UNIVERSITY OF TURKU

# Corporate Data Sharing: A Qualitative Analysis of Factors Influencing Data Sharing and the Impact of Privacy Enhancing Technologies

Information Systems Science/ Department Management and Entrepreneurship

Master's thesis

Author:
Felix Starnecker

16.12.2024

Turku

Master's thesis

**Subject**: Information Systems Science; Double Degree Digital Enterprise Management/ Information Systems (University of Passau)
**Author**: Felix Starnecker
**Title**: Corporate Data Sharing: A Qualitative Analysis of Factors Influencing Data Sharing and the Impact of Privacy Enhancing Technologies
**Supervisor(s)**: Prof. Jukka Heikkilä, Prof. Thomas Widjaja
**Number of pages**: 77 pages (without Appendices)
**Date**: 16.12.2024

**Abstract**

Companies are often reluctant to share data. Reasons for this include data protection issues, legal uncertainties and concerns about inadvertently disclosing trade secrets. Privacy Enhancing Technologies (PETs), such as Secure Multiparty Computation, which make it possible to share data without disclosing it to other parties, could therefore have a significant impact on companies' data sharing decisions. This paper examines the factors that companies consider when deciding to share data and how PETs can influence this process. The study is based on a systematic literature review and 20 expert interviews. The results show two models that depict the decision factors for sharing data - with and without PETs. They show that PETs change the factor of trust in the sharing of data between companies: Part of the interpersonal trust can be replaced by technological trust. However, due to the complexity, low prevalence and limited positive application examples of PETs, interpersonal trust remains central. In addition, the level of effort plays a crucial role in the sharing of data with PETs, as companies often decide against this type of technology due to a lack of time, expertise or convenience, despite the technological advantages. Practical implications include the use of consulting companies as intermediaries to build trust and closer cooperation between PET providers to promote the acceptance of these technologies.

**Key words**: Privacy Enhancing Technologies, PETs, Secure Multiparty Computation, Data Sharing, Data Exchange, Decision Factors

# Table of contents


**List of Figures** 5

**1 Introduction** 6

**2 Theoretical Basics** 9

**2.1 Data Disclosure Decisions** 9

**2.2 Individual and Organizational Information Privacy: an Overview** 11

**2.3 Privacy Enhancing Technologies** 12

**2.4 Data Sharing: Viewpoints and Interpretations** 14

**3 Method of Systematic Literature Review** 16

**4 Systematic Literature Review** 17

**4.1 Factors Influencing Company Decisions on Data Sharing** 17

4.1.1 The Risk-Cost-Benefit Analysis 17

4.1.2 The Importance of Trust and Control in Data Sharing 18

4.1.3 The Role of Interoperability in Cross-Company Data Sharing 20

4.1.4 The Influence of Internal Corporate Factors on Data Sharing 22

4.1.5 The Influence of External Factors on Data Sharing 25

**4.2 The Influence of Privacy Enhancing Technologies on Data Sharing Factors** 26

4.2.1 Influence of PETs on the Factors of Trust, Control, and their Interactions 26

4.2.2 The Influence of PETs on the Trade-off Between Risks, Costs, and Benefits 28

4.2.3 Internal Factors in Connection with PETs 30

4.2.4 External Factors in Connection with PETs 31

**5 Method of Interviews** 34

**6 Results of the Interviews** 36

**6.1 Factors that Influence Data-Sharing Decisions** 36

6.1.1 Requirements for Successful Data Sharing 36

6.1.2 The Importance of Trust and Control for Data Sharing 37

6.1.3 The Trade-off Between Risks and Benefits when Sharing Data 38

6.1.4 The Influence of Internal Factors on the Willingness to Share Data 41

6.1.5 The Influence of External Factors on the Willingness to Share Data 43

**6.2 The Impact of PETs on Data Sharing and its Influencing Factors** 44

6.2.1 Benefits and Risks of PET-Based Data Sharing 45

6.2.2 The Importance of Trust and Control in the Context of PETs 47

6.2.3 Internal Factors Influencing PET-Based Data Sharing 48

6.2.4 External Factors Influencing PET-Based Data Sharing 50
6.2.5 Challenges in the Acceptance of PETs on the Market 50

**7 Discussion 53**

**7.1 Key Findings 53**
7.1.1 Factors Influencing Corporate Data Sharing 53
7.1.2 The Impact of PETs on Data Sharing and its Influencing Factors 54

**7.2 Theoretical Implications 56**
7.2.1 Corporate Data Sharing Model 56
7.2.2 Corporate PET-Based Data Sharing Model 60
7.2.3 Level of Privacy with PETs 63

**7.3 Practical Implications 63**

**7.4 Limitations and Outlook 66**

**References 68**

**Appendices 78**

**Appendix A 78**
A.1 Concept Matrix ‘Corporate Data Sharing’ 78
A.2 Concept Matrix ‘Corporate Data Sharing with PETs’ 79

**Appendix B Interview Questionnaire 81**

**Appendix C Interview Analysis 82**

**Appendix D Interview Transcripts 120**

# List of Figures

Figure 1: Factors influencing data sharing ..... 17
Figure 2: Factors influencing the sharing of data with PETs ..... 26
Figure 3: From Interviews: Factors influencing data sharing ..... 36
Figure 4: From Interviews: Factors influencing the sharing of data with PETs ..... 45
Figure 5: Corporate data sharing model ..... 56
Figure 6: Corporate PET-based data sharing model ..... 60

# 1 Introduction

Data unfolds its greatest value when combining many data sets from different actors (Gelhaar & Otto, 2020). Therefore, it is becoming increasingly important for companies to share their data (Gelhaar & Otto, 2020). However, companies are reluctant to share data (Bitkom, 2024). The main reasons are data protection issues, legal uncertainties, a lack of data compatibility, and concerns about accidentally disclosing trade secrets (Bitkom, 2024). The reluctance of companies to share data with other companies is justified. Cyberattacks originating from actors within a company's data ecosystem, such as suppliers or other business partners, cause the highest costs compared to other attack methods at an average of 4.99 million US dollars (IBM, 2024). This means that shared data usage also entails the greatest risks. Therefore, companies must carefully weigh the benefits of data sharing against the associated risks (Arora et al., 2016; Witte et al., 2020). In addition to risks and benefits, many other factors also affect companies' willingness to share their data (Dahlberg & Nokkala, 2019). However, investigating these factors is challenging, as they are often shaped by subtle organizational dynamics and complex, difficult-to-penetrate interrelationships within companies (Smith et al., 2011). Moreover, the existing knowledge about these factors is based on a limited database and should, therefore, be consolidated and expanded through further research (Dahlberg & Nokkala, 2019; Müller et al., 2020). One aim of this paper is thus to analyze in more detail the factors that influence companies' decisions to share data. To this end, the following research question is posed: What factors do companies weigh when sharing data with other companies?

The technology used for data sharing also plays an important role in determining the factors influencing data sharing between companies (Müller et al., 2020). Emerging technologies can influence the factors that companies weigh up when sharing data (Müller et al., 2020). In this context, Privacy Enhancing Technologies (PETs) are increasingly becoming the focus of attention (Gartner, 2024; World Economic Forum, 2024). They enable data to be shared while ensuring privacy, security, and data sovereignty (United States Government, 2022). PETs could access data without disclosing sensitive information and without de-identification risks (United Nations, 2023). As a result, Privacy Enhancing Technologies fundamentally change the use of data and the factors that organizations consider when deciding on data sharing (Agahari et al., 2022). However, the generalizability of the research results to date is limited, which underlines the need for further research in this area (Agahari et al., 2022; Hasani et al.,

2023). There is a call to further investigate data sharing between companies in the context of such new technologies (Gelhaar & Otto, 2020; Müller et al., 2020), particularly regarding how PETs influence the willingness of companies to share data (Agahari et al., 2021). This thesis aims to contribute to closing this research gap by investigating the impact of PETs on companies' data-sharing behavior. As Privacy Enhancing Technologies encompass a wide range of different technologies (Hasani et al., 2023), the focus will be placed specifically on Secure Multiparty Computation to narrow down the technologies. The following research question is posed: How do Privacy Enhancing Technologies, particularly Secure Multiparty Computation, influence the decision-making factors and usage patterns for sharing data among companies, and to what extent do these technologies promote more intensive data sharing?

The thesis is divided into four blocks: theoretical basics, method and results of the systematic literature review, method and results of the interviews, and the discussion. In the theoretical basics, the extended APCO model by Dinev et al. (2015) is explained as a model that represents the decision-making process of individuals when disclosing data. This will serve as the basis for the creation of two models of data sharing in the corporate context. In addition, the theoretical basis provides a more detailed explanation of the terms privacy and data sharing and a basic understanding of PETs. In the second part of the thesis, a systematic literature review is carried out to analyze existing findings on data sharing between companies and the influence of PETs (Webster & Watson, 2002). The search string '(data sharing OR data exchange OR information sharing OR information exchange) AND (corporate OR enterprise OR company)' was used to search the EBSCO database to identify papers on corporate data-sharing behavior. The overarching factors of interoperability, trust, control, risks, costs, benefits, internal factors, and external factors that influence data sharing between companies emerge from the existing literature. A second search string '(data sharing OR data exchange OR information sharing OR information exchange) AND (Privacy Enhancing Technologies OR Secure Multiparty Computation OR Privacy Preserving Technologies OR PETs OR MPC)' was used to search the EBSCO database for papers on the impact of PETs on corporate data sharing. With the papers identified from this, changes in the previously identified factors can be determined. The results of the systematic literature review are captured in two concept matrices (see Appendix A). Based on the results of the systematic literature review, 20 interviews were conducted with experts from the IT sector. Of the 20 experts, ten were experts specifically in the field of Privacy Enhancing Technologies. The

results of the interviews are presented in the third part of the thesis. The interviews enabled factors from the systematic literature review to be confirmed and new factors to be identified. Finally, in the discussion section, the results from the previous chapters were used to create two models based on the extended APCO model (corporate data sharing model; corporate PET-based data sharing model). Furthermore, practical implications for suppliers of PETs, companies, and legislators are presented, the limitations of the work are discussed, and an outlook is given.

# 2 Theoretical Basics

## 2.1 Data Disclosure Decisions

When considering data sharing, studies need to distinguish between the perspectives of individuals and organizations (Smith et al., 2011). At the individual level, much research has been conducted on data-sharing decisions (Smith et al., 2011). One of the best-established models emerging from this is the Antecedents-Privacy Concerns-Outcomes (APCO) model (Smith et al., 2011). It is a fundamental model for analyzing the behavior of individuals about data disclosure decisions (Smith et al., 2011). At the center of this model is the variable of privacy concerns (Smith et al., 2011). Privacy concerns can be described as individual concerns about the disclosure of personal information (Bansal et al., 2016). Privacy experiences, privacy awareness, personality differences, demographic differences, and culture/climate influence the development of privacy concerns (Smith et al., 2011; Bansal et al., 2016). Privacy concerns, in turn, influence individuals' data-sharing behavior by changing the perception of the risks and benefits of data-sharing (Smith et al., 2011). Furthermore, privacy concerns have a reciprocal relationship with trust. On the one hand, trust can reduce privacy concerns; on the other hand, privacy concerns can also influence trust (Smith et al., 2011).

Trust also directly influences the decision to share data (Smith et al., 2011). Trust is the belief that shared personal information will not be used opportunistically (Dinev & Hart, 2006). It encompasses the three main aspects of competence, reliability, and security (Dinev & Hart, 2006). Thus, trust is not just the belief that the recipient of the personal information has the ability and reliability to perform its promised actions but also that the information will be kept secure (Dinev & Hart, 2006). How well trust is established depends on the personality traits of the individual (Bansal et al., 2016).

In addition to the factors of trust and privacy concerns, the privacy calculus is the third decisive factor that plays a role in the data disclosure decision (Smith et al., 2011). This factor describes the individual's weighing up between the privacy risks, costs, and benefits of disclosing data (Smith et al., 2011). If the benefits outweigh the privacy risks and costs, the person is more willing to share their data (Smith et al., 2011). Privacy risks are the expectation of potential losses resulting from the sharing of data (Aleem et al., 2017; Xu et

al., 2011). The main benefits of data sharing for individuals are financial rewards, personalization, and social adjustment benefits (Smith et al., 2011).

The construct of privacy uncertainty is closely related to the privacy risk (Aleem et al., 2017). Therefore, it's important to mention this construct in connection with the privacy calculus as well. Privacy uncertainty refers to "... consumers' inability to evaluate privacy risk due to imperfect information" (Aleem et al., 2017, p. 10). The insufficient availability of information as a cause of privacy uncertainty can be divided into pre-purchase and post-purchase information asymmetry (Aleem et al.,2017). Pre-purchase information asymmetry is the customer's perception that the seller is not fully transparent about what personal data he collects about the customer and is referred to as 'hidden information' (Aleem et al., 2017). Post-purchase information asymmetry can be divided into 'hidden action' and 'hidden effort' (Aleem et al., 2017). Hidden action is the consumer's perception that the seller does not always act as he claims. With hidden effort, the consumer has the impression that the seller does not devote enough resources to adequately protect the consumer's data (Aleem et al., 2017). Post-purchase information asymmetry, in particular, harms the consumer's intention to share their data (Aleem et al., 2017). Privacy uncertainty can be reduced through trust and social presence (Pavlou et al., 2007).

However, the assumption of the APCO model that individuals carry out a rational cost-benefit analysis when deciding whether to disclose data is unlikely (Dinev et al., 2015). For example, positive mood has also been found to have an impact on individuals' decision-making (Alashoor et al., 2022). An extended APCO model was thus established (Dinev et al., 2015). The distinctive feature of the extended APCO model is that it does not assume that individuals always make full use of their cognitive capacities and knowledge when making decisions about behavior regarding the disclosure of data (Dinev et al., 2015). Instead, in the extended APCO model, the so-called 'level of effort' and other external influences such as 'peripheral cues, biases, heuristics, misattribution' are also considered (Dinev et al., 2015). The 'level of effort' is determined by the individual's feelings, cognitive resources, motivation, and time constraints (Dinev et al., 2015). While data protection-relevant information is processed thoroughly and logically at a 'high level of effort', external influences such as peripheral clues, prejudices, heuristics, and misattributions dominate the decision-making process at a 'low level of effort' (Dinev et al., 2015). Rational trade-offs between risks, costs, and benefits are neglected in this situation (Dinev et al., 2015).

Similar to individuals, in companies, decision-making regarding data sharing is also often very subjective and based on personal judgments (Fassnacht et al., 2023; Fricker & Maksimov, 2017). Data privacy concerns can be considered at both an individual and an organizational level (Belanger & Crossler, 2011). It can be seen that the two levels are interrelated in certain areas and influence each other (Belanger & Crossler, 2011). For example, the data privacy concerns of companies are influenced by external factors in the same way as those of individuals (Belanger & Crossler, 2011). Similarly, organizational data-sharing decisions differ in some aspects (Smith et al., 2011; Belanger & Crossler, 2011). However, as studies at the organizational level are more complex and less well suited to rapid data collection methods, research at this level is limited (Smith et al., 2011). There is no clear understanding of the extent to which the two levels are related and how they differ. Therefore, this paper compares the decision-making of individuals based on the extended APCO model with the factors that influence organizations when sharing data.

## 2.2 Individual and Organizational Information Privacy: an Overview

It has already been shown that privacy concerns are a central element in individuals' data-sharing decisions (Smith et al., 2011). So, to compare individuals' data-sharing behavior with the behavior of companies, the understanding of privacy and the differences between these two levels must be highlighted.

When looking at definitions of privacy, a distinction can be made between a value-based and a cognate-based approach (Smith et al., 2011). The value-based view considers privacy as a right to which everyone is entitled (Smith et al., 2011). Cognate-based definitions consider privacy as a state that is described by different dimensions (Smith et al., 2011). This also implies that various degrees of privacy can be achieved (Smith et al., 2011). Furthermore, two types of privacy can be distinguished: Physical Privacy and Information Privacy (Smith et al., 2011). In this paper, the focus is on information privacy.

In the individual context, information privacy can be described as "... ability to manage information about oneself." (Belanger et al., 2002, p.249). Information privacy thus emphasizes the desire of individuals to retain control over their data (Belanger & Crossler, 2011). Four main information privacy concerns can be identified: Concern about the collection and storage of personal data, unauthorized re-use, unauthorized access, and the risk of errors (Smith et al., 1996). Despite the desire for information privacy, it can be observed that individuals pass on personal data and thus give up control over it (Smith et al., 2011).

This paradox illustrates the economic dimension of information privacy: it can also be seen as a tradable commodity and exchanged for other benefits (Smith et al., 2011). Several regulations have been introduced to strengthen individual information privacy. The strictest data protection law in the world is the General Data Protection Regulation (GDPR), adopted by the European Union in 2018 (GDPR, n.d.). The regulation aims to control the collection and use of European citizens' data by companies (GDPR, n.d.).

In the previous explanations, information privacy was considered from the perspective of individuals. However, information privacy is a multilevel concept that should be analyzed on several levels (Belanger & Crossler, 2011). In addition to individual information privacy, group and organizational information privacy can be identified (Skinner et al., 2006). These levels are related in certain aspects, but in other aspects, they differ from each other (Belanger & Crossler, 2011). While individual information privacy concerns are often influenced by characteristics such as gender, age, and level of education, organizational information privacy concerns may be more strongly determined by external factors such as the industry, the competitive position of the product, and legal regulations (Belanger & Crossler, 2011). Organizational information privacy concerns are described as "... the overall concern that organizational leaders have regarding the privacy of the information the organization possesses and has access to." (Belanger & Crossler, 2011, p.1033). The traditional understanding of information privacy as a right only individuals have to protect themselves from more powerful actors such as companies must be expanded (Agrawal et al., 2021). In the organizational context, information privacy includes meanings such as competitive secrecy, protection of corporate data, and state security (Agrawal et al., 2021). So-called Privacy Enhancing Technologies can be used to strengthen the information privacy of organizations (Belanger & Crossler, 2011). These technologies are explained in more detail in the following section.

## 2.3 Privacy Enhancing Technologies

Initial definitions of PETs presented them as technologies that support individuals having more control over their data (Seničar et al., 2003). With the help of PETs, the balance of power between the individual and the many actors in the online environment should become more equal (Seničar et al., 2003). As the responsibility for personal data has changed over time, the definitions of PETs have also changed their focus (OECD, 2023). In the past, individuals themselves were responsible for their data. Today, this responsibility lies with the

organizations. Therefore, current definitions of PETs increasingly include the role of organizations (OECD, 2023). They are defined as "… a broad scope of technologies and applications that are planned to improve the protection and information security of both individual ... and corporate users ... in web-based exercises and correspondences." (Hasani et al., 2023, p.3). However, this description of PETs include countless technologies (Hasani et al., 2023). To limit the number of technologies for this thesis, a definition of PETs is chosen that narrows them down to "… privacy-preserving data sharing and analytics technologies, which describes the set of techniques and approaches that enable data sharing and analysis among participating parties while maintaining disassociability and confidentiality." (United States Government, 2022). This definition includes only a special type of PET.

For a better overview of the different types of PETs, they can be categorized according to application areas (Garrido et al., 2022). A distinction can be made between the communication, storage, verification, sovereignty, and processing layers (Garrido et al., 2022). When they ensure confidentiality during data transmission, they are assigned to the communication layer (Garrido et al., 2022). The storage layer contains all PETs that ensure data privacy at rest. PETs in the verification layer guarantee the authenticity and integrity of data (Garrido et al., 2022). The sovereignty layer comprises PETs that give data owners options for better control over their data, such as privacy policies or smart contracts (Garrido et al., 2022). When PETs support privacy during data processing, they are assigned to the processing layer (Garrido et al., 2022). These PETs are also described as privacy-preserving computation technologies and aim to ensure privacy while performing computational analyses (Agrawal et al., 2021). Homomorphic encryption, secure multiparty computation, and differential privacy fall into this category (Agrawal et al., 2021; Garrido et al., 2022; Choi & Butler, 2019). These technologies could enable organizations to share personal data and perform joint analyses without violating regulations such as the GDPR (Scheibner et al., 2021; Helminger & Rechenberger, 2022). This would allow privacy to remain protected while facilitating the free exchange of personal data (Helminger & Rechenberger, 2022). The thesis refers only to these privacy-preserving computation technologies.

