
Lack of standards	4.1	1.29
Costs exceed benefits	3.9	0.99
Lack of valuation and pricing approaches	3.9	1.29
Danger of disclosing business secrets to competitors	3.7	1.49
Lack of data security	3.6	1.17
Lack of data sovereignty	3.3	1.70
Lack of knowledge on request	3.2	1.03
Lack of knowledge on internal data and information management	3.1	1.20
Lack of legal certainty regarding use and further processing of data products	2.9	1.29
Investment costs for digital technologies	2.6	1.27
State of digitization in company insufficient	2.2	1.23

Table 3. Aggregated answers to the question “How important do you rate the following hurdles towards selling data products?” (1 = not important, 5 = very important)

The results show main issues concerning standards, value and pricing, as well as security and sovereignty, which corresponds with further analysis results. Lacking standards of data (products) with respect to format and quality is rated as most the important barrier to participation in data sharing (mean = 4.1). One expert stated that this is a major challenge, as standards are difficult to enforce, especially in international relations where standards are very different between countries. As companies process data in different formats, several experts point out that data ecosystems or marketplaces should set these standards for data sharing, e.g., through partnering with standardization institutes. Data value and pricing models in data sharing are often unclear. Some experts state that missing reference values and prices lead to uncertainties. This is also a reason for the great importance the experts placed on the problem that costs in data sharing could exceed the benefits. In addition, the danger of disclosing business secrets to competitors, data security, as well as lack of data sovereignty are rated between “rather important” and “important”. In this context, interviewees rate the danger of disclosing business secrets to competitors as most important (mean = 3.7). However, they rate a possible lack of data sovereignty only as “rather important” (mean = 3.3). For fear of disclosing business secrets, the interviewees would rather sell less sensitive data or data from which no conclusions can be drawn about their origin. Regarding data security, the experts fall into two camps: they either rate this as very important and difficult to implement, or as less important. The latter argue that data security is always a topic, but not a big issue due to existing technologies. The experts rate lack of legal certainty regarding use and further processing of data products as less important. They point out that this simply needs to be addressed contractually. Three experts state that many gray areas exist from legal perspective. One expert, which rated this aspect as not important at all states that "this must be implemented by the platform in advance", so he expects legal issues to be resolved before data sharing takes place, e.g., through a data ecosystem provider. Investment costs for digital technologies (mean = 2.6) and insufficient state of digitization (mean = 2.2) are rated as less important. The companies consider their digitization state as sufficient and do not expect investment costs for digital technologies as critical issues for data sharing.

We asked the experts whether they expect further issues in data sharing. Again, they emphasize the issue of sensitive data being accessed by third parties. Furthermore, they mention integration with existing and development of new business models as potential issue. For some experts, data sharing is not part of their company's current business model. They require further information on potential benefits before joining data sharing. In particular, this is due to the fact that entering data sharing ecosystems requires resources, especially to be able to scale the new business models. Another of the highlighted aspect is the importance of standardization, e.g., of data formats. The experts also point to the importance of automation degree, e.g., uploading and offering data should not require hours of data cleansing but follow a clear, semi-automized process. Still, they want to avoid full automation to prevent third parties from unauthorized access or from proceeding with actions not intended by the seller. They want to keep control of their data, which in turn underscores the importance of data sovereignty.

Concepts for Realizing Data Sharing within Food Industry Data Ecosystems

Based on challenges and opportunities of data sharing, extracted through the coding process, we matched the results with concepts from existing information systems (IS) theories. We found three major concepts relevant for data sharing in data ecosystems, arising from empirical results in accordance with the technology acceptance model (TAM) and social exchange theory. The concepts are *usefulness, trust* and *value*, which have a major influence on data sharing within data ecosystems. *Trust* hereby involves different factors, including data security, data sovereignty, trust between all actors in the ecosystem, as well as authentication and anonymity. *Usefulness* is considered in a broader sense: the concept includes general user interface aspects (usability) and technical feasibility (functionality) for data sharing, e.g., through data product and API standards within ecosystems. *Value* also encompasses a variety of factors, ranging from value enhancement and new revenue streams, to lack of understanding - and needed support - in creating value through data use and support in calculating the value and prices of data products, to new revenue streams through data ecosystems. Building on the IS theories mentioned and our empirical results, we developed a conceptual model representing underlying theoretical concepts of data sharing in data ecosystems (cf. Figure 1).