A special focus in this work is placed on Secure Multiparty Computation (SMPC). SMPC is a cryptographic method that allows multiple parties to jointly process data without disclosing their respective data (Zhao et al., 2019). It means that data remains protected even during computation (Zhao et al., 2019). This can be compared to an ideal world in which there is a trustworthy third party to which various companies can hand over their data (Helminger &

Rechenberger, 2022). The trusted party then performs calculations on the data and only passes the final result to the companies. The original data remains protected throughout the process (Helminger & Rechenberger, 2022). Through advanced encryption-like techniques, SMPC can replace this trusted party of the ideal world (Helminger & Rechenberger, 2022).

Even though the focus of this thesis is on the PETs subcategory of privacy-preserving computation technologies and SMPC in particular, the term Privacy Enhancing Technologies (PETs) will continue to be used in the following chapters for the sake of simplicity.

## 2.4 Data Sharing: Viewpoints and Interpretations

In the literature, the term ‘data sharing’ is often associated with a focus on different characteristics of data sharing (Jussen et al., 2023). The focus of the investigation can be placed on the infrastructure, the service, the data type, the technology, or the groups involved. In terms of the used infrastructure, a distinction can be made between data marketplaces and data-sharing platforms. Data-sharing services can be ‘match-making’, ‘privacy-as-a-service’, or ‘analytics’. Depending on the use case, ‘anonymous personal data’, ‘metadata’, or ‘aggregated data’ can be shared (Jussen et al., 2023). A variation of the term blockchain-based data sharing or cloud data sharing emphasizes the technology used to share the data (Jussen et al., 2023). Business-, government-, or personal data sharing emphasizes the groups involved in the exchange (Jussen et al., 2023).

As this paper looks at data sharing in a corporate context and compares it to the decision-making of individuals when sharing data, the focus is on the groups involved. This includes understanding the differences between business data-sharing practices and personal data-sharing behaviors. When individuals disclose their data, data sharing usually refers to the scope and depth of the data disclosed (Krasnova et al., 2010). In the corporate context, data sharing can be described very fundamentally as the exchange of protected information between trading partners (Li & Lin, 2006). Protected information is usually data from central business processes (Li & Lin, 2006). The shared data must be useful to the other party and exchanged at the appropriate time (Tong & Crosno, 2016). Furthermore, the influence of PETs on data sharing between companies plays an important role in this work. Therefore, when PETs are used for data sharing, the term PET-based data sharing is introduced based on the previously described terms blockchain-based data sharing and cloud data sharing. This is intended to emphasize the difference between normal data sharing and data sharing with the inclusion of Privacy Enhancing Technologies.

In addition, an understanding of data sharing also includes related activities, such as processes and practices that support data-sharing relationships or are required for successful data sharing (Fassnacht et al., 2023). Factors that influence the decision to share data must also be taken into account when understanding data sharing (Fassnacht et al., 2023). As this paper looks at the data-sharing behavior of companies from different perspectives and factors of all kinds that influence companies' data-sharing decisions are included, a broader definition of the term has been chosen. When all relevant aspects are included, data sharing can be defined as

> … the domain-independent process of giving third parties access to the data sets of others. .... The expectation is to be compensated financially or through other benefits (e.g., receiving data) for providing the data. What the data may be used for and how it is made available is determined within the framework of the (legal) agreements between the data providers, data consumers, and other roles, depending on the use case. (Jussen et al., 2023, p.4).

# 3 Method of Systematic Literature Review

To answer the research questions “What factors do companies weigh when sharing data with other companies?” and “How do Privacy Enhancing Technologies, particularly Secure Multiparty Computation, influence the decision-making factors and usage patterns for sharing data among companies, and to what extent do these technologies promote a more intensive data sharing?”, the systematic literature review method is used (Webster & Watson, 2002).

To conduct a systematic literature review, the EBSCO database was searched for suitable literature. The search in the databases was limited to the journals of the AIS Senior Basket of Journals, the FT50, and the VHB complete list from A+ to B. This ensured that only high-quality and relevant sources were considered. In addition, gray literature was also included to capture new and practical developments. Google Scholar was used as a source for gray literature. The search string for identifying relevant literature regarding the question of which factors companies consider when sharing data was: “(data sharing OR data exchange OR information sharing OR information exchange) AND (corporate OR enterprise OR company). To narrow down the results, the first part of the string was searched in the title, and the second part in the abstract. This resulted in 128 papers in the EBSCO database. Filtering the papers by title reduced the number to 42 remaining papers. Of these papers, 19 papers that appeared relevant to the subject area remained after reading the abstract. To obtain literature on the influence of PETs on the data-sharing behavior of companies, the search string was changed to “(data sharing OR data exchange OR information sharing OR information exchange) AND (Privacy Enhancing Technologies OR Secure Multiparty Computation OR Privacy Preserving Technologies OR PETs OR MPC)”. The first part of the string was searched in the abstract, and the second part in the complete text. The search string produced 40 results in the EBSCO database. After filtering the literature by reading the title and abstract, six suitable articles remained. Due to the rather small number of results matching the context, the forward and especially backward search was another major part of the structured literature search. The following section explains the results obtained from the literature found. In addition, two concept matrices have been created to present the results of the literature search. The first shows the factors companies generally consider when sharing data (see Appendix A.1). The second outlines the factors that are important when sharing data with PETs (see Appendix A.2).

# 4 Systematic Literature Review

## 4.1 Factors Influencing Company Decisions on Data Sharing

The following sections analyze the factors that influence the decision-making process for cross-company data sharing. The figure below shows the concepts that influence cross-company data sharing in the same structure as they were summarized in the sub-chapters.

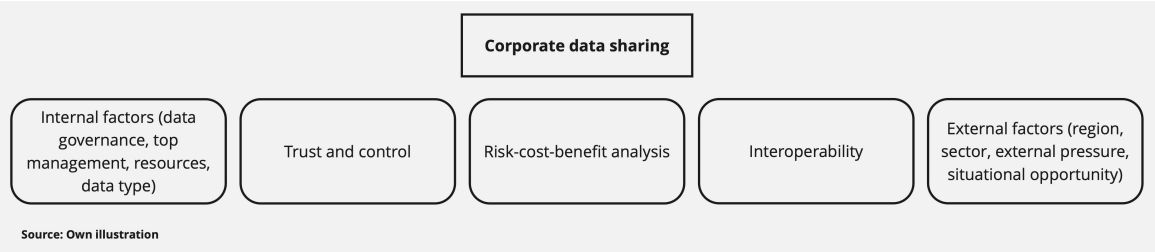


*Figure 1: Factors influencing data sharing*

### 4.1.1 The Risk-Cost-Benefit Analysis

The fundamental trade-off between risks, costs, and benefits in data-sharing decisions, which is already emphasized at the individual level, can also be observed in the corporate context (Smith et al., 2011; Dahlberg & Nokkala, 2019; Penttinen et al., 2018). If companies decide to share their data, they can reap numerous '**benefits**' (Gelhaar & Otto, 2020). For example, data sharing creates new opportunities to better monitor supply chain processes and thus increase efficiency (Dahlberg & Nokkala, 2019). Access to real-time data across companies makes it possible to optimize processes through data analyses, demand balancing, and predictive analyses (Müller et al., 2018). This enables companies to react more quickly to unforeseen events within the supply chain (Müller et al., 2020). Especially in times of geopolitical tensions, pandemics, and climate change, sharing data as a supply chain risk mitigation and resilience strategy is a key advantage (Bechtsis et al., 2022). In addition, connectivity along the value chain enables companies to achieve greater customer reach, simplify the communication process for orders and their fulfillment, and facilitate payment transactions (Müller et al., 2018). Aside from these operational benefits, exchanging data can also help achieve strategic business goals (Müller et al., 2020). Sharing data between companies not only offers potential for innovations but can even lead to the development of completely new business models (Enders et al., 2020; Müller et al., 2020). There is also the possibility that companies create joint business models when sharing data (Müller et al., 2020).

However, these benefits of data sharing are also accompanied by '**risks**' and '**costs**' that companies must carefully weigh up. If the costs exceed the benefits, this is a decisive factor for companies against data sharing (Penttinen et al., 2018). Sharing data can often be associated with considerable costs, which stand in comparison to uncertain and difficult-to-assess future benefits (Dahlberg & Nokkala, 2019). It is striking that even low costs associated with data sharing between companies reduce these costs disproportionately (Hoffmann et al., 2020). In addition, cross-company data sharing is associated with competition-related-, misuse-related-, end-user-related data privacy-, and reputation-related risks (Agahari et al., 2022). Of these, the 'competitiveness risk' represents one of the greatest risks (Agahari et al., 2022). It describes the risk of sensitive data falling into the hands of competitors (Gelhaar & Otto, 2020). But even if data falls into the hands of non-competing companies, there is a competitive risk (Müller et al., 2018). For example, sharing data along the value chain can lead to increased transparency in production processes, which can be exploited by other companies (Müller et al., 2018). As 'transparent suppliers', they could be exposed to increased price pressure, which could undermine their competitiveness (Müller et al., 2018). Particularly in highly competitive markets, this has a negative impact on the willingness of companies to share data (Hoffmann et al., 2020). Another risk is the 'data misuse risk' (Agahari et al., 2022). This is because as soon as a company passes on its data, it loses control over how and for what purposes it is used (Agahari et al., 2022). The 'end-user privacy risk' is also of great importance for companies, as a breach of customer data privacy can result in severe penalties under the GPPR (Agahari et al., 2022). The 'reputation risk' is also closely linked to this. If the privacy of customer data is violated, this can hurt the image or brand of the company (Agahari et al., 2022). In addition, there is always the risk of technological failures, hacker attacks, or unauthorized access when sharing data (Dahlberg & Nokkala, 2019). To encourage companies to share data despite these risks, it is important to make them aware of the benefits of data sharing (Gelhaar & Otto, 2020).

### 4.1.2 The Importance of Trust and Control in Data Sharing

In the case of data disclosure decisions at an individual level, the factor of trust has a positive influence on data sharing (Smith et al., 2011). Also, in companies, the factor of '**Trust**' plays a decisive role (Gelhaar & Otto, 2020; Holler et al., 2019; Li & Lin, 2006). Data-sharing ecosystems are easier to create if companies have known each other for a longer period and thus reinforce the impression of a trustworthy environment (Gelhaar & Otto, 2020). It can also be stated that trust and data sharing have a mutual causal relationship (Dyer & Chu,

2003). Companies that trust each other are more willing to share data. One of the reasons for this is that the need to draw up detailed contracts and agreements is reduced, which lowers transaction costs (Dyer & Chu, 2003). At the same time, intensive data sharing further promotes mutual trust, which strengthens the relationships between the companies involved and optimizes the flow of data (Dyer & Chu, 2003). More intensive data sharing not only increases trust but can also be associated with reduced opportunistic behavior (Tong & Crosno, 2016). However, it can also be observed that the negative correlation between data sharing and opportunistic behavior decreases in longer relationships (Tong & Crosno, 2016). This can be explained by the honeymoon effect (Tong & Crosno, 2016). In the earlier stages of a relationship, both parties are more concerned about protecting shared data. This motivation diminishes as the relationship progresses, and the likelihood of opportunistic behavior increases (Tong & Crosno, 2016).

Just as trust can increase data sharing, a lack of trust can also reduce it (Li & Lin, 2006). For example, companies' uncertainty about the reliability of suppliers reduces their willingness to share data (Li & Lin, 2006). Other concepts that are closely related to trust, such as social interaction and shared vision, also influence the sharing of data between companies (Müller et al., 2020). The three factors of trust, social interaction, and shared vision can be summarized as social capital (Müller et al., 2020).

In situations where trust between companies is not sufficient to enable data sharing, '**control**' is an alternative (Bons et al., 2012). With the help of control mechanisms, data-sharing relationships can run securely and reliably, thereby restoring the lack of trust between companies (Bons et al., 2012; Agahari et al., 2022). Control helps to ensure that the companies involved in data sharing work towards their common goals and do not just pursue their own interests (Agahari et al., 2022). Companies can apply different control mechanisms depending on their position in the data ecosystem (Lis & Otto, 2020). Here a distinction can be made between 'contract-based control', 'structural-based control', and 'technology-based control' (Agahari et al., 2022). In contract-based control, data-sharing agreements, contracts, or authorization mechanisms are used to regulate the relationship between data providers and users concerning data access and use (Agahari et al., 2022). The specific design of such contract-based control mechanisms influences the willingness of companies to share data (Holler et al., 2019). This is because well-formulated contracts limit the scope for opportunistic behavior by the other party (Poppo & Zenger, 2002). 'Structural-based control' refers to the organization of relationships between data owners and data users, whereby data

either remains with the owner and only limited access is granted or is shared in a closed ecosystem (Agahari et al., 2022). If technical solutions such as anonymization, encryption, and aggregation are used to exercise control, this is referred to as 'technology-based control' (Agahari et al., 2022). However, one disadvantage of technology-based control is that such control mechanisms can also lead to restrictions on the usability of data (Agahari et al., 2022).

In the previous sections, the factors of risk, trust, and control have been considered individually. Although they are three different concepts, they are closely related and influence each other (Agahari et al., 2022). Trust is the degree to which one party is willing to make itself vulnerable to the actions of another party (Mayer et al., 1995). This means that higher trust leads to taking greater risks (Mayer et al., 1995). Strengthening control mechanisms reduces the risk of data sharing and increases trust between the parties (Agahari et al., 2022). However, control mechanisms can also increase the risk of opportunistic behavior, while non-mediated forms of power, such as trust, reduce this risk (Handley & Benton, 2012). Therefore, a combination of a trust-based relationship between companies and appropriate control mechanisms complement each other perfectly to reduce the risk of opportunistic behavior (Poppo & Zenger, 2002).

### 4.1.3 The Role of Interoperability in Cross-Company Data Sharing

The future success of companies will depend heavily on collaboration with other companies (Coltman et al., 2015). Ensuring a strong network between organizations is key to innovation (Coltman et al., 2015). Therefore, when strategically aligning their IT, organizations should focus on how they can best collaborate with different IT resources and capabilities (Coltman et al., 2015). A crucial point here is to seamlessly integrate information flows between organizations (Coltman et al., 2015). In this context, '**interoperability**' plays a central role, especially when it comes to data sharing between companies. It's a fundamental prerequisite and influencing factor for sharing data within a federated environment (Bastiaansen et al., 2020; Fassnacht et al., 2023). A lack of interoperability due to poor quality of internal data or a fragmented status of internal information systems lowers the willingness to share data (Dahlberg & Nokkala, 2019). Interoperability can be defined as "The ability of two or more systems or components to exchange information and to use the information that has been exchanged." (IEEE, 1990, p.42). This means data sharing is only possible if a certain degree of interoperability is achieved (Bastiaansen et al., 2020). Three different approaches are described for achieving interoperability between systems (Curry, 2012). The 'integrated

approach' defines a common data format that all systems must implement. In the unified approach, there is a superordinate model that translates the different data formats of the systems (Curry, 2012). The federated approach translates or adapts the data when it is shared between the systems in real time (Curry, 2012). Each of these approaches has specific advantages and disadvantages in terms of complexity and implementation costs (Curry, 2012). The biggest challenge for organizations is to create interoperability that enables successful long-term data sharing while minimizing complexity and cost (Curry, 2012).

However, a standardized data format is not the only important factor for a successful interoperability strategy. The overarching factor of interoperability can be divided into organizational, technical, semantic, and legal aspects (Bastiaansen et al., 2020). Organizational interoperability is the degree to which companies adapt and align their processes and structures to work together more efficiently and to achieve common goals better (Bastiaansen et al., 2020). Technical interoperability describes the ability for different applications to interact smoothly with each other, for services to be interconnected, for infrastructures to be compatible with each other, and for communication processes between systems to run securely (Bastiaansen et al., 2020). Technically interoperable systems can correctly interpret and implement formal agreements and contracts that regulate how data may be shared and used between parties (Bastiaansen et al., 2020). Too many actors in the network increase complexity and make technical interoperability difficult to implement (Bastiaansen et al., 2020). Technical interoperability is particularly emphasized as a challenge in the healthcare sector (Witte et al., 2020). Due to several individual technical solutions and many different network actors with different interests regarding data usage, there is a lack of uniform data standards for the sharing of health data (Witte et al., 2020). This leads to reduced data sharing and slows down innovation in this area (Witte et al., 2020). Technical interoperability is also emphasized as a challenge in industrial manufacturing. Machines often outlive technological developments due to their long life cycles, resulting in compatibility problems with new IT systems and communication standards (Lis & Otto, 2020). In addition, different degrees of machine automation increase complexity (Müller et al., 2018). To ensure technical interoperability, outdated machines must be modernized, and the recorded machine data must remain consistent and usable despite changing technologies and standards (Müller et al., 2018; Lis & Otto, 2020). Legal interoperability is used to design and harmonize legal regulations, guidelines, and contracts in such a way that cooperation and data sharing between these organizations are made possible without legal conflicts arising (Bastiaansen et al.,

2020). This is achieved, among other things, through the creation of data-sharing agreements and user contracts (Bastiaansen et al., 2020). One challenge in this context, for example, is the unclear ownership of machine-generated data or data analyzed by third-party providers (Lis & Otto, 2020). This can lead to uncertainties in the use of data and shows that existing legal regulations need to be further developed to ensure legal interoperability (Lis & Otto, 2020). Semantic interoperability allows data to be shared in such a way that the format and meaning of the shared data is preserved and correctly understood by all parties involved (Bastiaansen et al., 2020). This ensures that companies from different environments understand each other (Bastiaansen et al., 2020).

### 4.1.4 The Influence of Internal Corporate Factors on Data Sharing

In addition to interoperability, which particularly refers to system compatibility between organizations, factors within organizations also influence data-sharing behavior. '**Data governance**', '**top management**', '**resources**', and '**data-type**' can be identified here. These are summarized in this paper as internal factors.

'**Data governance**' can be described as a comprehensive framework that aims to manage data as a strategic asset within an organization (Abraham et al., 2019). It includes the definition of decision-making rights, responsibilities, and the development and monitoring of policies, standards, and procedures for the secure and efficient use of data to enable cross-organizational collaboration (Abraham et al., 2019). Although data governance has external dimensions, its foundation is anchored in the internal processes and structures that enable an organization to participate effectively and securely in inter-organizational data activities (Lis & Otto, 2020). It can, therefore, be considered an internal factor. Intra-organizational data governance is emphasized as the basis for successful data sharing (Lis & Otto, 2020). This is because new challenges arising from data sharing, such as the lack of rules and processes for handling external data, creating incentives for others to share data, and identifying useful and profitable data, can be overcome through well-structured internal data governance (Lis & Otto, 2020). Another challenge is the usually unclear ownership of data, which becomes even more complex when data is used across systems (Fassnacht et al., 2023). This leads to uncertainties about who is responsible for maintaining the data, who is authorized to access it, and who is ultimately authorized to make decisions (Fassnacht et al., 2023). Data governance can be used to define clear responsibilities and accountabilities for data access and use (Fassnacht et al., 2023). This can strengthen the social system within the organization and

thus increase the sense of psychological ownership among employees (Cheng & Du, 2015). The feeling of psychological ownership, in turn, has a positive effect on the use, management, and sharing of data (Cheng & Du, 2015).

Another internal factor is the '**top management**' of a company (Fassnacht et al., 2023). If a company's top management understands and supports data sharing and its benefits, this has a positive effect on the sharing of data across organizations (Li & Lin, 2006). With long-term support from top management, data sharing can be expanded in a targeted manner, and associated internal processes and external user experiences can be improved (Enders & Benz et al., 2020). In contrast, the lack of support for data sharing from the top management results in a decline in support from middle management and thus harms data sharing (Enders & Benz et al., 2020). Through inter-organizational relationship management, the management level can actively promote data sharing. Good relationships between managers of different companies strengthen trust and thus have a positive influence on data sharing (Wang et al., 2014). The development of a data strategy and embedding it in an overarching corporate strategy by top management can also help to develop a better understanding of the potential benefits of data sharing (Enders & Benz et al., 2020). This is particularly important because the benefits of data sharing are often difficult to recognize in companies (Gelhaar & Otto, 2020). Furthermore, embedding the data strategy into an overarching corporate strategy is of great importance, as responsibility for data is often incorrectly assigned to a separate department within an organization (Peppard, 2018). However, the value created from data can only be optimized if data management is integrated into every department of the company and data from every area can be used (Peppard, 2018). New data governance concepts such as Data Mesh attempt to implement this idea by giving individual departments greater responsibility for their data and thus managing data in a more decentralized manner (IBM, 2023).

The availability of '**resources**' is another important internal factor that influences the willingness of companies to share data. Both human and financial resources have a positive influence on collaboration between companies and data sharing (Müller et al., 2020). In particular, a lack of competent employees can reduce the willingness to share data, as the technical requirements involved are considered very complex (Dahlberg & Nokkala, 2019). Companies often lack the necessary financial and human resources to share data securely with other companies (Müller et al., 2018). The lack of specialists in this area can be illustrated by a study by IBM (2024). It states that cybersecurity teams in organizations are constantly

understaffed and that the shortage of cybersecurity experts in 2024 has even increased by 26.2% compared to the previous year (IBM, 2024).

Another internal factor is the '**data type**'. It has been shown that decision criteria regarding data sharing are weighted differently depending on the characteristics of the dataset (Enders et al., 2020). Some dataset metrics correlate strongly with certain decision criteria. For example, the proximity of the data to the core business (coreness) influences how critically companies assess the potential loss of their competitiveness if they share this data (Enders et al., 2020). In contrast, however, this metric shows no correlation with the decision criterion of data protection, which describes the risk of data being traced back to individuals or companies (Enders et al., 2020). Other important metrics that can significantly influence the decision to share data include the up-to-dateness of the data (currentness), the extent of the shared data set (extent), the level of detail (granularity), the interoperability and data quality, structure and availability (Enders et al., 2020; Fassnacht et al., 2023). At the same time, however, it can also be stated that the respective relationships between dataset metrics and decision criteria also depend very much on the use case and content of the dataset (Enders et al., 2020). The relationships between dataset metrics and decision criteria change depending on the purpose for which the dataset is used and the data it contains (Enders et al., 2020).

More specifically, it can be stated that data about competitive advantages, skills, and expertise is usually shared less to prevent customers from copying and passing on this data (Dahlberg & Nokkala, 2019). Data that reveals prices, profit margins, and costs is also shared less (Dahlberg & Nokkala, 2019). Planning data, on the other hand, was considered shareable by most, as was the automatic transmission of invoices and payments (Dahlberg & Nokkala, 2019). The benefits of sharing planning data are seen in the associated process optimization and adherence to deadlines. The sharing of invoice and payment data forms the basis for the automation of supply chains and is, therefore, of great importance (Dahlberg & Nokkala, 2019). When deciding about sharing less sensitive data, such as planning and capacity data, factors, such as financial and human resources, as well as existing IT connections, play a key role (Müller et al., 2020). The decision to share more sensitive data, such as production-, and process data, is more strongly influenced by factors like trust, shared vision, or benefits (Müller et al., 2020). Good IT connections and benefits are particularly important for design data, as these are usually more difficult to share due to different formats, and a quid pro quo is expected for sharing potentially critical trade secrets (Müller et al., 2020).

### 4.1.5 The Influence of External Factors on Data Sharing

External factors such as '**region**', '**sector**', '**external pressure**', and '**situational opportunity**' also influence companies' decisions to share data.

'**External pressure**' positively affects the willingness to share data (Dahlberg & Nokkala, 2019). This can arise, for example, from the growing demands of customers who expect optimized inter-organizational processes (Dahlberg & Nokkala, 2019). In addition, regulatory requirements such as data protection regulations also generate external pressure (Fassnacht et al., 2023; Holler et al., 2019). The emergence of trends and new developments can also generate external pressure (Fassnacht et al., 2023).

The '**sector**' in which a company operates also influences its willingness to share data with other companies. Highly competitive sectors are usually more reluctant to share data (Hoffmann et al., 2020). For example, companies in the automotive industry often aim to achieve a monopoly position in the market and, therefore, want to avoid the risk of losing their unique selling points through knowledge spillovers when sharing data (Agahari et al., 2022). Nevertheless, many studies on data sharing specialize in the automotive industry, as it is characterized by high market uncertainty, which also increases the importance of data sharing (Dyer & Chu, 2003; Agahari et al., 2022). Furthermore, it can be stated that the sensitivity of data types varies for different sectors (Dahlberg & Nokkala, 2019). For example, certain types of data that are often shared in the biorefinery sector are not shared in the maritime sector due to their high sensitivity (Dahlberg & Nokkala, 2019). In general, some sectors, such as healthcare, operate with much more sensitive and more regulated data than other sectors (Witte et al., 2020). So, data sharing poses a greater challenge here (Witte et al., 2020).

In addition to the sector, the '**region**' also affects the objectives and influencing factors of data sharing and the willingness to share data (Müller et al., 2020). For example, German suppliers are less aware of the benefits of sharing data than other regions. In contrast, US providers are very focused on the benefits and goals of data sharing (Müller et al., 2020). Especially in Germany, the factor of trust promotes the willingness to share data, while in China social interaction-, and in India shared vision and resources are considered more important (Müller et al., 2020). Overall, it can be seen that Chinese companies are the most open to collaboration and data sharing (Müller et al., 2020). Therefore, the introduction of test applications has the best chance of being accepted in this region (Müller et al., 2020).

A final external factor is referred to as '**situational opportunity**' (Dahlberg & Nokkala, 2019). Unexpected situations can promote the willingness to share data (Dahlberg & Nokkala, 2019). The current global situation or new technological developments can put companies in situations where the decision to share data could bring new opportunities (Dahlberg & Nokkala, 2019). This situational factor also includes the actions of a company's environment. Only if companies can recognize successful data-sharing processes in other companies will they decide to use them themselves (Dahlberg & Nokkala, 2019).

## 4.2 The Influence of Privacy Enhancing Technologies on Data Sharing Factors

The cross-organizational use of sensitive data requires new ethical, technical, and legal approaches to ensure that sensitive information remains protected (Müller et al., 2018). Privacy Enhancing Technologies are one way to make data sharing between companies more efficient and, at the same time, more secure (Scheibner et al., 2021). PETs could fundamentally change the definition and significance of the initial influencing factors for successful data sharing (Agahari et al., 2022). In the following sections, the influence of PETs on the previously identified factors is presented.

The figure below illustrates the factors that influence PET-based data sharing between companies. Compared to corporate data sharing without PETs, two factors (interoperability and data governance) that could not be identified in connection with PETs were removed, and one new factor (organizational readiness) was added.

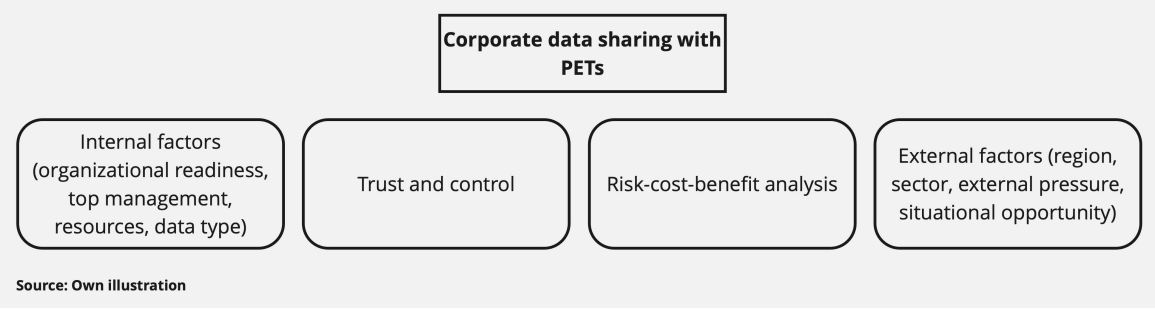


*Figure 2: Factors influencing the sharing of data with PETs*

### 4.2.1 Influence of PETs on the Factors of Trust, Control, and their Interactions

By using PETs, companies can process data jointly while ensuring data privacy at the same time (Garrido et al., 2022). This makes it possible to retain '**control**' over one's data even when sharing data (Agahari et al., 2022). Of the various control options, the integration of

PETs specifically strengthens ‘technology-based control’ (Agahari et al., 2022). In addition to actual control, perceived control is also strengthened (Agahari & Reuver, 2022). It is striking that the specific type of Privacy Enhancing Technology does not influence the increase in perceived control (Agahari & Reuver, 2022). Each Privacy Enhancing Technology increases the perceived control to the same extent (Agahari & Reuver, 2022). When, on the other hand, data is shared via trusted third parties without the integration of PETs, no increase in perceived control can be determined (Agahari & Reuver, 2022).

When companies can better control their data-sharing activities with other stakeholders, this also increases ‘**trust**’ between the parties (Agahari et al., 2022). Therefore, PETs also positively affect trust between companies by creating better control options (Agahari & Reuver, 2022). In initial relationships between companies, where comparatively little information is known about the other company, even the appearance of good control mechanisms can create an illusion of trustworthiness (McKnight et al., 1996). Especially if the benefits of the data-sharing relationship are considered to be high, this reinforces the illusion of trustworthiness (McKnight et al., 1996). This deceptive trust is also evident in the comparison of real and nonexistent Privacy Enhancing Technology (Agahari & Reuver, 2022). Both technologies increase the trust of the participants equally (Agahari & Reuver, 2022). The specific type and functionality of the Privacy Enhancing Technology is not decisive (Agahari & Reuver, 2022). It can even be observed that trust among users is higher if the exact functioning of PETs is not understood (Wen et al., 2023). This is because a better understanding of the technology leads to greater consideration of potential risks, which can reduce trust (Wen et al., 2023).

Furthermore, PETs are changing the way companies interact with each other and, thus, also the definition of trust between companies (Lumineau et al., 2020). The need for interpersonal trust between actors when sharing data is increasingly being replaced by trust in a system and the technology it uses (Lumineau et al., 2020). Without the integration of PETs, interpersonal trust has been identified as a decisive factor for successful cross-organizational data sharing (Gelhaar & Otto, 2020; Dyer & Chu, 2003; Müller et al., 2020; Li & Lin, 2006). However, if data is shared with the help of PETs, trust between the actors loses relevance (Agahari et al., 2022). It is no longer necessary to trust other companies or a central party to monitor data sharing (Körner et al., 2022; Agahari et al., 2022). Instead, trust in the technology becomes more important, as the security of data sharing now depends mainly on the correct functioning of the technology (Agahari et al., 2022). Despite an increase in the importance of trust in

technology, trust in people or organizations is not eliminated (Lumineau et al., 2020). There remains a need to trust developers who have designed the technology and actors who supply the system with data to a certain extent (Lumineau et al., 2020).

### 4.2.2 The Influence of PETs on the Trade-off Between Risks, Costs, and Benefits

The factors of control and trust are closely related to risk (Agahari et al., 2022). While increased trust can lead to taking greater risks, an improvement in control can reduce risks (Mayer et al., 1995; Agahari et al., 2022). Consequently, by influencing trust and control, PETs also change the factor of risk when sharing data (Agahari et al., 2022).

On the one hand, the perceived '**risk**' of sharing data can be reduced by integrating PETs (Agahari & Reuver, 2022). On the other hand, actual risks can also be reduced. In particular, the risk of data misuse can be reduced without having to accept restrictions on the usability of the data (Garrido et al., 2022). When deciding on suitable control measures, companies usually have to think about the trade-off between data usability and data misuse risk (Agahari et al., 2022). This is because although greater control can reduce the risk of data misuse, it also reduces the usability and, thus, the value of the data (Agahari et al., 2022). By integrating PETs, this trade-off can be avoided to a certain extent (Garrido et al., 2022). In addition, the competitive risk, which is seen as one of the greatest risks when sharing data, can also be reduced by integrating PETs (Agahari et al., 2022). By maintaining the confidentiality of the data, knowledge spillovers can be prevented, which limits the risk of losing competitive advantages against other companies (Agahari et al., 2022; Garrido et al., 2022). Also, the risk of violating the privacy of end users and damaging the company's image can be reduced through the use of PETs (Agahari et al., 2022). Furthermore, the risk of the so-called copying problem can be avoided (Garrido et al., 2022). By using PETs, data is not disclosed even during processing (Zhao et al., 2019). This means that companies are no longer exposed to the risk of data being copied and resold when sharing data, thereby losing the value of its uniqueness (Garrido et al., 2022). Data owners can grant other companies one-time access to data for agreed calculations without them being able to copy the data or use it for non-agreed purposes (Garrido et al., 2022). The value of the data is thus preserved for the original owner (Garrido et al., 2022).

However, other risks remain or shift to other actors in the data ecosystem. The original risks associated with the disclosure of sensitive data can now shift from the data owner to the data user with the integration of PETs (Agahari et al., 2022). This is because PETs only protect the

input privacy of the data owner, but not the output privacy (Choi & Butler, 2019). As a result, there is a risk for data users that the data owners will find out their strategic interests by viewing the data queries (Agahari et al., 2022). At the same time, however, there is also a risk for data owners that the results of the queries can be used to draw conclusions about the underlying data (Garrido et al., 2022). A combination of different PETs can help to prevent these re-identification attacks (Garrido et al., 2022).

Some of the risks that arise from data sharing and are changed by PETs have already been highlighted. At the same time, however, there are other non-data-sharing related risks in companies that could be reduced by sharing data with PETs. In the energy sector, increased PET-based data sharing between different actors in the energy market could help to identify hidden system risks (Körner et al., 2022). In the now very decentralized energy sector, with many small, distributed players, the various parties can merge their data in encrypted form and perform calculations on it without giving competitors an advantage (Körner et al., 2022). By pooling data, developments can be continuously monitored, and the risk of a power grid collapse can be reduced (Körner et al., 2022). In the financial sector, the increased sharing of data through the integration of PETs makes it possible to better identify fraud (Sangers et al., 2019). The use of PETs can also reduce the risk of cyberattacks (World Economic Forum, 2020). For individual companies, it is very difficult to recognize all cyber threats in a rapidly changing digital world (World Economic Forum, 2020). PETs could make it possible for companies to share cyber information while protecting this highly sensitive data (World Economic Forum, 2020).

In terms of '**costs**', the use of PETs can, on the one hand, save the expenditure for a central party that processes the data (Kanger & Pruulmann-Vengerfeldt, 2015). On the other hand, PETs also require clear regulations to ensure the quality of data input. This leads to coordination costs for the parties involved (Kanger & Pruulmann-Vengerfeldt, 2015).

The '**benefits**' factor is also of great importance when sharing data with PETs, as companies ultimately always want to maximize their profits (Agahari et al., 2022). If the integration of PETs provides clear benefits that increase the company's profits and, at the same time, the perceived costs of adoption are low, the decision is made to share data (Agahari et al., 2022; Hasani et al., 2023; Kanger & Pruulmann-Vengerfeldt, 2015). As companies are reluctant to take the risks of initial adoption, convincing use cases that demonstrate the benefits of PET-based data sharing are crucial for the adoption of PETs (Kanger & Pruulmann-Vengerfeldt,

2015). Benefits can include personalization, service improvements, productivity increases, or direct monetization through data sales (Agahari et al., 2022; Hasani et al., 2023). However, safe benchmarking, predictive maintenance, and carbon footprint calculations are also specific areas of application for PETs that could provide benefits for companies (SAP, 2024). PETs can also extend the functionality of data marketplaces beyond pure data sharing (Agahari et al., 2021). They act as so-called boundary resources that enable different actors to work together securely on a shared platform without privacy concerns (Agahari et al., 2021). One example would be a platform provided by a hospital: Third-party providers can develop innovative healthcare applications on it by accessing the hospital's data without ever having direct access to the raw data (Agahari et al., 2021). This enables completely new applications (Agahari et al., 2021).

### 4.2.3 Internal Factors in Connection with PETs

Data governance is emphasized as a fundamental factor for the successful data sharing between companies (Lis & Otto, 2020; Fassnacht et al., 2023). In the context of PETs, data governance is not specifically highlighted, but '**organizational readiness**' is emphasized as a condition for the successful use of PETs (Agahari et al., 2022; Hasani et al., 2023). This includes a clear data governance structure and, in particular, data pre-processing capabilities such as data cleaning or data harmonization (Agahari et al., 2022). If this is not in place, companies will find it difficult to exploit the benefits of PETs (Agahari et al., 2022). Similarly, 'organizational fit' is also mentioned as a condition for the adoption of PETs (Kanger & Pruulmann-Vengerfeldt, 2015). This describes how easily PETs can be integrated into existing organizational processes (Kanger & Pruulmann-Vengerfeldt, 2015). If PETs can be easily included in existing routines, work processes, rules and regulations, this has a positive effect on the adoption of PETs (Kanger & Pruulmann-Vengerfeldt, 2015).

The influence of '**top management**' becomes even more important for corporate PET-based data sharing. The support, innovative spirit, and attitude of top management have a significant influence on the decision to use PETs and their successful implementation (Hasani et al., 2023). Similarly, management can also be one of the biggest obstacles to the introduction of PETs if they are not open to this type of technology (Bamford, 2020). Often, changes in a company also increase the willingness to introduce new technologies (Kanger & Pruulmann-Vengerfeldt, 2015). A recent change in management can, therefore, positively affect the adoption of PETs (Kanger & Pruulmann-Vengerfeldt, 2015).

In terms of the '**data type**', companies are usually more cautious when sharing sensitive, core business-relevant data, as this entails higher risks (Dahlberg & Nokkala, 2019). Even with the integration of PETs, this caution remains when sharing sensitive data. Data that is considered sensitive-, relates to the core business-, or is of strategic importance is shared only hesitantly, even with the integration of PETs (Agahari et al., 2022). This kind of data is only shared with a high level of trust in the other parties (Müller et al., 2020). However, the sharing of generic and non-sensitive data can be encouraged through the integration of PETs (Agahari et al., 2022). Here, PETs are perceived as useful and support increased data sharing (Agahari et al., 2022). In contrast, other literature emphasizes the importance of data sensitivity when integrating PETs (Kanger & Pruulmann-Vengerfeldt, 2015). The use of PETs only makes sense if the shared data is sensitive enough and regular cross-company data processing is necessary (Kanger & Pruulmann-Vengerfeldt, 2015). Furthermore, data quality becomes more important when data is shared with PETs (Kanger & Pruulmann-Vengerfeldt, 2015).

It has already been established that the availability of '**resources**' has a positive influence on data sharing between companies (Müller et al., 2020). Human resources are becoming even more important for PET-based data sharing (Hasani et al., 2023). This is because PETs are a very complex subject area, even for IT specialists (Agrawal et al., 2021; Körner et al., 2022). However, a clear understanding of the technology is a prerequisite for adoption (Kanger & Pruulmann-Vengerfeldt, 2015). The "ease of use" of PETs is, therefore, also described as a key factor for their adoption (Hasani et al., 2023). The financial resources required for corporate PET-based data sharing are difficult to estimate (Kanger & Pruulmann-Vengerfeldt, 2015). However, it could be that only large companies have the financial resources available for the costs associated with the implementation of PETs (Kanger & Pruulmann-Vengerfeldt, 2015). So, a lack of both financial resources and experienced personnel could hinder PET-based data sharing (Hasani et al., 2023).