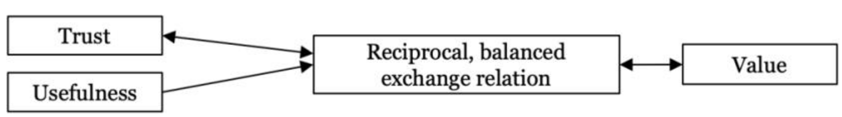

Figure 1: Underlying Concepts of Data Sharing Field in Data Ecosystems

The concepts influence each other, which is represented by the directed connections. We derived these directions based on our results in accordance with the above-mentioned theories. We thereby confirm social exchange theory in conjunction with TAM (cf. (Cook 1977; Kankanhalli et al. 2005) in the context of data sharing. Data sharing in data ecosystems is realized through *reciprocal, balanced exchange relations* (cf. (Cook 1977). We show this concept to link trust, usefulness and value in the context of data sharing. *Usefulness* represents a pre-condition for reciprocal, balanced exchange relations in the context of data sharing. *Trust* influences the general emergence of data sharing practices in terms of a reciprocal, balanced exchange relation. Once realized, a *reciprocal, balanced exchange relation* generates *value* for the involved actors. Through *value* generation, the exchange relation is in turn strengthened, if all actors consider this value to be fairly distributed. Building on this "proof" for *reciprocal, balanced exchange relations, trust* is in consequence enhanced. We discuss the conceptual model derived in the next section in more detail.

Discussion

Data sharing builds on interorganizational exchange relations, anchored within the social exchange theory (Cook 1977). Exchange relations are interactive and exist between at least two parties based on reciprocal reinforcement; exchange relations represent a series of transactions (Emerson 1972). A reciprocal, balanced exchange relation in the context of data sharing means that all stakeholders contribute something of equal value within the exchange process. A balanced exchange relationship means that actors within the setting have equal power or equal levels of dependence. For example, in an unbalanced relationship, the exchange ratio changes in favor of the actor who has a power advantage. In context of emergence of data sharing in data ecosystems, exchange relations should be reciprocal and balanced to create nearly equivalent perceived added-value for the participating actors. This is supported by our empirical results. The experts often describe the reciprocal, balanced exchange relation as "the need of a win-win" situation for all actors within data sharing. Furthermore, such exchange relations influence the level of trust between all actors involved.

Trust represents the willingness of one party to rely on another party (Söllner et al. 2016) and plays a major role in data sharing, which requires a trustworthy and secure environment (Gelhaar and Otto 2020). For data sharing in food industry data ecosystems, trust-based relationships between organizations are crucial. One study showed that trust represents a crucial antecedent within strategic information flows in inter-firm logistics relationships (Klein and Rai 2009). Strong trust implies a general belief in good intent of the other parties. Actors in data ecosystems trusting each other are likely to believe that recipients of their data products will not misuse them (Kankanhalli et al. 2005; Putnam 1993). Within data sharing, trust supports to create reciprocal, balanced exchange relations. Furthermore, the occurrence of these relations, which generate value for each party, is capable to influence trust – positively and negatively. If shared data are used by the receiving party as agreed, e.g. within an electronic contract (cf. (Maass 2022)) and without misuse, trust can grow. In case of misuse, the occurrence of which is feared by most of the experts, trust might decrease dramatically and potentially lead to one or more actors quitting data sharing.