### 4.2.4 External Factors in Connection with PETs

'External pressure' on data-sharing behavior can be generated by regulatory requirements, emerging trends, or growing customer demands (Dahlberg & Nokkala, 2019; Fassnacht et al., 2023; Holler et al., 2019). Perceived external pressure can positively influence the intention to integrate PETs into data-sharing processes (Hasani et al., 2023). Therefore, providers of PETs should focus particularly on companies that are under external pressure to change (Kanger & Pruulmann-Vengerfeldt, 2015). Regulatory requirements are one form of external pressure

that can positively affect PET-based data sharing, as PETs could help to meet these regulatory requirements (Scheibner et al., 2021). They could be a solution to the EU GDPR's goal of strengthening data protection while facilitating the free movement of personal data (Helminger & Rechenberger, 2022). The use of PETs can reduce the risk of non-compliance with the GDPR when data is shared (Scheibner et al., 2021). One example is the principle of data minimization, which is prescribed in the GDPR (Bamford, 2020). It says that the collection, use, or disclosure of personal data should be limited to what is strictly necessary. PET-based data sharing fulfills this requirement and could, therefore, be a means for companies to meet regulatory requirements (Bamford, 2020). However, other forms of external pressure can also negatively affect the decision to share data with PETs (Kanger & Pruulmann-Vengerfeldt, 2015). Especially in larger companies, there is usually pressure from the outside not to take unnecessary risks. When risks such as the introduction of new technologies are taken, they always have to be justified first (Kanger & Pruulmann-Vengerfeldt, 2015). This can inhibit the adoption of PETs (Kanger & Pruulmann-Vengerfeldt, 2015).

The influence of the ‘sector’ changes for PET-based data sharing. In healthcare, for example, data is highly sensitive and regulated, so sharing data without PETs is a challenge (Witte et al., 2020). However, data sharing involving PETs is more likely to occur in the healthcare sector than in other sectors (Witte et al., 2020; Hasani et al., 2023). In general, sectors that deal with larger amounts of information have a higher adoption rate of PETs (Hasani et al., 2023). Of the IT, services, manufacturing, finance, and wholesale sectors, the IT and service sectors (healthcare, consulting, retail, and telecommunications) are most likely to adopt PETs (Hasani et al., 2023). Industry structures & standards also affect the adoption of PETs (Hasani et al., 2023). Uniform standards within industries and low competition have a positive impact on the introduction of new technologies (Hasani et al., 2023). According to a study, the Internet ($5.1 billion), software ($1.7 billion), electronics ($1.5 billion), healthcare ($429 million) and industrials ($189 million) sectors invested the most in PETs in the years 2021-2023 (World Economic Forum, 2024).

‘Regional differences’ can also be observed concerning investments in PETs. The USA provided the most business funding in this area with $4.6 trillion (World Economic Forum, 2024). China is in second place with $1.6 billion in business funding, followed by the Netherlands with $1.2 billion (World Economic Forum, 2024).

As resources are an important influence for PET-based data sharing, the ‘situational opportunity’ in the form of the macroeconomic situation also plays a decisive role, as it influences the availability of these resources (Kanger & Pruulmann-Vengerfeldt, 2015). In addition, organizational changes within the company can offer an opportunity for the integration of PETs, as such restructuring often creates new opportunities for innovation and the introduction of new technologies (Kanger & Pruulmann-Vengerfeldt, 2015).

# 5 Method of Interviews

Building on the knowledge gained in the structured literature review about factors influencing data sharing between organizations and the impact of PETs on these factors, semistructured interviews with experts were conducted. To uncover the subtle organizational dynamics in data-sharing decisions, interviews are cited as the most appropriate method (Smith et al., 2011). This method is intended to provide in-depth qualitative insights and flexibility in exploring specific topics. To this end, 20 interviews were conducted over two months. Potential candidates for interviews were contacted via LinkedIn SalesNavigator. The selection of experts was based on Meuser & Nagel's (1991) conditions for expert status. Both IT experts and specialists for Privacy Enhancing Technologies were selected to explore the decision-making process of companies when sharing data and the influence of PETs. Of the 20 interviews, ten were conducted with experts in the IT sector. Four of these were active in IT security-, and six in other areas such as data governance, cloud, and data spaces. The other ten interviews were conducted with experts in the field of Privacy Enhancing Technologies. These included five founders from currently leading European Privacy Enhancing Technology providers, four employees from companies that offer PETs, and one research expert in this field.

As suggested by Myers & Newman (2007), the interview guide included a brief introduction to the topic, key questions, and a suitable conclusion (see Appendix B). The key questions were divided into two major parts. The first part analyzed the general data-sharing behavior of companies and which factors influence it. In the second part, the focus of the questions was placed on Privacy Enhancing Technologies and their influence on data sharing. As advised by Myers & Newman (2007), a certain flexibility was applied in the interviews. This means that questions were adapted, omitted, or added based on the development of the conversation. Depending on the expertise, the focus of the interview was adapted to the respective knowledge of the expert.

At the beginning of each interview, consent was asked for the interview to be transcribed and used in this paper in anonymized form. To transcribe the interviews, Microsoft Teams and the Maxqda software were used. The transcripts only contain the question part, the introduction and conclusion were not transcribed. General information about the interview and the expert can be found in a table at the beginning of each transcript. The transcripts are divided into numbered sections to be able to refer to specific passages (see Appendix D).

The interviews were analyzed according to the method of Meuser and Nagel (1991). After transcription, the interviews were paraphrased, whereby the statements that were relevant to this work were written out in summarized form. Subsequently, headings were created for the interview statements, which allowed the statements to be further summarized. The headings of different interviews were compared with each other, divided into categories, and then assigned to a superordinate concept. The concepts were mainly drawn from the theoretical background of the systematic literature review. However, new concepts also emerged that could not yet be identified in the systematic literature review. The paraphrasing of the interviews, headings, and assigned concepts are shown in a table (see Appendix C). The results of the interviews are presented in the following chapter.

# 6 Results of the Interviews

Similar to the interview questionnaire, the results of the interviews are also divided into two parts. The first part relates to the research question of what factors companies weigh when deciding to share their data with other companies. The second part refers to the second research question about the influence of PETs on the factors influencing data sharing and the extent to which PETs promote more intense data sharing.

## 6.1 Factors that Influence Data-Sharing Decisions

The headings created from the paraphrased interview statements have been assigned to various concepts (see Appendix C). The majority of the concepts could be transferred from the systematic literature review. However, the concepts of ‘data governance’, ‘resources’, and ‘cost’ from the systematic literature review could not be found in the interviews. Instead, the new concepts of ‘digital infrastructure’ and ‘corporate culture’ were identified. The following figure shows the concepts identified in the interviews. Newly added concepts compared to the factors identified in Chapter 4.1 have been highlighted.

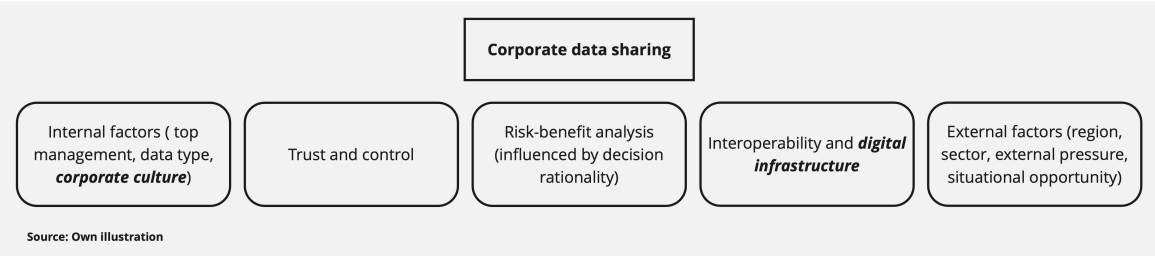


*Figure 3: From Interviews: Factors influencing data sharing*

### 6.1.1 Requirements for Successful Data Sharing

In the systematic literature review, interoperability was identified as a fundamental prerequisite for the sharing of data (Bastiaansen et al., 2020; Fassnacht et al., 2023). ‘**Interoperability**’ is also mentioned in the interviews as a prerequisite for cross-company data sharing (B1,24). Interoperability would be the tool to enable data sharing (B12,22). In particular, the implementation of technical interfaces would be an important prerequisite (B5,26; B8,17). If the interfaces of machines in companies were outdated, this would make data sharing more difficult (B5,26). In addition, data quality in the form of a uniform, standardized format is described in the interviews as a prerequisite (B1,24; B5,24; B19,14). A standardized data format is of great importance to achieve interoperability and is therefore

also assigned to the concept of interoperability (Curry, 2012). Furthermore, the '**digital infrastructure**' of a company, which is closely linked to interoperability, is mentioned as a prerequisite (B5,26; B19,14,16). Only if it is possible to share data within the company, data could also be shared across companies (B1,28).

In addition, '**control**' measures in the form of contracts would be an important prerequisite and essential component of data sharing (B2,8; B8,17; B7,8; B6,8). Contracts such as non-disclosure agreements would be agreed before any data-sharing activity (B2,8; B3,16; B6,8). By naming these 'contract-based control' mechanisms as an essential prerequisite, the interviews emphasize the great importance of this form of control for data sharing. In contrast, the importance of control is less emphasized in the literature. It is only described as a measure to support data sharing and establish trust (Bons et al., 2012; Agahari et al., 2022).

### 6.1.2 The Importance of Trust and Control for Data Sharing

Some interviewees see '**trust**' between companies as a very influential factor (B5,10; B8,7; B10,18; B12,6; B13,12; B15,17; B18,5). As decision-making processes in companies regarding the sharing of data are very subjective, mutual trust would play an important role (B10,18). Companies would often be dependent on assessing statements from third parties, and trust could be a decisive factor here (B13,12). Despite control mechanisms, it would be ultimately necessary to trust that the shared data will not be misused (B8,7). One interviewee describes Catena-X, a data ecosystem in the automotive industry, as an example of how data sharing is based almost exclusively on trust (B19,8). There would be no technical solutions to secure the data. No one would be able to prevent the parties involved from copying the data and using it for their purposes (B19,8). Furthermore, it is emphasized that, especially in medium-sized companies, data is often not shared due to a lack of trust (B5,10). Therefore, in this kind of company, there is the greatest discrepancy between the potential of sharing data and creating value from it and the actual implementation (B5,10). Various ways of generating trust were mentioned in the interviews. Partnership-based relationships on the management level or previous business relationships between companies could strengthen trust (B8,9; B18,5). But also technical expertise, proven by publications or titles, successful similar use cases, and positive experiences of other organizations with a potential data partner could strengthen the trust (B7,16; B13,14). If companies have things in common, such as the same country of origin or a common goal, this could also generate trust (B10,14). Just as in the literature (Bons et al., 2012; Agahari et al., 2022), the interviews emphasize control as a trust-

building measure. Contracts and inspection catalogs could create trust between companies (B8,7,9; B17,8). It would be important for companies to rely on the protection of their interests (B17,8). Both in the interviews and the literature, the importance of the design of contracts is highlighted in this context (Poppo & Zenger, 2002; B17,8).

In other interviews, trust is seen as less crucial. Companies would pay more attention to economic benefits and potential business risks than to trust (B20,6; B4,8,10). Instead of blind trust, companies would rely on trust based on a careful risk assessment (B17,8,23). The decision-making process for data sharing would depend on many different factors, such as company size, market position, and geographical location, rather than personal relationships or trust (B3, 28).

As already mentioned, '**control**' is seen by many experts as essential for data sharing (B1,12; B2,8; B3,16; B6,8; B7,8; B8,17; B9,8; B11,8; B20,7). Control through contracts that protect the interests of the company and guarantee the security of the data would play a greater role than trust (B1,12; B4,8,10). As companies generally do not trust each other, control would be the only way to enable cooperation between companies (B20,7).

### 6.1.3 The Trade-off Between Risks and Benefits when Sharing Data

The interviews indicate that the '**benefits**' are a decisive influencing factor for sharing data. Some interviews emphasized that the decision to share data is largely a trade-off between benefits and costs (B1,14; B20,9). There should be an economic benefit for companies to share their data (B4,24; B13,12; B15,35). This could be achieved either by reducing risk and lowering costs or by increasing sales (B15,35). The technology chosen for data sharing would only play a supporting role, while the decision to share data would be primarily based on economic considerations (B15,35). Economic advantages could include deeper insights into business processes and joint cost sharing, especially for complex and expensive tasks (B4,24; B12,38). The German automotive industry, for instance, would share data to carry out CO2 calculations, as in-house calculations are very costly (B12,38).

Privacy risk is the expectation of potential losses resulting from the sharing of data (Aleem et al., 2017; Xu et al., 2011). At an individual level, there is a tendency to overestimate potential losses and thus to be unable to rationally assess privacy risk (Kahneman & Tversky, 1979). To find out whether this theory could also apply to the corporate context, the interview question was asked whether decisions on data sharing in companies are made objectively and

rationally or whether it is a subjective process similar to individual decision-making behavior. The findings from this question were assigned to the overarching concept of '**decision rationality**'. The answers were very mixed. Interviewees who described data-sharing decisions as subjective explained this by saying that people need to be convinced to share data. They noted that the ability of the party offering the data to persuade others, as well as the subjective judgment of the receiving party, plays a significant role in the decision (B4,12). Another interviewee pointed out a use case in the financial sector in which there were no rational arguments against sharing data with competitors, and yet the data was not shared in the end (B7,10). This also shows that the corporate decision here was not made based on a rational risk assessment. Even with the help of technologies that circumvent privacy risks, there would always be a tendency to be suspicious of data sharing (B10,18). There would exist a culture of risk avoidance, where it is better not to share any data at all than to take any kind of risk (B10,28). It is also emphasized that outdated solutions are often retained due to regulations, even though there are better alternatives (B7,12). This caution in data sharing decisions is described by the saying, "You never get fired for buying IBM" (B7,14) and once again illustrates the subjective factor in data sharing decisions. Many other experts saw the decision-making process as a combination of subjective and objective elements. On the one hand, the corporate culture as a subjective component would influence the willingness to take new risks, such as sharing data (B12,10). On the other hand, the financial perspective on the benefits of data sharing would represent an objective component of the decision (B12,10). A progressive mindset and the courage to take the risk of sharing data would be required, but at the same time, the necessary compliance checks must be carried out within the company before data sharing can be implemented (B18,7). Although the decision-making process might be more objective due to the large number of parties involved, it would always involve people, meaning that a subjective element could never be entirely avoided (B15,19). One interviewee emphasized the current cautious and dismissive attitude towards the sharing of data and the associated overestimation of risks, which would be intensified by regulations such as the GDPR (B5,12). At the same time, however, he noted that we would find ourselves in a transition phase towards a more objective assessment in which rational considerations were becoming increasingly important (B5,12). However, some experts also saw the decision-making process as very objective and rational (B1,14; B3,20; B13,6; B17,10; B20,9). It is seen as a largely rational weighing up of costs and benefits (B1,14; B20,9). Decisions would always be reviewed by at least two people, but usually by many more people and different departments (B3,18; B13,16). There would usually be fixed parameters, such as the type of

business relationship, sensitivity of the data, or business area, to decide whether data should be shared (B17,10). It is also emphasized that employees would be reluctant to take individual responsibility and would, therefore, prefer to have decisions approved by several departments (B3,20). The decision-making process is seen as being as objective as possible, given that people are involved (B3,20). Furthermore, influences on decision rationality were highlighted in the interviews. One interviewee described the size of the company as a decisive factor influencing the degree of rationality in decisions about data sharing. The larger a company, the more rational the decision-making processes would be (B8,13). The top management of a company and the corporate culture would also influence rationality (B1,14; B9,14). Regardless of the degree of rationality in risk assessments, interviewees emphasized the difficulty of correctly assessing risks. The risks of data sharing wouldn't be tangible and, therefore, difficult to assess (B20,11,13). There wouldn't exist a framework for assessing risks, which means that they are often misjudged (B12,16).

In addition to the degree of rationality in the assessment of risks, the actual '**risks**' are also important when deciding to share data. In the interviews, three central risks were identified when sharing data: legal risks, reputational risks, and the risk of disclosing business secrets. The risks are in line with those highlighted in the literature (Agahari et al., 2022; Gelhaar & Otto, 2020). The legal risk in the form of fines due to legal violations in the handling of data is seen by many experts as one of the greatest risks (B1,20; B6,14,16; B7,18; B10,22,24; B12,14; B15,29; B18,9; B20,11). A major problem here would be that there are only a few contracts that have already been tested in court and whose effectiveness has been confirmed under real conditions (B5,22). Although legal action is theoretically possible in the event of data misuse, there would be a lack of standardized contracts that can be legally relied upon (B5,22). Greater legal certainty would lead to a reduction in legal risk (B5,22). In connection with this, the literature shows that perceived uncertainty leads to a higher perception of risk (Pavlou et al., 2007). Furthermore, the risk of a loss of reputation is also frequently mentioned in the interviews (B1,20; B7,18; B10,22,24; B12,14; B15,29; B20,11). One specific example of this is the sharing of data in the area of cybersecurity. Here, there would be a risk that shared sensitive data could reveal that a company has taken inadequate security precautions, which could lead to considerable reputational damage (B14,22). The third major risk highlighted in the interviews is disclosing business secrets (B1,22; B6,14,16; B10,22,24; B12,14; B18,9; B20,11). This would pose a particularly high risk in the case of company sales, as highly sensitive data must be disclosed to potential buyers without any certainty

about whether they will acquire the company (B2,18). The risk is further heightened by laws such as the USA Patriot Act, which expands the powers of US authorities to access data, even if it is stored abroad. This means that sensitive business information could potentially be accessed by foreign authorities under certain conditions (B6,14,15; Deutscher Bundestag, 2020). Another threat associated with data sharing is the loss of control over the data as soon as it is passed on (B16,11). Even if the other company had no malicious intentions, there would be a risk that they handle the data less carefully and thus reduce data security (B16,11). The company with which data is shared must always have implemented the same security standards as oneself (B8,11,15). In particular, the uncertainty caused by the loss of control over one's data is linked to the concept of post-purchase information asymmetry and its components, ‘hidden action’ and ‘hidden effort’ (Aleem et al., 2017). The fear that other companies will not act as contractually regulated and misuse the data can be classified as a ‘hidden action’. The perception that the other company is not investing enough in the security of the data describes the ‘hidden effort’ (Aleem et al., 2017). Both hurt the willingness to share data (Aleem et al.,2017).

However, in connection with the question about the risks of data sharing, the risk that arises when data is not shared is also mentioned (B9,9; B20,13). There would be a risk of missing out on opportunities for innovation if data is not shared (B20,13). For example, very few companies would own the type of data sets that are needed to implement artificial intelligence. Only through data cooperation would they be able to generate the required data volumes and thus successfully develop AI solutions (B9,9).

About the general consideration of the risks and benefits of sharing data, the value-risk dilemma is seen by one interviewee as a challenge (B17,10,12). The benefits of data are often difficult to assess, while the risks, such as misuse or loss, are clear and immediate. This uncertainty would often make companies reluctant to share data (B17,10,12). This is in line with the literature, which emphasizes that companies often do not recognize the benefits of data sharing (Dahlberg & Nokkala, 2019; Gelhaar & Otto, 2020).

### 6.1.4 The Influence of Internal Factors on the Willingness to Share Data

As part of the systematic literature review, the influence of data governance, the type of data, company resources, and top management on the willingness to share data was assigned to the category of internal factors. In connection with data governance, only one interviewee described a form of internal, decentralized data organization, the data mesh concept, as an

internal factor that increases data sharing within the company but restricts cross-company data sharing (B12,42). The interviewees did not mention the availability of company resources in the form of financial means or specialists as an influencing factor for the fundamental willingness to share data. However, the factors '**top management**' and '**data type**' were confirmed as concepts in the interviews. In addition, the concept of '**corporate culture**' was identified in the interviews.

'**Top management**' is seen as an important factor in the willingness to share data (B1,26; B7,20). It would be more likely that data will be shared if top managers are open to the topic or are well versed in it (B7,20). The influence of top management would be one reason why decision-making processes regarding data sharing contain a subjective component (B1,14). In general, decision-making structures within a company would influence the willingness to share data. A lot would depend on whether only one person in the company or a board of directors decides about data sharing (B5,26).

The '**data type**' would also play an important role in company decisions regarding data sharing. Depending on the category of data, there could be different regulatory requirements, which are defined in particular by legal frameworks such as the AI Act or the Digital Services Act (B1,58). As a rule, data without direct business value would be shared, while sensitive data, such as trade secrets, would be withheld (B3,24). Data in the area of cybersecurity could be shared more easily, as the focus here is on joint protection against threats (B3,24). In this context, companies would share data on threats (threat intelligence) in particular, as this would be beneficial for all parties involved without risking competitive disadvantages (B3,24).