Usability contributes to the overall functionality of a system by making it accessible to users and facilitating effective use of its functional capabilities. This is anchored in engineering and designing software (Goodwin 1987). TAM goes further by providing scales to measure and validate perceived usefulness and ease of use to make statements about user acceptance of a system (Davis 1989). Our results imply that a high perceived usefulness and ease of use, supported by high usability and functionality, influences the willingness of actors to participate in data sharing. For the experts, user interface design and functionality are important aspects, and they place value on the implementation of common, cross-national data formats, data product descriptions and API standards as prerequisites to get involved in data sharing. Similar to the work of (Kankanhalli et al. 2005), the aspect of technology acceptance, which also includes usefulness and usability aspects, may partially explain attitudes towards data sharing, i.e., willingness to join data sharing. Still, we did not identify an influence from reciprocal, balanced exchange relations back to the aspect of usefulness (cf. Figure 1). This means, usefulness here represents an enabler for the general emergence of data sharing.

Studies point to the fact that interorganizational information and knowledge sharing has a positive correlation with business value (Mandrella et al. 2020). Based on our results, data sharing, i.e., a reciprocal, balanced exchange relationship, should lead to business value for all actors involved. Value thereby can take different forms, i.e., monetary or business value. Coming back to the data sharing example (cf. Introduction) where oil mills, grain mills and farmer participate as actors in data sharing within a data ecosystem. Here, value is represented by (1) monetary value, i.e., oil and grain mills earn money for providing their data and (2) business value for the farmer, as they can optimize his planting activities and prevent losses, as well as (3) business value for the oil and grain mills, as they obtain security of supply from the farmer, thanks to optimized cultivation planning and the resulting output of raw materials based on their data. Further aspects of business value include increase in efficiency, speed, reliability or efficiency. Still, the generation of business value through data products is tied to resulting utilization concepts, as data value is dependent on data usage (Stein and Maass 2022). Usage possibilities, i.e., data-driven business model opportunities need to be made clearer to actors in data sharing, as this will have a major influence on the amount of value generated. Furthermore, they need support with calculating potential monetary values based on the utilization of data products within data-driven business models. The collaborative generation of value within data sharing influences reciprocal, balanced exchange relationships, if perceived as equal by the actors. In consequence, this rising stability of the exchange relationship increases trust.

Conclusion and Future Work

This paper is subject to several limitations which need to be overcome by future work. First, we conducted interviews with ten qualified experts in the German food industry with knowledge of data analytics and data sharing practices. Food supply chains are complex, diverse, and transnational. Even though the interviews with German experts represent a solid starting point, further studies need to include experts from various countries to draw a more complete picture of expectations towards data sharing in the food industry. Also, extending studies to further industries such as manufacturing could generating interesting new insights. Due to their expertise, the experts might assess or perceive hurdles differently than companies with less experience in these areas. Future work needs to clarify our findings through larger, quantitative studies. Second, in our results and discussion sections, we focused on the concepts of trust, usefulness and value in data sharing. These concepts are mainly relevant for the emergence of food industry data ecosystems. Within the interviews, additional aspects, such as future of food industry and data sharing, sustainability, acquiring and retaining customers or acquiring suitable data products arose. These aspects focus rather on perpetuation of data ecosystems and need to be considered in more detail in future work on food industry data ecosystems. Third, we mainly focused on the identification and theoretical alignment of concepts for implementing data sharing practices in food industry data ecosystems. Future research needs to specify how to realize these concepts and overcome hurdles in terms of data ecosystem design, acceptance and perpetuation of data ecosystems, e.g., by formulating possible pathways. Further theories exist that could generate insights into data sharing but did not arise from our results, such as George Akerlof's information asymmetry theory. In future research, this and related theories need to be further examined in data sharing.

Our work contributes to both theory and practice. Building on our results, we developed a conceptual model representing the relationships between theoretical concepts for implementing data sharing in food industry data ecosystems. The empirical results support the concepts of IS theories TAM in conjunction with social exchange theory in the context of data sharing. The results offer insights into current opportunities and barriers in context of data sharing in the food industry. Thereby, we support the future emergence of data sharing in food industry data ecosystems.

Acknowledgement

This work is part-funded by German Federal Ministry for Economics and Climate Action (BMWK) within Smart Data Economy under grant number 01MT19003B (EVAREST) & 01MD19010C (Future Data Assets).

REFERENCES

Abbas, A.E., Agahari, W., van de Ven, M., Zuiderwijk, A., and de Reuver, M. 2021. "Business Data Sharing through Data Marketplaces: A Systematic Literature Review," *J. Theor. Appl. Electron. Commer. Res.* (16:7), pp. 3321-3339.

Anand, A., Coltman, T., and Sharma, R. 2016. "Four Steps to Realizing Business Value from Digital Data Streams," *MIS Q. Exec.* (15:4).

Arnaut, C., Pont, M., Scaria, E., Berghmans, A., and Leconte, S. 2018. "Study on Data Sharing between Companies in Europe," in: *A study prepared for the European Commission, European Commission.*

Bogner, A., Littig, B., and Menz, W. 2014. *Interviews Mit Experten*. Springer-Verlag.

Cappiello, C., Gal, A., Jarke, M., and Rehof, J. 2020. "Data Ecosystems: Sovereign Data Exchange among Organizations (Dagstuhl Seminar 19391)," in: *Dagstuhl Reports*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik.

Cook, K.S. 1977. "Exchange and Power in Networks of Interorganizational Relations," *Sociol Q* (18:1), pp. 62-82.

Davis, F.D. 1989. "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology," *MIS Q* (13:3), pp. 319-340.

Döring, N., and Bortz, J. 2016. *Forschungsmethoden Und Evaluation*. Wiesbaden: Springer:

Durrant, A., Markovic, M., Matthews, D., May, D., Enright, J., and Leontidis, G. 2022. "The Role of Cross-Silo Federated Learning in Facilitating Data Sharing in the Agri-Food Sector," *Comput. Electron. Agric.* (193), p. 106648.

Emerson, R.M. 1972. "Exchange Theory, Part I: A Psychological Basis for Social Exchange," *Sociol Theory* (2), pp. 38-57.

European-Commission. 2017. "Building a European Data Economy ". Brussels.

Eurostat. 2020. "Employment in Food Supply across Eu Regions." Retrieved 2022-02-07, from https://ec.europa.eu/eurostat/de/web/products-eurostat-news/-/ddn-20200522-2

Frimpong, J. 2021. "Food Report 2021," Statista, Statista, p. 10.

Gasco-Hernandez, M., Feng, W., and Gil-Garcia, J.R. 2018. "Providing Public Value through Data Sharing: Understanding Critical Factors of Food Traceability for Local Farms and Institutional Buyers," *Proc.of 51st HICSS*.

Gelhaar, J., Gürpinar, T., Henke, M., and Otto, B. 2021. "Towards a Taxonomy of Incentive Mechanisms for Data Sharing in Data Ecosystems," *Proc. of 25th PACIS*), p. 121.

Gelhaar, J., and Otto, B. 2020. "Challenges in the Emergence of Data Ecosystems," *Proc. of 24th PACIS*, p. 175.

Gibbs, G.R. 2007. "Thematic Coding and Categorizing," *Analyzing qualitative data* (703), pp. 38-56.

Gläser, J., and Laudel, G. 2009. *Experteninterviews Und Qualitative Inhaltsanalyse*. Springer-Verlag.

Goodwin, N.C. 1987. "Functionality and Usability," *Commun. ACM* (30:3), pp. 229–233.

Kankanhalli, A., Tan, B.C., and Wei, K.-K. 2005. "Contributing Knowledge to Electronic Knowledge Repositories: An Empirical Investigation," *MIS Q* (29:1), pp. 113-143.