One concept that was identified in the interviews but did not yet come up in the systematic literature review is the '**corporate culture**'. It would significantly influence the willingness to take new risks, such as sharing data between companies (B12,10). Non-technology-oriented companies would be more emotionally driven and less rational when making data-sharing decisions (B9,14). Technology-oriented companies would usually be more open in this area due to a better understanding (B9,14).

### 6.1.5 The Influence of External Factors on the Willingness to Share Data

‘**Region**’, ‘**sector**’, ‘**external pressure**’, and ‘**situational opportunity**’ were identified as external factors in the structured literature review. Each of these factors was also identified as a concept in the interviews.

‘**External pressure**’ from regulatory requirements was particularly emphasized in the interviews. Laws and regulations would be one of the strongest influencing factors on corporate data sharing (B15,37). Whether this influence is positive or negative would depend on the respective law or regulation. The European General Data Protection Regulation (GDPR), for example, would have led many companies to act cautiously and refrain from sharing data to avoid legal risks (B5,10). In most cases, laws are formulated very obscurely, which would create uncertainty for companies (B6,28). In response to this uncertainty, companies' internal policies would often implement the laws even more strictly than required to be legally on the safe side (B10,28). This would also make data handling processes very complex and lengthy (B10,28). However, some laws positively influence data sharing. The European Union's Data Act legislative proposal was highlighted in some interviews (B5,14, B12,40,42; B13,12). This would force companies to make data accessible or share it in certain situations (B5,14, B12,40,42; B13,12). The Supply Chain Due Diligence Act is also described as a major driver of data sharing between companies (B5,4). It obliges companies to ensure human rights and environmental standards throughout their supply chain (Haupt et al., 2021). This requires increased data sharing to identify risks and take appropriate measures (Haupt et al., 2021). Indirectly, however, it would also be a welcome means for companies to obtain valuable data from their suppliers to better control them and spur competition (B5,8). Furthermore, the restrictions on the use of third-party cookies would also lead to changes in data-sharing behavior (B4,6). Companies would now be forced to use alternative ways, such as data sharing between advertisers and publishers, to gain insights into the efficiency of advertising campaigns (B4,6; B9,11). In the pharmaceutical industry, new laws could allow real-world evidence data to be used for the approval of new drugs (B9,11). This could also increase data sharing in the healthcare sector (B9,11).

The ‘**sector**’ in which a company operates would also influence the extent to which data is shared. Different sectors would have different degrees of competition, cooperation, and regulation (B10,16). Therefore, data-sharing behavior would differ depending on the sector (B10,16). The literature particularly emphasizes healthcare as a sector where data sharing

poses significant challenges due to the sensitive and highly regulated nature of the data (Witte et al., 2020). Also, in the interviews, the healthcare sector is described as highly regulated (B5,28). However, at the same time, a great deal of data would be shared here, for example, between hospitals and insurance institutions (B10,10). Furthermore, the financial sector is identified as another highly regulated sector (B5,28). The risk of reputational damage as a result of data sharing is seen as particularly critical here (B7,22). However, the healthcare and financial sectors, as the most regulated sectors, would also offer the greatest potential for data sharing (B7,24).

'**Situational opportunities**' would also influence companies to share data. For example, the German automotive industry would share data due to strong international competition (B3,50; B6,20). Common goals, for example, against competitors or to fulfill legal requirements, would make data sharing sensible (B3,50). Another example is the demographic trend in industrialized countries toward an aging population, which means that more healthcare is needed (B16,4,13). This would lead to more data being shared in this area to make care more efficient (B16,4,13). The examples from the interviews support the literature, showing that current global circumstances place companies in unexpected situations where data sharing can create new opportunities (Dahlberg & Nokkala, 2019).

Furthermore, the '**region**' is also highlighted in the interviews as an influence to share data. For example, companies in the Netherlands would be more open to innovative data collaboration technologies, while in Germany, data protection is a top priority (B18,25). In addition, it is emphasized above all that regional differences in data regulations would represent a major challenge for cross-country data sharing (B19,6).

## 6.2 The Impact of PETs on Data Sharing and its Influencing Factors

In the interview analysis, the interviewees' statements about the influence of PETs on data sharing were assigned to superordinate concepts (see Appendix C). For the most part, existing concepts from the systematic literature review were identified (see Chapter 4.2). However, the new concepts of 'size' and 'corporate culture' were added. In contrast, the concepts of 'organizational readiness', 'data type', and 'costs' were not mentioned in the interviews. The following figure shows all concepts that have been identified about the influence of PETs on data sharing. Newly added concepts were highlighted.

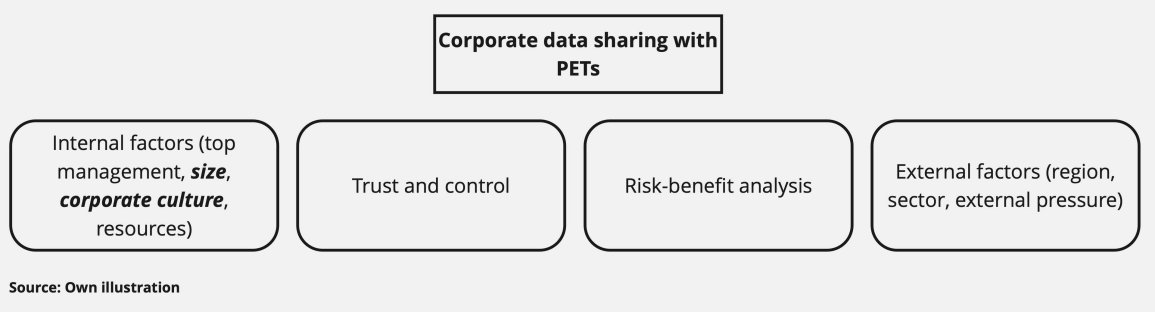


*Figure 4: From Interviews: Factors influencing the sharing of data with PETs*

### 6.2.1 Benefits and Risks of PET-Based Data Sharing

When deciding whether to share data with PETs, the benefits that can be derived from it play a decisive role for companies (Agahari et al., 2022; Hasani et al., 2023). The concept of '**benefits**' was also identified in the interviews. Some interviewees do not see any real use cases in PETs that could bring benefits. Companies would initially be reluctant to use PETs, as the benefits are often unclear (B1,36; B2,4; B3,38). There are thousands of interesting technologies out there, but the decisive factor would be that they can also be used in practice and come with benefits (B3,38). For PETs there would be a lack of concrete use cases that could be solved (B2,30). When sharing data, it would usually be important to be able to view the data. Not much could be done with data encrypted by PETs (B2,4). In many cases, the costs would outweigh the benefits. Particularly, large companies, which already have extensive databases, would see little added value in using external data sources or sharing data with other companies (B3,40). Instead, the focus would be on optimizing internal data processes to use their data more efficiently (B3,40). In most cases, companies wouldn't be interested in collaborating with competitors. They would rather invest in their solutions than share data with competitors and make progress together (B3,44).

Despite these negative aspects, some interviewees also show how PETs could change and expand data sharing and the associated benefits. In the future, more data would be shared, as the value of data could only be fully exploited by combining it (B15,43; B16,15). PETs could help to do this in a privacy-preserving way (B16,15). They would, therefore, gain importance in the future (B15,43). PETs could make it possible to share data that was previously impossible or very risky (B10,30). As soon as companies realize that they can share sensitive data securely with PETs, new data collaborations will emerge, or existing collaborations will be made more efficient (B9,19; B11,26,28). More sensitive data could be shared (B18,19). However, general data sharing between companies would also increase (B11,26,28). With the integration of PETs, companies could merge encrypted data and use it for benchmarking

purposes, such as the cross-company comparison of KPIs (B6,26; B19,6). Another concrete example would be the calculation of CO2 emissions along the supply chain. PETs could be used to merge data in encrypted form and thus calculate aggregated CO2 emissions (B19,26). In the marketing sector, PETs could enable companies to access aggregated advertising data in compliance with data protection regulations now that the use of third-party cookies has been legally restricted (B4,26). PETs would also be very useful in the area of cybersecurity (B10,8). Instead of each company individually setting up expensive security operation centers for network monitoring, they could join forces (B8,21). However, as cyber security data is highly sensitive, this would only be possible with data encryption through the integration of PETs (B8,21). PETs could also be used in the area of threat intelligence to securely share sensitive data about typical cyberattack patterns (B8,21). Another promising use case would be fraud detection, particularly in the case of money laundering offenses (B5,30,38). Here, banks need to track a large number of account transactions, which would only be possible through the secure sharing of data between banks using PETs (B5,30,38). In addition to the use of PETs for cross-company data sharing, they could also be a solution for cross-country data sharing. If a company operating in different countries wants to share data across countries, there would often be data protection hurdles. These could be eliminated with the integration of PETs (B19,5,24).

The interviews show that PETs can change the benefits of data sharing to the extent that new use cases arise that would not have been possible before. However, these new use cases only arise because PETs would ensure more secure data sharing (B9,19; B11,26,28). So, actually, PETs only directly influence the risks and costs of sharing data (Agahari & Reuver, 2022). PETs would minimize the '**risks**' of sharing data by preventing the other party from misusing the data. This can also be referred to as hard privacy (B4,36,40). With soft privacy, on the other hand, trust would be placed in contracts and NDAs (B4,36,40). Although PETs minimize the risks of data sharing, they would also increase the potential costs of damage if a security breach occurs, as data shared with PETs is usually more sensitive (B12,59).

Even if PETs minimize the risk of data misuse, there would still be a certain residual risk. Some interviewees expressed concern that the technology could have security gaps (B5,50; B6,32; B8,31; B17,25). In addition, the complexity of the technology could lead to misuse by users (B9,24; B10,60). This is because PETs would only guarantee input privacy but not output privacy (B13,24; B19,36). Unintentional incorrect queries by users could lead to sensitive information becoming known despite the guarantee of technical security (B13,24).

Many technologies would be technically perfect, but the risks often arise at the weakest link in the chain, which primarily refers to human errors and vulnerabilities (B12,48). Due to the lack of output privacy, there would also be a risk of reverse engineering, in which attackers could reconstruct the original data set by repeatedly querying slightly modified data sets and thus extract sensitive information from an otherwise protected data set (B5,50; B10,56; B20,23). The risks associated with the lack of output privacy are also addressed in the literature (Agahari et al., 2022; Garrido et al., 2022; Choi & Butler, 2019). Data manipulation is named as a further risk. There would always be a risk when working with other companies that they may contribute manipulated data or data of poor quality (B3,40; B10,62; B15,53; B20,17,23). Another risk would be that PETs often act as a kind of 'black box', which means that companies would receive results based on algorithms and data that they cannot see and understand (B10,58). In many situations, however, it would be crucial to be able to reconstruct calculations to verify results (B10,62; B17,27). One interviewee also expressed a risk for the market, as better data availability with PETs could expose companies to greater price pressure (B6,32). In a similar way to the Supply Chain Due Diligence Act, improved data availability could be used to better control suppliers and increase competition (B5,8).

### 6.2.2 The Importance of Trust and Control in the Context of PETs

The concept of '**control**' in connection with PETs is not discussed in detail in the interviews, but there is a general view that PETs increase control. However, the concept of '**trust**' is a controversial topic. On the one hand, the interviews show that the trust required for data sharing between actors is significantly reduced by the integration of PETs. PETs would increasingly replace trust in other actors with trust in the technology itself (B13,32). Only mathematics would have to be trusted (B11,8,10). One interviewee, therefore, also describes PETs as 'zero trust solutions', as they would replace interpersonal trust with technology (B5,10).

At the same time, however, the reduction in interpersonal trust through the use of PETs is also linked to the condition that the technology is trustworthy. Only if the technology guarantees that data is secure, less trust in the other actors would be required (B2,38). In addition, it is not possible to see how PETs calculate their results. This would also make trust in the technology and its algorithms essential (B20,21). However, generating this technological trust is a challenge. Technological trust could only arise if PETs are understood by companies (B1,54; B16,26; B20,19). However, company departments often lack the time and technical

capacity to familiarize themselves with such complex technologies (B19,22). So, in the end, companies have to trust the providers of PETs (B19,22; B4,36). Yet, companies would generally reject this, as they do not simply trust the statements of technology vendors when making such decisions (B3,46,48). The Privacy Enhancing Technology would have to be approved by internal company experts, the legislator, or an independent third party to be trusted (B3,46,48). In addition, new technologies are generally often met with mistrust or rejection in companies (B19,22). The interviews thus show the importance of trust in technology and the simultaneous challenge of establishing it. This is in line with findings from the literature that despite the increase in the need for technological trust, there is still a need to trust the developers of the technology (Lumineau et al., 2020).

Trust in the technology could be established through use cases in which PETs are proven to work (B8,33,35). However, use cases only exist if the technology is trusted. This problem can be described as a chicken and egg problem (B8,33,35). Once this initial problem has been overcome, trust in the technology will increase the more established PETs become on the market (B5,42,46). It could be compared to self-driving cars. Only when people are used to doing it they would start to trust the technology (B12,54). In addition, trust in the technology would be strengthened if the technology was certified and the provider verified (B18,21). Once trust in the technology has been established, less trust would be needed when sharing data (B18,21).

### 6.2.3 Internal Factors Influencing PET-Based Data Sharing

Organizational readiness and organizational fit influence whether companies are willing to share data with PETs (Agahari et al., 2022; Hasani et al., 2023; Kanger & Pruulmann-Vengerfeldt, 2015). In the interviews, organizational readiness is addressed indirectly in the form of '**corporate culture**'. Companies would need to be open to technologies and have a certain technical understanding (B4,16; B13,26,28). In most cases, it would be the innovation departments in companies that invest in PETs rather than the operational areas (B10,66). Security in handling data and a willingness to invest money in innovative technologies would also positively affect PET-based data sharing (B16,20). The more digitalized a company is, the more willing it will be to invest in PETs (B19,34).

Another internal factor mentioned in the interviews is the '**size**' of the company. However, opinions on the influence of size on PET-based data sharing were divided in the interviews. On the one hand, larger companies would be more likely to invest in PETs (B19,34). They

would have a large network of suppliers and could, therefore, derive the greatest benefit from PETs if they use these technologies to share data with their suppliers (B11,38). On the other hand, larger companies would also often have more fixed processes and structures, which negatively affects the adoption of PETs (B19,24). In the departments of the company, it would be preferable to retain the familiar structures, as this is more convenient (B19,24). As long as data is secured by familiar means such as contracts and all regulations are complied with, there would be little incentive for departments to implement the company's interest in securing data technologically (B19,24). Furthermore, in the interviews, it is highlighted that it would be easier for smaller companies to invest in PETs, as only a few stakeholders need to be convinced, and this would speed up decision-making processes (B15,51).

The attitude of top management was identified in the literature as another important influencing factor for PET-based data sharing (Fassnacht et al., 2023; Hasani et al., 2023; Kanger & Pruulmann-Vengerfeldt, 2015). The importance of '**top management**' is also emphasized in an interview. One challenge when introducing PETs would be that they must be considered from a legal, organizational, and technical perspective (B10,72). This would mean that not only two companies are involved but also many different company departments. Top management would be crucial for coordinating this cross-departmental collaboration (B10,72).

Furthermore, the availability of IT specialists is highlighted in the literature as a decisive influence on the adoption of PETs (Hasani et al., 2023). The importance of specialists is also emphasized in the interviews due to the high complexity of PETs (B4,36; B12,63; B13,20; B16,22). Statements in this regard are assigned to the concept of '**resources**' in the interview analysis. A reason why many companies do not use PETs would be a lack of knowledge about this technology and a lack of capacity to deal with it (B2,34; B8,29; B19,22). To introduce PETs, the company must have the technical expertise to evaluate this technology (B7,50). Also, a high level of technical expertise would be required to implement the technology (B13,20). Even if PET vendors were to offer to do this, companies would be very reluctant to give external parties access to their sensitive IT systems, especially if they do not understand the technology (B10,60). This shows that company-internal technical expertise is crucial for the introduction of PETs.

### 6.2.4 External Factors Influencing PET-Based Data Sharing

The interviews emphasize '**external pressure**' in the form of laws and regulations as an important influencing factor. Laws and regulations would make PETs increasingly important (B15,49). For example, according to the GDPR, the highest level of privacy must always be sought that is within the scope of current technical possibilities. This would mean that PETs have to be used in the future following the GDPR (B10,48). At present, the only major problem is that it is not yet completely certain whether PETs can guarantee data processing in compliance with data privacy regulations (B19,24).

Furthermore, the '**sector**' was identified in the literature as an external factor influencing the introduction of PETs (Hasani et al., 2023). The IT and services sectors (healthcare, consulting, retail, and telecommunications) have higher adoption rates of PETs as they deal with large amounts of information (Hasani et al., 2023). The sector was also highlighted in the interviews as an important influencing factor for the adoption of PETs. The usage of PETs would be particularly beneficial for the financial, healthcare, and education sectors, as they handle a great deal of sensitive data (B5,48; B14,26). Particularly in the healthcare sector, innovative technologies for the privacy-preserving sharing of data and joint statistical analysis would be sought (B16,4; B19,34). The increased availability of data through the use of PETs could lead to more precise statistical analyses and significantly simplify the process of drug approval (B1,50). Healthcare insurance companies would also be interested in PETs, as they would greatly benefit from increased efficiency in the healthcare sector (B10,66).

Concerning the '**region**', the interviews emphasized that the European market would offer more potential for PETs than the American market, as the data protection regulations in Europe are more demanding (B9,22). Within Europe, the Netherlands, in particular, would be very open to these technologies (B18,25). In Germany, on the other hand, the focus would be more on data protection than on collaboration with the help of PETs (B18,25). This is also consistent with a study that shows that the Netherlands is the European country with the highest level of investment in PETs (World Economic Forum, 2024).

### 6.2.5 Challenges in the Acceptance of PETs on the Market

In the interviews, it was emphasized that the integration of PETs could lead to new data collaborations and make existing collaborations more efficient (B9,19). The integration of PETs would allow both more data and more sensitive data to be shared (B5,42; B11,26,28;

B18,19). However, for PETs to be able to change the factors influencing companies' data-sharing decisions and increase data sharing, they must first be accepted and used on the market. The interviews highlighted several challenges in the acceptance of PETs on the market. The statements in this regard were assigned to the new concept of **'Acceptance of PETs'** in the interview analysis.

The first major challenge in the acceptance of PETs is the lack of understanding of the technology. A clear understanding of the technology is a prerequisite for the adoption of PETs (Kanger & Pruulmann-Vengerfeldt, 2015). PETs would often be seen as something almost magical that people do not understand. This lack of understanding would inhibit their broad acceptance (B16,26). Only when companies understand how PETs work and what benefits they could bring would they be prepared to invest in the technologies (B1,54; B20,19). To overcome this challenge, providers of PETs would need to communicate with potential users in easy-to-understand language (B4,16).

Closely linked to the lack of understanding of PETs is the lack of awareness and positive examples of the technology. So far, there would be still too little evidence that PETs work (B7,44). This would make it difficult to assess the risk, which deters many companies, especially in the highly sensitive IT sector (B7,44). For PETs to be accepted, there must first be a critical mass of early adopters who use PETs (B17,27). Clear regulations from legislators would also help (B19,30).

Another challenge is finding the right use case for the technology. Investments in PETs would only be made if companies could derive a clear economic benefit from them (B15,49). As a provider, the focus should, therefore, be on the use case that can be solved with PETs and not on the technology (B7,36). This is because if the technology were used to advertise, only a small group of companies that could understand the topic at all would be reached. With an understandable use case, on the other hand, many more companies could be reached (B7,36). This is why many suppliers of PETs would initially have advertised their technology but then switched to emphasizing the benefits of PETs (B10,54).

PETs would also be a collaborative technology that requires several players (B10,70). This would mean that the solution has to be sold to at least two companies at the same time, which is often made more difficult by different priorities and schedules (B10,70). In addition, not only technical but also organizational and legal aspects must be considered, which means that

several departments would be involved (B10,70). Therefore, many different stakeholders would have to be convinced and agree before PETs could be used (B13,18,30; B19,16,28).

# 7 Discussion

## 7.1 Key Findings

The findings from the systematic literature review and the interviews on factors that influence data sharing between companies and the effects of PETs on this are summarized below.

### 7.1.1 Factors Influencing Corporate Data Sharing

The following part of the thesis explains the findings regarding which factors companies weigh up when sharing data with other companies and thus answers the first research question.

Both in the interviews and the literature, interoperability and associated characteristics such as data quality and technical interfaces are described as a prerequisite and important influencing factor in the decision to share data (Bastiaansen et al., 2020; Fassnacht et al., 2023; B1,24; B5,24,26; B8,17 B12,22; B19,14). The digital infrastructure is also mentioned as a prerequisite for data sharing (B5,26; B19,14,16). These results underline the role of what may be termed technical framework as an essential prerequisite for corporate data sharing.

Additionally, risks and benefits can be confirmed as the main factors influencing the cross-company sharing of data. Both in the literature and the interviews, the importance of benefits in the decision of companies to share data is highlighted (Dahlberg & Nokkala, 2019; Agahari et al., 2022; B1,14; B4,24; B13,12; B15,35; B20,9). At the same time, however, it is also emphasized that companies can only poorly assess the benefits of data sharing (Dahlberg & Nokkala, 2019; Gelhaar & Otto, 2020; B17,10,12). The risks, on the other hand, are perceived even more strongly (B17,10,12). The ‘competitiveness risk‘, the ‘data misuse risk‘, the ‘end-users privacy risk‘, and the ‘reputation risk‘ can be identified as the main risks of data sharing (Agahari et al., 2022). Especially the ‘end-users privacy risk’, paraphrased in the interviews as ‘legal risk’ is highlighted (B1,20; B6,14; B7,18). Due to the uncertain legal situation, the risk would appear to be even higher (B5,22).

Whether trust influences the decision to share data is the subject of controversial debate: In the literature, trust is seen as an important influencing factor for the willingness to share data (Dyer & Chu, 2003; Li & Lin, 2006; Gelhaar & Otto, 2020). In the interviews, opinions were divided on the importance of trust. On the one hand, trust is important because companies often have no choice but to trust that data will not be misused (B8,7; B13,12). On the other

hand, the decision-making process in companies is based on a careful risk assessment and less on trust (B3,28; B17,8,3; B20,6). Closely linked to trust is the factor of control. Control mechanisms can create trust between companies (Bons et al., 2012; Agahari et al., 2022; B8,7,9; B17,8). This is why control positively affects the willingness to share data (Bons et al., 2012; Agahari et al., 2022). In the interviews, the form of control referred to as "contract-based control" is even described as a prerequisite for data sharing (B2,8; B6,8; B7,8; B8,17; B13,16).

Internal company factors also play an important role when deciding about data sharing. Intra-organizational data governance (Lis & Otto, 2020), the availability of financial and human resources (Müller et al., 2020; Dahlberg & Nokkala, 2019), the importance of top management (Fassnacht et al., 2020; Li & Lin, 2006; Enders & Benz et al., 2020) and the type of data (Enders et al., 2020; Müller et al., 2020; Dahlberg & Nokkala, 2019) influence companies in their willingness to share data. In the interviews, the importance of top management (B1,26; B7,20) and type of data (B1,58; B3,24) is confirmed. Furthermore, the corporate culture is highlighted as a decisive factor in the willingness to share data (B9,14; B12,10).

External pressure in particular from regulatory requirements (Fassnacht et al., 2023; Holler et al., 2019; B4,6; B5,4,10,14; B6,28; B9,11; B10,28; B12,40,42; B13,12; B15,37), the sector in which a company operates (Hoffmann et al., 2020; Witte et al., 2020; B5,28; B7,22,24; B10,10,16), situational circumstances (Dahlberg & Nokkala, 2019; B3,50; B6,20; B16,4,13) and regional differences (Müller et al., 2020; B18,25; B19,6) were emphasized as external influences for data sharing both in the literature and in the interviews.

### 7.1.2 The Impact of PETs on Data Sharing and its Influencing Factors

This section addresses the second research question about the influence of PETs on corporate data sharing.

The integration of PETs changes both the benefits and the risks of data sharing. Concerning the change in the risk factor, it can be stated on the one hand that risks such as the risk of competition, the risk of violating the privacy of end users, or the risk of the copying problem are reduced (Agahari et al., 2022; Garrido et al., 2022; B4,36,40; B12,59). On the other hand, new risks may also arise or shift to other actors (Agahari et al., 2022). For example, the lack of output privacy of PETs can pose a risk for both data owners and data users (Agahari et al.,

2022; Garrido et al., 2022; B5,50; B10,56; B13,24; B19,34; B20,23). In addition, the risk of manipulation of the input data by one of the actors is emphasized in the interviews (B3,40; B10,62; B15,53; B20,17,23). Regarding the benefit factor, its importance for the adoption of PETs is particularly emphasized. For companies, clear visibility of benefits is crucial when sharing data with PETs (Agahari et al., 2022; Hasani et al., 2023; B1,36; B2,30; B3,38). Moreover, the benefits resulting from data sharing change with the integration of PETs. New types of collaboration are emerging, creating additional benefits from data sharing (Körner et al., 2022; Agahari et al., 2021; B4,26; B5,30,38; B6,26; B8,21; B9,19; B10,30; B11,26,28; B16,15; B19,5,6,24,26). Consequently, PETs also impact the extent of data sharing: more data is shared overall, including more sensitive data (Agahari & Reuver, 2022; Körner et al., 2022; B11,26,28; B18,19).

Both the control factor and the trust factor are influenced by PET-based data sharing. PETs can strengthen the type of "technology-based control" (Agahari et al., 2022). As a result, the importance of interpersonal trust is decreasing, while the importance of technological trust is increasing (Agahari et al., 2022; B5,10; B11,8,10; B13,32). However, interpersonal trust in the providers of such technologies is still crucial, especially at present, when PETs are still very new and unknown (B4,36; B19,22; Lumineau et al., 2020).

When sharing data with PETs, the internal factors of top management and the availability of financial and human resources become particularly important (Müller et al., 2020; Hasani et al., 2023). In addition, "organizational readiness" is emphasized as an important influence on the willingness to share data with PETs (Hasani et al., 2023; Kanger & Pruulmann-Vengerfeldt, 2015). The type of data shared also changes with the involvement of PETs (Agahari et al., 2022; Kanger & Pruulmann-Vengerfeldt, 2015). In line with the literature, the interviews highlight the importance of top management (B10,72) and, in particular, human resources (B4,36; B7,50; B12,63; B13,20; B16,22). Furthermore, the corporate culture (B4,16; B19,34) and the company size (B11,38; B15,51; B19,24,33) are emphasized as influencing factors on the willingness to share data with PETs.

Concerning external factors, data sharing with PETs is primarily influenced by external pressure (Hasani et al., 2023; Scheibner et al., 2021; Kanger & Pruulmann-Vengerfeldt, 2015). Particularly, laws and regulations put pressure on companies to share data with PETs (Scheibner et al., 2021; Helminger & Rechenberger, 2022; B10,48; B15,49; B19,24). Therefore, PETs can also change the sectors in which data is most frequently shared, as they

could be particularly useful in highly regulated areas such as healthcare (Witte et al., 2020; Hasani et al., 2023; B16,4).

## 7.2 Theoretical Implications

### 7.2.1 Corporate Data Sharing Model

Based on the extended APCO model by Dinev et al. (2015), a modified corporate data-sharing model is proposed (see Figure 5).

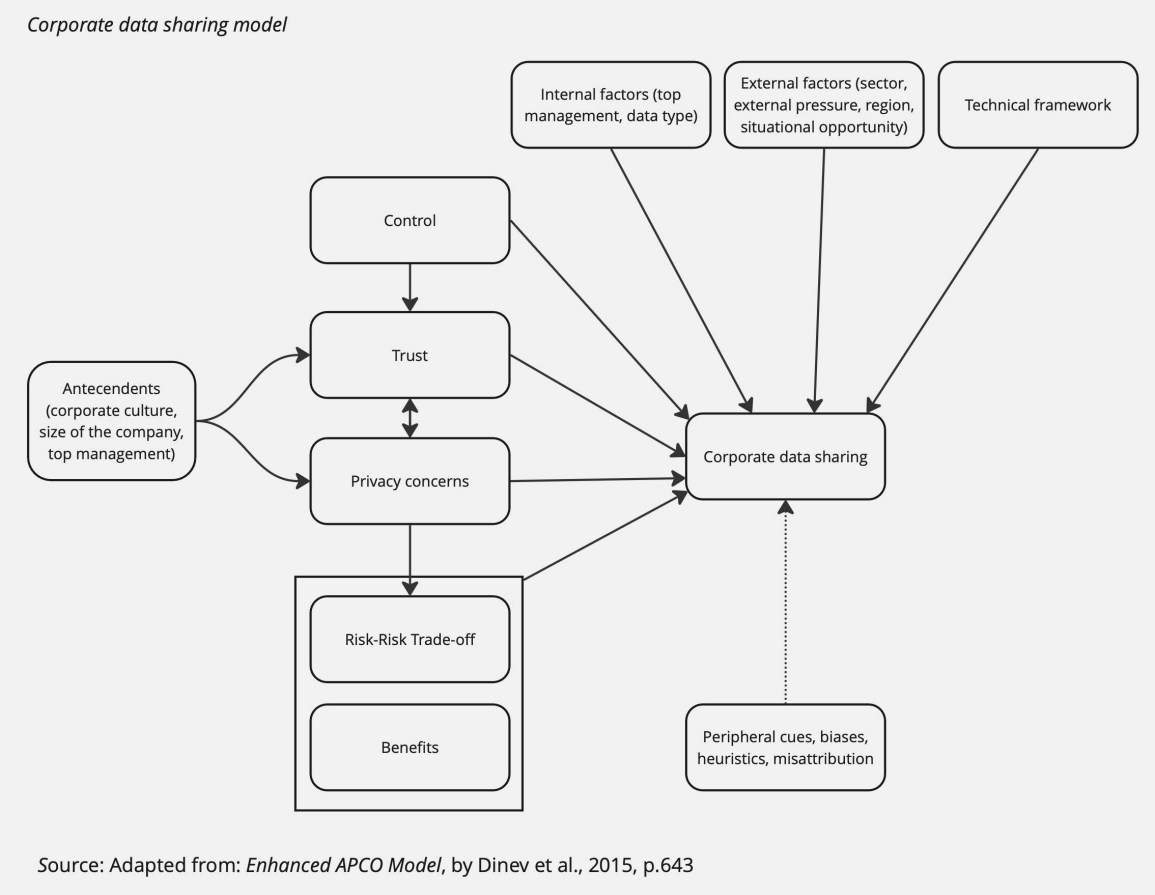


*Figure 5: Corporate data sharing model*

In the following, the factors presented in the corporate data-sharing model and their relationship to each other are explained.

In the extended APCO model, the trade-off between risks/costs and benefits is presented as a central aspect of individuals' decisions regarding the sharing of data (Smith et al., 2011; Dinev et al., 2015). Risks, costs, and benefits also influence companies in their willingness to share data (Dahlberg & Nokkala, 2019). However, the three factors are more closely interrelated than in the individual context. According to one interviewee, economic benefit could be

achieved through a reduction in risk as well as a reduction in costs or an increase in sales (B15,35). In most cases, the main benefits of data sharing are the reduction of other risks. For example, the benefits of better supply chain monitoring or predictive analysis resulting from data sharing reduce the risk of failures (Dahlberg & Nokkala, 2019; Müller et al., 2020; Bechtsis et al., 2022). Therefore, the greatest risk of sharing data is not to share data (B9,9; B20,13). However, there are also risks, such as the "competitiveness risk" or "data misuse risk" that arise from sharing data (Agahari et al., 2022). This results in a risk-risk trade-off for companies when making data-sharing decisions, as the risks are changed with the sharing of data (Hansen et al., 2008). Some risks arise from sharing data, and other risks can be reduced. In addition to this risk-risk trade-off, actual benefits such as process optimization, greater customer reach, or increased potential for innovation also influence the willingness to share data (Müller et al., 2018; Müller et al., 2020; Dahlberg & Nokkala, 2019). However, in contrast to data disclosure decisions of individuals, for companies, the cost factor is less important when deciding about data sharing (Smith et al., 2011; Dahlberg & Nokkala, 2019). In the interviews, the cost factor is not mentioned at all. This suggests that companies are making a fundamental trade-off between risks that arise, risks that are reduced, and additional benefits when sharing data. Therefore, the trade-off between risks, costs, and benefits regarding the data-sharing decisions of individuals in the extended APCO model is replaced by a risk-risk benefit trade-off in the corporate data-sharing model (see Figure 5).

At an individual level, privacy concerns are presented as a further factor influencing data disclosure decisions (Dinev et al., 2015; Smith et al., 2011). Privacy concerns also exist in companies in the form of concerns for the protection of sensitive company data, the preservation of business and competitive secrets, and compliance with regulatory requirements (Agrawal et al., 2021). In the extended APCO model for individuals, privacy concerns influence how strongly a person perceives and evaluates the risks of sharing data, whereby higher concerns increase the perceived risks (Dinev et al., 2015). Although privacy concerns are not directly mentioned as an influencing factor for data sharing between companies, there is much to suggest that this is also the case in the corporate context. A large proportion of interviewees described the decision-making process of companies to share data as subjective or at least partially subjective (B4,12; B7,10; B10,18; B12,10; B15,19; B18,7). Risks are usually overestimated and avoided (B5,12; B7,14; B10,28). The lack of rationality in decision-making processes and the overestimation of risks indicate that privacy concerns also influence risk perception and, thus, the decision to share data in companies. Therefore,

the privacy concerns factor is added to the corporate data sharing model (see Figure 5). Just like the data privacy concerns of individuals, the data privacy concerns of companies are also influenced by external factors (Belanger & Crossler, 2011). Three factors are identified that influence the privacy concerns of companies indirectly via the impact on rationality in the decision-making process. The larger a company is, the more rational the decision-making processes are (B8,13). Also, non-technology-oriented companies are more emotionally driven and less rational when making data-sharing decisions (B9,14). Moreover, the top management and the decision-making structures in companies in general would have a major influence on the rationality of the decision-making process (B1,14; B5,26). The impact of a company's size, corporate culture, and top management on rationality in decision-making suggests that these factors may influence the extent of corporate privacy concerns. They are, therefore, added to the corporate data-sharing model as factors influencing privacy concerns (see Figure 5).

For individuals, trust is another important factor that influences the willingness to share data (Dinev et al., 2015; Smith et al., 2011). In the literature, trust is also shown to be an important influence for companies when making decisions regarding the sharing of data (Dyer & Chu, 2003; Gelhaar & Otto, 2020). In the interviews, opinions are divided on whether trust is important for companies when sharing data. It is striking that four of the five interviewees who consider trust to be less important describe the decision-making process as rationally controlled (B1; B3; B17; B20). By contrast, of the seven interviewees who consider trust to be a decisive factor, only one (B13) rated the decision-making process as predominantly objective and rational. This indicates a correlation between the degree of rationality in the corporate decision-making process and the importance of trust. As already stated, the size-, culture-, and top management of the company could have an impact on the rationality of the decision-making process. This indicates that the importance of trust in data-sharing decisions could also depend on the size-, culture-, and top management of the company. The size of the company, the corporate culture, and the top management are, therefore, added to the corporate data-sharing model as factors influencing trust (see Figure 5).

The control factor is not included in the extended APCO model (Dinev et al., 2025). For companies, however, control plays an important role in the decision to share data (Bons et al., 2012; Agahari et al., 2022). In particular, the “contract-based control” identified by Agahari et al. (2022) is emphasized as a prerequisite and thus essential for data sharing (B2,8; B6,8;

B3,16; B7,8; B8,17). Therefore, the corporate data-sharing model is expanded to include the factor of control (see Figure 5).

Of the identified company-internal influencing factors of data governance, resources, top management, and data type, only the latter two were confirmed in the interviews as significant for data sharing (see Chapter 7.1.1). Therefore, only top management and data type are added to the corporate data-sharing model as internal factors (see Figure 5).

As already described in the previous chapter (7.1.1), the external factors sector, external pressure, region, and situational opportunity are identified as influencing factors for data-sharing decisions both in the literature and the interviews. In addition, factors summarized under the term the technical framework are seen as a prerequisite for cross-company data sharing (see Chapter 7.1.1). Both external factors and the technical framework must, therefore, be added to the corporate data-sharing model (see Figure 5).

The importance of the factors that individuals weigh up concerning the disclosure of data depends on the ‘level of effort’ (Dinev et al., 2015). This influence cannot be determined for companies. Even if some interviewees see the decision-making process as subjective, decisions still undergo compliance checks and are usually reviewed by many different departments (B3,18; B13,16; B18,7). In particular, as employees do not want to bear responsibility for wrong decisions, decisions are always approved by several departments (B3,20). This shows that, in contrast to individuals, no hasty decisions are made in companies based on feelings or motivation. The influence of the ‘level of effort’ can, therefore, be neglected in the corporate data-sharing model.

Finally, the extended APCO model also presents peripheral cues, biases, heuristics, and misattribution as influences on individuals' decisions (Dinev et al., 2015). These influences are not mentioned directly about companies. However, there are indications that these factors are also important in the corporate context. For example, companies often assess the trustworthiness of other actors by categorizing them according to specific dimensions and making decisions about trustworthiness based on these categories (Lumineau et al., 2020). Peripheral cues could influence these categorizations. Furthermore, categorization could be seen as a heuristic, as it represents a simplified decision-making process. However, as there is no direct confirmation of these impacts, the influence is only shown as a dashed line in the corporate data-sharing model (see Figure 5).

### 7.2.2 Corporate PET-Based Data Sharing Model

If companies have to decide about data sharing with PETs, changes can be observed in the factors that influence this decision. Therefore, the corporate data sharing model described in section 7.2.1 must be adapted for PET-based data sharing (see Figure 6).

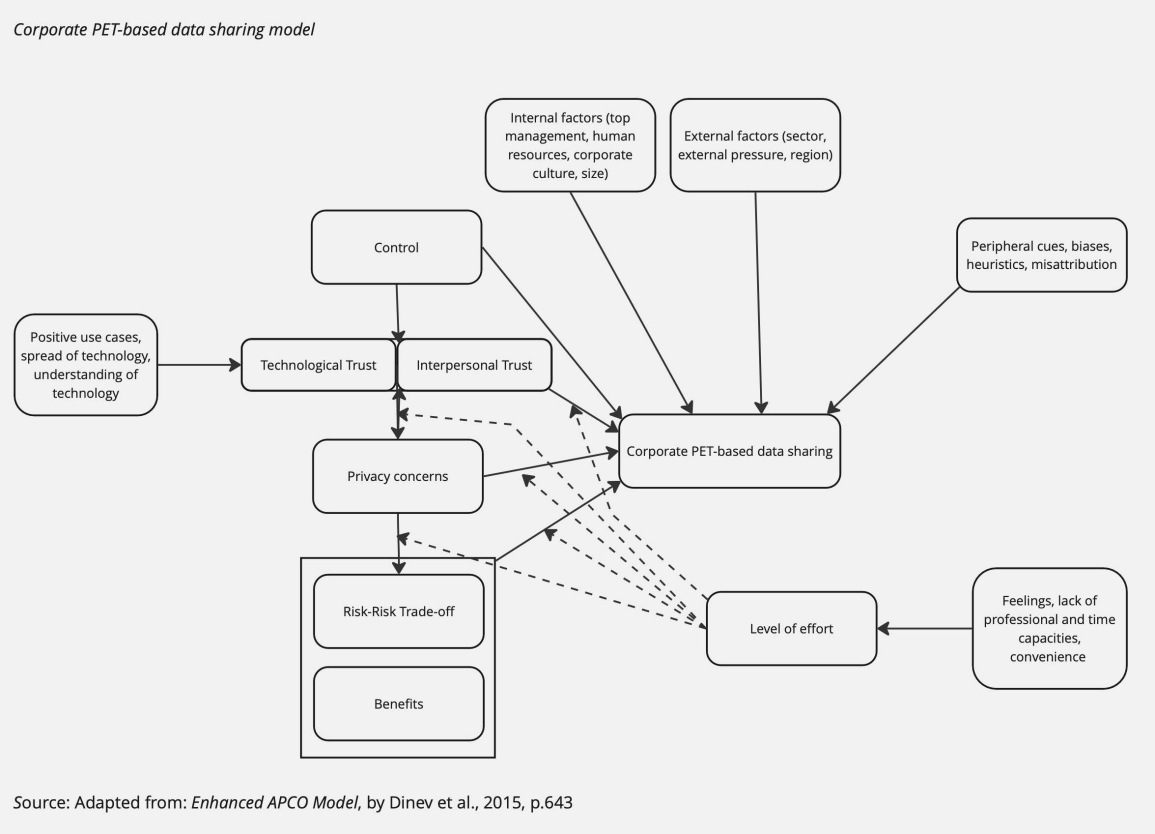


*Figure 6: Corporate PET-based data sharing model*

In the following, the changes presented in the corporate PET-based data-sharing model in comparison to the corporate data-sharing model are discussed.