Klein, R., and Rai, A. 2009. "Interfirm Strategic Information Flows in Logistics Supply Chain Relationships," *MIS Q* (33:4), pp. 735-762.

KPMG. 2016. "Survival of the Smartest - Studie Zur Digitalen Transformation."

Krippendorff, K. 2018. *Content Analysis: An Introduction to Its Methodology*. Sage publications.

Likert, R. 1932. "A Technique for the Measurement of Attitudes," *Archives of psychology* (22:1).

Lindner, M., Straub, S., and Kühne, B. 2021. "How to Share Data? ."

Lorenz, J.-T. 2018. "The Rise of Ecosystems and Platforms: What Role Can Insurers Play and How Can They Get Started?" Retrieved 2022-04-22, from https://www.mckinsey.com/

Maass, W. 2009. *Elektronische Wissensmärkte*. Springer-Verlag.

Maass, W. 2022. "Contract-Based Data-Drivin Decision Making in Federated Data Ecosystems " in: *Proc. of 55th HICSS*.

Mandrella, M., Trang, S., and Kolbe, L.M. 2020. "Synthesizing and Integrating Research on It-Based Value Cocreation: A Meta-Analysis," *J. Assoc. Inf.* (21:2), pp. 388-427.

Maxwell, J.A. 2012. *Qualitative Research Design: An Interactive Approach*. Sage publications.

Myers, M.D., and Newman, M. 2007. "The Qualitative Interview in Is Research: Examining the Craft," *Inf Organ* (17:1), pp. 2-26.

Oliveira, M.I.S., Barros Lima, G.d.F., and Farias Lóscio, B. 2019. "Investigations into Data Ecosystems: A Systematic Mapping Study," *Knowl Inf Syst* (61:2), pp. 589-630.

Oliveira, M.I.S., and Lóscio, B.F. 2018. "What Is a Data Ecosystem?," in: *Proc. of 19th Annual International Conference on DGO*. p. Article 74.

Oliveira, M.I.S., Oliveira, L.E.R., Batista, M.G.R., and Lóscio, B.F. 2018. "Towards a Meta-Model for Data Ecosystems," *Proc. of 19th Annual International Conference on DGO*, pp. 1-10.

Putnam, R. 1993. "The Prosperous Community: Social Capital and Public Life," *Am prospect* (13).

Rix, C., Stein, H., Chen, Q., Frank, J., and Maass, W. 2021. "Conceptualizing Data Ecosystems for Industrial Food Production," in: *Proc. of IEEE CBI*.

Röhl, K.-H., Bolwin, L., and Hüttl, P. 2021. "Datenwirtschaft in Deutschland."

Selander, L., Henfridsson, O., and Svahn, F. 2013. "Capability Search and Redeem across Digital Ecosystems," *J. Inf. Technol.* (28:3), pp. 183-197.

Shapiro, C., Carl, S., and Varian, H.R. 1998. *Information Rules*. Harvard Business Press.

Söllner, M., Benbasat, I., Gefen, D., Leimeister, J.M., and Pavlou, P.A. 2016. "Trust," *MIS Q Research Curations* (1:2).

Stein, H., and Maass, W. 2022. "Requirements for Data Valuation Methods," in: *Proc. of 55th HICSS*. Hawaii: pp. 6155 - 6164.

Tamm, T., Seddon, P.B., and Shanks, G. 2020. "How Do Different Types of Ba Users Contribute to Business Value?," *Commun Assoc Inf Syst* (46:1), p. 28.

Tao, Q., Ding, H., Wang, H., and Cui, X. 2021. "Application Research: Big Data in Food Industry," *Foods* (10:9), p. 2203.

Zech, H. 2017. "Data as a Tradeable Commodity–Implications for Contract Law," *Proc. of 18th EIPIN*.

Zuiderwijk, A., Janssen, M., and Davis, C. 2014. "Innovation with Open Data: Essential Elements of Open Data Ecosystems," *Inf. polity* (19:1-2), pp. 17-33.