The interviews show that the benefits that could be achieved with PETs are often unclear for companies (B1,36). Many companies would not recognize any specific use cases in which these technologies could offer direct added value (B2,30). One interviewee, therefore, describes PETs as “a solution, looking for a problem” (B20,17). The lack of knowledge among companies about the benefits and use cases of these technologies shows that there is no direct demand for these technologies but that providers of PETs must proactively approach companies and convince them of the benefits and suitable applications. However, providers face several challenges in doing so. Companies often lack the technical and time capacities to deal with PETs and the possible solutions proposed by providers in more detail (B2,34;

B8,29; B19,22). In addition, employees have no real incentive to introduce PETs. It is more convenient to maintain familiar structures (B19,24). Also, new technologies are generally met with mistrust and rejection (B19,22). Thus, the decision to adopt PETs is influenced by mistrust, convenience, and a lack of professional and time capacities. These influences on the decision to share data with PETs show similarities to the influences on data-sharing decisions of individuals. Just as feelings, cognitive and time constraints, and motivation, summarized in the ‘level of effort’, influence individuals' data-sharing decisions (Dinev et al., 2015), the decision to adopt PETs is also influenced by feelings such as mistrust, a lack of professional and time capacities, and convenience. In the previous chapter, the assumption was made that the ‘level of effort’ has no significance for companies' decisions regarding data sharing. However, for PET-based data-sharing decisions, the ‘level of effort’ could become more important again. For this reason, the ‘level of effort’ is included in the corporate PET-based data-sharing model (see Figure 6). As in the extended APCO model, the 'level of effort' has no direct impact on the decision to share data, but influences the significance of the other factors (Dinev et al., 2015). This is also shown in the corporate PET-based data-sharing model with dashed lines (see Figure 6). In addition, the factors influencing the ‘level of effort’ – ‘feelings’, ‘lack of professional and time capacities’, and ‘convenience’ - are added to the model (see Figure 6).

With a low level of effort, external influences such as peripheral clues, prejudices, heuristics, and misattributions are more important factors in decision-making than rational consideration of benefits and risks (Dinev et al., 2015). In the case of PET-based data sharing, use cases were also observed in which a decision was made against the use of these technologies despite clear advantages and without rational counterarguments (B7,6). This indicates that a low level of effort also leads to a higher influence of peripheral clues, prejudices, heuristics, and misattributions when sharing data with PETs. These external influences are, therefore, also taken into account in the corporate PET-based data-sharing model (see Figure 6).

Another important factor in connection with PETs is trust. Both the literature and the interviews show that interpersonal trust is becoming less important with PETs and that trust in technology is becoming more important instead (Agahari et al., 2022; B5,10; B13,32). Technological trust is often presented as given by the interviewees who work in the field of PETs (B9,20; B11,8,10; B13,32; B20,21). At the same time, however, four interviewees from the IT sector with no connection to PETs expressed their concern that PETs could contain security gaps (B5,50; B6,32; B8,31; B17,25). This suggests that trust in the technology is not

a given for companies, even though PET providers advertise that everything is mathematically provable (B11,8,10). To generate technological trust, positive use cases, a certain spread of the technology, and a good understanding of the technology could help (B5,42,46; B8,33,35; B19,22). As already shown in the results of the interviews, none of this is the case for PETs. This means that, due to the lack of technological trust, a great deal of interpersonal trust is required. Less in the party with whom data is shared but more in the party offering the technology. This confirms the findings of Lumineau et al. (2020) that despite the increase in the importance of trust in technology, the need to trust the developers of the technology remains. One interviewee stated that the greater the technological trust in PETs, the less important interpersonal trust would become in the future (B18,21). The explanations thus indicate that there must always be a certain degree of trust, which is composed of technological and interpersonal trust. The larger one of the two types of trust is, the less of the other type is needed to maintain the required level of trust. In the corporate PET-based data-sharing model, the trust factor is, therefore, represented as a composition of technological and interpersonal trust (see Figure 6). The previously identified factors, “positive use cases, a certain spread of the technology, and a good understanding of the technology”, are presented in the corporate PET-based data-sharing model as an influence on the proportion of technological trust (see Figure 6).

Concerning the control factor, a distinction can be made between technology-based control, structural-based control, and contract-based control (Agahari et al., 2022). The literature shows that technology-based control, in particular, is increased by PETs (Agahari et al., 2022). Nevertheless, the interviews emphasize that the other types of control also remain important for PET-based data sharing (B20,17). This is because the introduction of PETs can be facilitated by existing agreements and contracts between different actors (B20,17). This means that all control types remain important even with the integration of PETs. The control factor from the corporate data-sharing model, therefore, remains unchanged in the corporate PET-based data-sharing model.

The influence of top management and human resources on PET-based data sharing is emphasized both in the literature and the interviews (see Chapter 7.1.2). In addition, the size of the company and the corporate culture were mentioned in the interviews as further internal influences (see Chapter 7.1.2). These four internal factors are added to the corporate PET-based data-sharing model (see Figure 6).

As external factors, the interviews confirm the factors of external pressure, sector, and region identified in the literature research (see Chapter 7.1.2). The situational opportunity factor cannot be confirmed in the interviews and is therefore omitted for the corporate PET-based data-sharing model (see Figure 6). The 'technical framework' factor is also omitted for the corporate PET-based data-sharing model, as this factor was not addressed in the interviews.

### 7.2.3 Level of Privacy with PETs

The term Privacy Enhancing Technology already emphasizes the increase in privacy through these technologies. It is argued that PETs would protect the privacy of data, and therefore, with the integration of PETs, personal data could be shared without violating the GDPR (Helminger & Rechenberger, 2022; Scheibner et al., 2021). Some approaches to defining privacy show that there are different degrees of privacy (Smith et al., 2011). According to the definition of information privacy, it is achieved when one retains control over one's data (Belanger & Crossler, 2011). However, PETs are associated with a very limited and technically focused understanding of privacy (Agrawal et al., 2021). Understandings of privacy, such as complete control over one's data, are neglected by PETs (Agrawal et al., 2021). Privacy is only understood as securing the confidentiality of data using cryptographic methods (Agrawal et al., 2021). However, PETs still make it possible to perform data analyses and generate ‘panoptic knowledge’, which can also be exploited (Agrawal et al., 2021). For example, results from data analyses could be used to exert increased price pressure on suppliers (Müller et al., 2018; B6,32). In particular, the new Supply Chain Due Diligence Act would already be used by companies to better control suppliers and spur competition (B5,8). PETs would open up many more opportunities to put companies under pressure (B6,32). This example shows how the panoptic knowledge gained through PETs could be exploited. Suppliers would secure the confidentiality of their data with PETs, but their privacy would still be violated, as other companies could use the knowledge gained from the data against them. This confirms the limited privacy understanding of PETs emphasized by Agrawal et al. (2021).

## 7.3 Practical Implications

The practical implications for providers of PETs arise from the existing challenges in the acceptance of these technologies. A key problem is the lack of positive use cases and the lack of awareness of the technologies (B7,44; B17,27). Only when companies see data-sharing use

cases that work for others will they be willing to share data themselves (Dahlberg & Nokkala, 2019). A critical mass of early adopters would have to be reached for PETs to be accepted (B17,27). In this context, it seems sensible for providers of PETs not to see other providers exclusively as competitors but rather to recognize the advantages of cooperation. Particularly in the current market situation, in which the number of PET providers is still limited, and at the same time, many potential companies could benefit from the use of PETs, cooperation between PET providers could be advantageous in certain respects. Cooperations could help to increase awareness of PETs and develop a wider range of positive use cases. PET providers need to be aware that the success of another provider in the current market situation does not necessarily represent a disadvantage for their own business but could rather contribute to the generation of a critical mass of early adopters, which ultimately benefits everyone.

Furthermore, the importance of trust in the technology and its providers for the acceptance of PETs has already been highlighted several times (Agahari et al., 2022; Lumineau et al., 2022). The challenge here is that companies usually do not trust PET vendors (B3,46,48). Therefore, the involvement of consulting firms could be helpful for providers of PETs. Consulting firms often enjoy a high level of trust among their customers and could, therefore, play a key role in building the initial trust in PETs that providers themselves may not be able to achieve on their own (B7,50,52). At the same time, consulting companies would benefit from this as they could expand their consulting expertise with the latest data security technologies.

Furthermore, the importance of benefits in data-sharing decisions shows that PET providers should focus on practical use cases rather than on the technology itself (B15,35). In this regard, one interviewee described how PET providers usually initially try to promote their technology but later switch to selling solutions for use cases. In doing so, they also adapt their terminology and increasingly fall back on familiar concepts to make the solutions more understandable and accessible (B10,54). Providers should, therefore, focus their marketing on the economic benefits of their solution and not get lost in complex technical details.

Another challenge for the adoption of PETs is that, as collaboration technologies, they require several players (B10,70). Providers have to sell their solutions to at least two companies at the same time, which is often complicated by different priorities and schedules (B10,70). In addition, technical, organizational, and legal aspects must be considered, which means that many different departments and interests must be involved (B13,18,30; B19,16,28). To avoid this challenge, providers should focus on existing alliances for the introduction of PETs.

Agreements have already been made here, and there is a certain level of trust between the players, which significantly reduces coordination effort. Examples of such alliances could be established data rooms or industry associations (B20,17). Focusing on large companies could also circumvent the problem. One interviewee emphasized the usefulness of PETs for large companies that are confronted with the difficulties of cross-border data sharing (B19,6,26). Such corporations often have locations in different countries with different data protection regulations, which makes internal data sharing more difficult (B19,6,26). PETs could circumvent this challenge, and at the same time, providers only have to convince one company of their technology.

In general, the results of this study show that providers of PETs should focus, in particular, on technologically oriented medium-sized companies. This is because smaller companies are often less able to benefit from PETs (B11,38), while larger companies are often less flexible to change due to long decision-making processes and diverse interests (B15,51; B19,24). Companies in the healthcare, finance, education, and IT sectors are particularly predestined for the introduction of PETs (Witte et al., 2020; Hasani et al., 2023; B1,50; B5,48; B10,66; B14,26; B16,4).

There are also practical implications for companies. It can be seen that the adoption of PETs positively affects firm performance (Hasani et al., 2023). Companies that use PETs record faster growth, which is achieved through increased sales, optimized profits, improved productivity, and increased customer satisfaction (Hasani et al., 2023). PET-based data sharing should, therefore, definitely be considered. Companies must also be aware that a major risk of PETs is their misuse by employees, as incorrect queries can reveal sensitive data due to the lack of output privacy (Garrido et al., 2022). Companies, thus, should train their employees in the correct use of the technology. Furthermore, companies should use, on the one hand, a combination of PETs to reduce risks (Garrido et al., 2022). On the other hand, despite the use of PETs, the importance of trust and trust-generating measures such as contracts should not be neglected. This is because a combination of trust and control mechanisms complement each other perfectly to reduce the risk of opportunistic behavior (Poppo & Zenger, 2002). In contrast, the use of PETs alone increases the risk of opportunistic behavior (Handley & Benton, 2012).

“Legal certainty reduces legal risk” (B5,22) - this principle illustrates the need for clear rules when handling data. Uncertainty about regulations is one of the main reasons why companies

withhold data or why processes are lengthy and inefficient (Bitkom, 2024; B5,10; B6,28; B10,28). To fully exploit the potential of data and not fall behind in competition with other nations, it is therefore crucial that legislators issue clear, well-formulated, and easily understandable regulations.

## 7.4 Limitations and Outlook

In this thesis about exploring corporate data sharing and the impact of PETs on it, several limitations need consideration.

A central problem of the systematic literature review in the field of data sharing lies in the diversity of the use cases examined. The analyzed studies often differ greatly in terms of sector, type of data shared, and company size, which can lead to results that are difficult to compare and sometimes inconsistent. This complicates the drawing of general conclusions. In addition, in technology-driven areas such as PET-based data sharing, even short time intervals between publications can lead to significant differences in results as the underlying technologies and approaches evolve rapidly. To reduce these limitations, future studies could segment the literature more according to specific application areas, sectors, or company sizes. This would allow for more targeted and comparable findings. In addition, the use of dynamically updated overviews, so-called ‘living reviews’, could help to continuously integrate new studies and ensure that the results are up to date in this rapidly developing field of research.

To mitigate the risk of interviewer and participant bias, neutral and open-ended questions were used, and a standardized interview protocol was followed to ensure consistency. Anonymity and confidentiality were also ensured to help participants feel comfortable providing honest answers. However, despite these measures, there is always the possibility that interviewers unintentionally influence participants' responses or that interviewees present their opinions in a way that aligns with how they wish to be perceived, thus introducing bias. In addition, some interviewees - particularly those working in companies that sell PETs - indicated that they were unable to answer certain questions or did not want to disclose confidential information for competitive reasons. This reluctance can also influence the results, as key insights from the company’s perspective on PETs may only be partially captured. Another limitation was the limited time available for the interviews. Although an average duration of around 30 minutes is normally sufficient for an interview, it may be perceived as too short for a topic as complex as Privacy Enhancing Technologies. Particularly

for the 10 interviewees who work exclusively in the IT sector and have no direct experience with PETs, the brevity of the interviews may have influenced their responses about PETs. To address these limitations, future research could incorporate quantitative methods to validate the findings, offering a complementary approach that may help to generalize results and reduce bias by capturing a broader spectrum of perspectives.

Despite these limitations, this work can provide valuable insights into the impact of Privacy Enhancing Technologies on companies' data-sharing behavior and thus be of great benefit to both providers and users of Privacy Enhancing Technologies. Nevertheless, there is a great need for further research, especially as this work has identified a lack of trust in the technology as one of the main barriers to the spread of PETs. More awareness and a better understanding, which could be created through further research, would assist in building trust. Therefore, research plays a crucial role in the establishment of PETs.

Due to the unfamiliarity of PETs to date, it was important to examine them in a general context to provide companies in particular with general knowledge about these technologies and to inspire them. As PETs become more widespread in the future, the need for more specific research will also increase. Companies will need insights from specific application areas as well as studies of specific PETs. Comparisons of different PETs for specific applications are also an interesting area of research. Future research should focus on these areas. In addition, the theoretical implications of this work have already addressed the limited understanding of data privacy conveyed by PETs. In particular, the danger that these technologies could be misused to exert excessive price pressure on companies was pointed out. Ethical considerations about PETs offer an exciting area for further research. Examining the moral implications and responsibilities of implementing these technologies could help to create a trustworthy framework for their use.

## References

Abbas, A. E., Agahari, W., Van De Ven, M., Zuiderwijk, A. & De Reuver, M. (2021). Business Data Sharing through Data Marketplaces: A Systematic Literature Review. *Journal Of Theoretical And Applied Electronic Commerce Research*, *16*(7), 3321–3339. https://doi.org/10.3390/jtaer16070180

Abraham, R., Schneider, J. & Brocke, J. V. (2019). Data governance: A conceptual framework, structured review, and research agenda. *International Journal Of Information Management*, *49*, 424–438. https://doi.org/10.1016/j.ijinfomgt.2019.07.008

Agahari, W. & De Reuver, M. (2022, June). *Rethinking consumers' data sharing decisions with the emergence of multi- party computation: an experimental design for evaluation* [Conference Paper].Thirtieth European Conference On Information Systems (ECIS 2022), Timișoara, Romania. https://www.researchgate.net/publication/360555237_Rethinking_consumers'_data_sharing_decisions_with_the_emergence_of_multi-party_computation_an_experimental_design_for_evaluation

Agahari, W., Dolci, R. & De Reuver, G. (2021). Business model implications of privacy-preserving technologies in data marketplaces: The case of multi-party computation. *European Conference On Information Systems*. https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1058&context=ecis2021_rp

Agahari, W., Ofe, H. & De Reuver, M. (2022). It is not (only) about privacy: How multi-party computation redefines control, trust, and risk in data sharing. *Electronic Markets*, *32*(3), 1577–1602. https://doi.org/10.1007/s12525-022-00572-w

Agrawal, N., Binns, R., Kleek, M. V., Laine, K. & Shadbolt, N. (2021, January 20). *Exploring Design and Governance Challenges in the Development of Privacy-Preserving Computation* [Conference Paper]. Proceedings Of The 2021 CHI Conference On Human Factors in Computing Systems, Yokohama, Japan. http://arxiv.org/abs/2101.08048

Alashoor, T., Keil, M., Smith, H. J. & McConnell, A. R. (2022). Too Tired and in Too Good of a Mood to Worry About Privacy: Explaining the Privacy Paradox Through the Lens

of Effort Level in Information Processing. *Information Systems Research*, *34*(4), 1415–1436. https://doi.org/10.1287/isre.2022.1182

Aleem, U., Cavusoglu, H. & Benbasat, I. (2017). An Empirical Investigation of the Antecedents and Consequences of Privacy Uncertainty in the Context of Mobile Apps. *Information Systems Research*, *31*(4), 1037–1063. https://doi.org/10.1287/isre.2020.0931

Arora, A., Athreye, S. & Huang, C. (2016). The paradox of openness revisited: Collaborative innovation and patenting by UK innovators. *Research Policy*, *45*(7), 1352–1361. https://doi.org/10.1016/j.r espol.2016.03.019

Bamford, S. (2020). Applications of privacy-enhancing technology to data sharing at a global pharmaceutical company. *Journal Of Data Protection & Privacy*, *Vol.3, 3*, 281–290. https://doi.org/10.69554/jcfu2737

Bansal, G., Zahedi, F. M. & Gefen, D. (2016). Do context and personality matter? Trust and privacy concerns in disclosing private information online. *Information & Management*, *53*(1), 1–21. https://doi.org/10.1016/j.im.2015.08.001

Bastiaansen, H., Dalmolen, S., Kollenstart, M. & Van Engers, T. (2020). User-Centric Network-Model for Data Control with Interoperable Legal Data Sharing Artefacts. *Pacific Asia Conference On Information Systems*, 172. https://aisel.aisnet.org/ cgi/viewcontent.cgi?article=1171&context=pacis2020

Bechtsis, D., Tsolakis, N., Iakovou, E. & Vlachos, D. (2022). Data-driven secure, resilient and sustainable supply chains: gaps, opportunities, and a new generalised data sharing and data monetisation framework. *International Journal Of Production Research*, *60*(14), 4397–4417. https://doi.org/10.1080/00 207543.2021.1957506

Bélanger, F. & Crossler, R. E. (2011). Privacy in the Digital Age: A Review of Information Privacy Research in Information Systems. *MIS Quarterly*, *35*(4), 1017–1041. https://doi.org/10.2307 /41409971

Belanger, F., Hiller, J. S. & Smith, W. J. (2002). Trustworthiness in electronic commerce: the role of privacy, security, and site attributes. *The Journal Of Strategic Information Systems*, *11*(3–4), 245–270. https://doi.org/10.1016/s0963-8687(02)00018-5

Bitkom. (2024). Deutsche Unternehmen nutzen ihre Daten kaum. *Bitkom*.

https://www.bitkom.org/Presse/Presseinformation/Datenoekonomie-Deutschland-2024

Bons, R. W. H., Lee, R. M. & Nguyen, V. H. (2012). Generating Procedural Controls to Facilitate Trade: The Role of Control in the Absence of Trust. *BLED 2012 – Special Issue.* https://domino.fov.uni-mb.si/proceedings.nsf/Proceedings/1C08968438F3EC39C1257A5A004102AE/$File/09_bons.pdf

Brocke, J. V., Simons, A., Niehaves, B., Niehaves, B., Reimer, K., Plattfaut, R. & Cleven, A. (2009). Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process. *European Conference On Information Systems*, 2206–2217. http://dblp.uni-trier.de/db/conf/ecis/ecis2009.html#BrockeSNRPC09

Cheng, J.-H. & Du, T. C. (2015, December 6). *A Socio-Technical System Perspective Of Psychological Ownership Toward Sharing IoT Data In Supply Chains* [Conference paper]. The Fifteenth International Conference On Electronic Business, Hong Kong. https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1008&context=iceb2015

Choi, J. I. & Butler, K. R. B. (2019). Secure Multiparty Computation and Trusted Hardware: Examining Adoption Challenges and Opportunities. *Security And Communication Networks*, *2019*, 1–28. https://doi .org/10.1155/2019/1368905

Coltman, T., Tallon, P., Sharma, R. & Queiroz, M. (2015). Strategic IT Alignment: Twenty-Five Years on. *Journal Of Information Technology*, *30*(2), 91–100. https://doi.org/10.1057/jit.2014.35

Curry, E. (2012, July 16). *System of systems information interoperability using a linked dataspace* [Conference paper]. IEEE 7th International Conference On System Of Systems Engineering, Genova, Italy. https://www.researchgate.net/publication/259740266_System_of_Systems_Information_Interoperability_using_a_Linked_Dataspace

Dahlberg, T. & Nokkala, T. (2019, June 16). *Willingness to Share Supply Chain Data in an Ecosystem Governed Platform – An Interview Study* [Conference paper]. 32nd Bled eConference – Humanizing Technology For A Sustainable Society, Bled, Slovenia. https://www.researchgate.net/publication/335613440_Willingness_to_Share_Supply_Chain_Data_in_an_Ecosystem_Governed_Platform_-_An_Interview_Study#fullTextFileContent

Deutscher Bundestag. (2020). *US-Datenrecht. Zugriff US-amerikanischer Behörden auf Daten*. https://www.bundestag.de/resource/blob/796102/ea53ffe8e08a9ab11e270719263d8c53/WD-3-181-20-pdf-data.pdf

Dinev, T. & Hart, P. (2006). An Extended Privacy Calculus Model for E-Commerce Transactions. *Information Systems Research*, *17*(1), 61–80. https://doi.org/ 10.1287/isre.1060.0080

Dinev, T., McConnell, A. R. & Smith, H. J. (2015). Research Commentary—Informing Privacy Research Through Information Systems, Psychology, and Behavioral Economics: Thinking Outside the "APCO" Box. *Information Systems Research*, *26*(4), 639–655. https://doi.org/10.1287/isr e.2015.0600

Dyer, J. H. & Chu, W. (2003). The Role of Trustworthiness in Reducing Transaction Costs and Improving Performance: Empirical Evidence from the United States, Japan, and Korea. *Organization Science*, *14*(1), 57–68. https://doi.org/10.1287/orsc.14.1.57.12806

Enders, T., Benz, C., Schüritz, R. & Lujan, P. (2020, June 15). *How to Implement an Open Data Strategy? Analyzing Organizational Change Processes to Enable Value Creation by Revealing Data* [Conference Paper]. Proceedings Of 28th European Conference On Information Systems (ECIS 2020)., Marrakesh, Marokko. https://www.researchgate.net/publication/341821156_How_to_Implement_an_Open_Data_Strategy_Analyzing_Organizational_Change_Processes_to_Enable_Value_Creation_by_Revealing_Data#fullTextFileContent

Enders, T., Wolff, C. & Satzger, G. (2020, June 15). *Knowing What to Share: Selective Revealing in Open Data* [Conference paper]. Proceedings Of 28th European Conference On Information Systems (ECIS 2020), Marrakesh, Marokko. https://www.researchgate.net/publication/341447249_Knowing_What_to_Share_Selective_Revealing_in_Open_Data

Fassnacht, M., Benz, C., Leimstoll, J. & Satzger, G. (2023, December). *Is Your Organization Ready to Share? A Framework of Beneficial Conditions for Data Sharing* [Conference paper]. A Framework Of Beneficial Conditions For Data Sharing. Proceedings Of The 44th International Conference On Information Systems (ICIS). https://www.researchgate.net/publication/374422448_Is_Your_Organization_Ready_to_Share_A_Framework_of_Beneficial_Conditions_for_Data_Sharing

Fricker, S. & Maksimov, Y. (2017, June). *Pricing of Data Products in Data Marketplaces* [Conference paper]. International Conference On Software Business, Essen, Germany. https://www.researchgate.net/publication/318492802_Pricing_of_Data_Products_in_Data_Marketplaces

Garrido, G. M., Sedlmeir, J., Uludağ, Ö., Alaoui, I. S., Luckow, A. & Matthes, F. (2022). Revealing the landscape of privacy-enhancing technologies in the context of data markets for the IoT: A systematic literature review. *Journal Of Network And Computer Applications*, *207*, 103465. https://doi.org/10.1016/j.jnca.2022.103465

Gartner. (2024). Gartner 2024 Hype Cycle for Emerging Technologies Highlights Developer Productivity, Total Experience, AI and Security. *Gartner*. Retrieved on 15. September 2024, of https://www.gartner.com/en/newsroom/press-releases/2024-08-21-gartner-2024-hype-cycle-for-emerging-technologies-highlights-developer-productivity-total-experience-ai-and-security

GDPR. (n.d.). What is GDPR, the EU's new data protection law? *GDPR.EU*. https://gdpr.eu/what-is-gdpr/

Gelhaar, J. & Otto, B. (2020, June 5). *Challenges in the Emergence of Data Ecosystems* [Conference paper]. Twenty-Third Pacific Asia Conference On Information Systems, Dubai. https://www.researchgate.net/publication/341930759_Challenges_in_the_Emergence_of_Data_Ecosystems

Handley, S. M. & Benton, W. (2012). The influence of exchange hazards and power on opportunism in outsourcing relationships. *Journal Of Operations Management*, *30*(1–2), 55–68. https://doi.org/10.1016/j.jom.2011.06.001

Hansen, S. F., Von Krauss, M. K. & Tickner, J. A. (2008). The precautionary principle and risk-risk tradeoffs. *Journal Of Risk Research*, *11*(4), 423–464. https://doi.org/10.1080/13669870801967192

Hasani, T., Rezania, D., Levallet, N., O'Reilly, N. & Mohammadi, M. (2023). Privacy enhancing technology adoption and its impact on SMEs' performance. *International Journal Of Engineering Business Management*, *15*, 184797902311728. https://doi.org/10.1177/18479790231172874

Haupt, S., Lichter, J. & May, F. C. (2021). SORGFALTSPFLICHTEN ENTLANG GLOBALER LIEFERKETTEN: Eine ökonomische Analyse. In *Handelsblatt RESEARCH INSTITUTE*.

Helminger, L. & Rechberger, C. (2022). Multi-Party Computation in the GDPR. In *Springer eBooks* (S. 21–39). https://doi.org/10.1007/978-3-031-09901-4_2

Hoffmann, R., Kittel, B. & Larsen, M. (2020). Information exchange in laboratory markets: competitionn, transfer costs, and the emergence of reputation. *Experimental Economics*, *24*(1), 118–142. https://doi.org/10.1007/s10683-020-09652-0

Holler, M., Vogt, H. & Barth, L. (2019). Exploring the Willingness-to-Share Data of Digitized Products in B2B Manufacturing Industries. *BLED Proceedings*. https://doi.org/10.18690/978-961-286-280-0.56

IBM. (2023, 3. Juli). *Was bedeutet Data Mesh?* Retrieved September 6, 2024. https://www.ibm.com/de-de/topics/data-mesh

IBM. (2024). *Cost of a Data Breach Report*. https://table.media/wp-content/uploads/2024/07/30132828/Cost-of-a-Data-Breach-Report-2024.pdf

IEEE. (1990). *IEEE Standard Glossary of Software Engineering Terminology*. https://doi.org/10.1109/IEEESTD.1990.101064

Jussen, I., Schweihoff, J., Dahms, V. & Möller, F. (2023, January). *Data Sharing Fundamentals: Definition and Characteristics* [Conference paper]. Proceedings Of The 56th Hawaii International Conference On System Sciences (HICSS), Maui, Hawaii. https://www.researchgate.net/publication/363769417_Data_Sharing_Fundamentals_Definition_and_Characteristics

Kahneman, D. & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*(2), 263. https://doi.org/10.2307/1914185

Kanger, L. & Pruulmann-Vengerfeldt, P. (2015). Social Need for Secure Multiparty Computation. *Applications Of Secure Multiparty Computation*, 43–57. https://doi.org/10.3233/978-1-61499-532-6-43

Körner, M., Sedlmeir, J., Weibelzahl, M., Fridgen, G., Heine, M. & Neumann, C. (2022). Systemic risks in electricity systems: A perspective on the potential of digital technologies. *Energy Policy*, *164*, 112901. https://doi.org/10.1016/j.enpol.2022.112901

Krasnova, H., Spiekermann, S., Koroleva, K. & Hildebrand, T. (2010). Online social networks: Why we disclose. *Journal Of Information Technology*, *25*(2), 109–125. https://doi.org/10.1057/jit.2010.6

Li, S. & Lin, B. (2006). Accessing information sharing and information quality in supply chain management. *Decision Support Systems*, *42*(3), 1641–1656. https://doi.org/10.1016/j.dss.2006.02.011

Lis, D. & Otto, B. (2020, August 10). *Data Governance in Data Ecosystems – Insights from Organizations* [Conference paper]. Americas Conference On Information Systems (AMCIS). https://www.researchgate.net/publication/343215188_Data_Governance_in_Data_Ecosystems_-_Insights_from_Organizations

Lumineau, F., Wang, W. & Schilke, O. (2020). Organizational Trust in the Age of the Fourth Industrial Revolution. *Journal Of Management Inquiry*. https://doi.org/10.1177/10564926221127852

Mayer, R. C., Davis, J. H. & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy Of Management Review*, *20*(3), 709–734. https://doi.org/10.5465/amr.1995.9508080335

McKnight, D. H., Cummings, L. & Chervany, N. (1996). Initial Trust Formation in New Organisational Relationships. *Academy Of Management Review*. https://doi.org/10.2307/259290

Meuser, M. & Nagel, U. (1991). ExpertInneninterviews - vielfach erprobt, wenig bedacht: ein Beitrag zur qualitativen Methodendiskussion. *Qualitativ-empirische Sozialforschung: Konzepte, Methoden, Analysen*. Westdt. Verlag.

Müller, J. M., Buliga, O. & Voigt, K. (2018). Fortune favors the prepared: How SMEs approach business model innovations in Industry 4.0. *Technological Forecasting And Social Change*, *132*, 2–17. https://doi.org/10.1016/j.techfore.2017.12.019

Müller, J. M., Veile, J. W. & Voigt, K. (2020). Prerequisites and incentives for digital information sharing in Industry 4.0 – An international comparison across data types. *Computers & Industrial Engineering*, *148*, 106733.

https://doi.org/10.1016/j.cie.2020.106733

Myers, M. D. & Newman, M. (2007). The qualitative interview in IS research: Examining the craft. *Information And Organization*, *17*(1), 2–26. https://doi.org/10.1016/j.infoandorg.2006.11.001

OECD. (2023). *EMERGING PRIVACY ENHANCING TECHNOLOGIES: CURRENT REGULATORY AND POLICY APPROACHES.* https://www.oecd.org/en/publications/emerging-privacy-enhancing-technologies_bf121be4-en.html

Pavlou, N., Liang, N. & Xue, N. (2007). Understanding and Mitigating Uncertainty in Online Exchange Relationships: A Principal-Agent Perspective. *MIS Quarterly*, *31*(1), 105. https://doi.org/10.2307/25148783

Penttinen, E., Halme, M., Lyytinen, K. & Myllynen, N. (2018). What Influences Choice of Business-to-Business Connectivity Platforms? *International Journal Of Electronic Commerce*, *22*(4), 479–509. https://doi.org/10.1080/10864415.2018.1485083

Peppard, J. (2018). Rethinking the concept of the IS organization. *Information Systems Journal*, *28*(1), 76–103. https://doi.org/10.1111/isj.12122

Poppo, L. & Zenger, T. (2002). Do formal contracts and relational governance function as substitutes or complements? *Strategic Management Journal*, *23*(8), 707–725. https://doi.org/10.1002/smj.249

Sangers, A., Van Heesch, M., Attema, T., Veugen, T., Wiggerman, M., Veldsink, J., Bloemen, O. & Worm, D. (2019). Secure Multiparty PageRank Algorithm for Collaborative Fraud Detection. In *Lecture notes in computer science* (S. 605–623). https://doi.org/10.1007/978-3-030-32101-7_35

SAP. (2024, 27. März). *Fully homomorphic encryption: data insights without sharing data*. SAP News Center. https://news.sap.com/2024/03/fully-homomorphic-encryption-insights-without-sharing-data/?source=email-sapflash-topic1-20240415&sap-outbound-id=A0F2ECD0322E162817289869CD9C9D1FE3C95F6C&smc_campaign_id=0000040773&source=email-smc

Scheibner, J., Raisaro, J. L., Troncoso-Pastoriza, J. R., Ienca, M., Fellay, J., Vayena, E. & Hubaux, J. (2021). Revolutionizing Medical Data Sharing Using Advanced Privacy-

Enhancing Technologies: Technical, Legal, and Ethical Synthesis. *Journal Of Medical Internet Research*, *23*(2), e25120. https://doi.org/10.2196/25120

Seničar, V., Jerman-Blažič, B. & Klobučar, T. (2003). Privacy-Enhancing Technologies—approaches and development. *Computer Standards & Interfaces*, *25*(2), 147–158. https://doi.org/10.1016/s0920-5489(03)00003-5

Skinner, G., Han, S. & Chang, E. (2006). An information privacy taxonomy for collaborative environments. *Information Management & Computer Security*, *14*(4), 382–394. https://doi.org/10.1108/09685220610690835

Smith, H. J., Milberg, S. J. & Burke, S. J. (1996). Information Privacy: Measuring Individuals' Concerns about Organizational Practices. *MIS Quarterly*, *20*(2), 167. https://doi.org/10.2307/249477

Smith, N., Dinev, N. & Xu, N. (2011). Information Privacy Research: An Interdisciplinary Review. *MIS Quarterly*, *35*(4), 989. https://doi.org/10.2307/41409970

Tong, P. Y. & Crosno, J. L. (2016). Are information asymmetry and sharing good, bad, or context dependent? A meta-analytic review. *Industrial Marketing Management*, *56*, 167–180. https://doi.org/10.1016/j.indmarman.2015.11.004

United Nations. (2023). *The Pet Guide: The united nations guide on privacy-enhancing technologies for official statistics.* https://unstats.un.org/bigdata/task-teams/privacy/guide/

United States Government. (2022, June 9). *Request for Information on Advancing Privacy-Enhancing Technologies*. Federal Register. https://www.federalregister.gov/documents/2022/06/09/2022-12432/request-for-information-on-advancing-privacy-enhancing-technologies#footnote-2-p35251

Wang, Z., Ye, F. & Tan, K. H. (2014). Effects of managerial ties and trust on supply chain information sharing and supplier opportunism. *International Journal Of Production Research*, *52*(23), 7046–7061. https://doi.org/10.1080/00207543.2014.932931

Webster, J. & Watson, R. T. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly*, *26*(2), 13–23. http://www.jstor.org/stable/4132319

Wen, Z. A., Jia, J., Yan, H., Yao, Y., Liu, Z. & Dong, C. (2023). The influence of explanation designs on user understanding differential privacy and making data-sharing decision. *Information Sciences*, *642*, 118799. https://doi.org/10.1016/j.ins.2023.03.024

Witte, A.-K., Fuerstenau, D. & Zarnekow, R. (2020, December). *Digital Health Ecosystems for Sensor Technology Integration - A Qualitative Study on the Paradox of Data Openness* [Conference paper]. Forty-First International Conference On Information Systems (ICIS), India. https://www.researchgate.net/publication/344407999_Digital_Health_Ecosystems_for_Sensor_Technology_Integration_-_A_Qualitative_Study_on_the_Paradox_of_Data_Openness

World Economic Forum. (2020). *Cyber Information Sharing: Building Collective Security*. https://www.weforum.org/publications/cyber-information-sharing-building-collective-security/

World Economic Forum. (2024). *Top 10 Emerging Technologies of 2024*. https://www.weforum.org/publications/top-10-emerging-technologies-2024/

Xu, H., Dinev, T., Smith, J. & Hart, P. (2011). Information Privacy Concerns: Linking Individual Perceptions with Institutional Privacy Assurances. *Journal Of The Association For Information Systems*, *12*(12), 798–824. https://doi.org/10.17705/1jais.00281

Zhao, C., Zhao, S., Zhao, M., Chen, Z., Gao, C., Li, H. & Tan, Y. (2019). Secure Multi-Party Computation: Theory, practice and applications. *Information Sciences*, *476*, 357–372. https://doi.org/10.1016/j.ins.2018.10.024

# Appendices

## Appendix A

### A.1 Concept Matrix 'Corporate Data Sharing'

| Literature | Concept matrix 1 (Corporate data sharing) | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Inter-oper-ability | Trust | Control | Risks/ costs | Benefits | Internal factors | | | | External factors | | | |
| | | | | | | Data Governance | Top management | Resources | Data type | Region | Sector | External pressure | Situational opportunity |
| Dahlberg & Nokkala, 2019 | x | | | x | x | | | x | x | | x | x | x |
| Penttinen et al., 2018 | | | | x | x | | | | | | | | |
| Gelhaar & Otto, 2020 | | x | | x | x | | | | | | | | |
| Müller et al., 2018 | x | | | x | x | | | x | | | | | |
| Bechtsis et al., 2022 | | | | | x | | | | | | | | |
| Müller et al., 2020 | | x | | | x | | | x | x | x | | | |
| Enders et al., 2020 | | | | | x | | | | | | | | |
| Hoffmann et al., 2020 | | | | x | | | | | | | x | | |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Agahari et al., 2022 | | x | x | x | | | | | | | x | | |
| Holler et al., 2019 | | x | x | | | | | | | | | x | |
| Li & Lin, 2006 | | x | | | | | x | | | | | | |
| Dyer & Chu, 2003 | | x | | | | | | | | | x | | |
| Bons et al., 2012 | | | x | | | | | | | | | | |
| Lis & Otto, 2020 | x | | x | | | x | | | | | | | |
| Bastiaansen et al., 2020 | x | | | | | | | | | | | | |
| Fassnacht et al., 2023 | x | | | | | x | x | | x | | | x | |
| Witte et al., 2020 | x | | | | | | | | | | x | | |
| Enders & Benz et al., 2020 | | | | | | | x | | | | | | |
| Wang et al., 2014 | | x | | | | | x | | | | | | |
| Enders et al., 2020 | | | | | | | | | x | | | | |

## A.2 Concept Matrix ‘Corporate Data Sharing with PETs’

| Literature | Concept matrix 2 (Corporate data sharing with PETs) | | | | | |
|---|---|---|---|---|---|---|
| | Trust | Control | Risks/ costs | Benefits | Internal factors | External factors |

| | | | | | Organisational readiness | Top management | Resources | Data type | Region | Sector | External pressure | Situational opportunity |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Garrido et al., 2022 | | x | x | | | | | | | | | |
| Agahari et al., 2022 | x | x | x | x | x | | | x | | | | |
| Agahari & Reuver, 2022 | x | x | x | | | | | | | | | |
| Wen et al., 2023 | x | | | | | | | | | | | |
| Lumineau et al., 2020 | x | | | | | | | | | | | |
| Körner et al., 2022 | x | | x | | | | x | | | | | |
| Kanger & Pruulmann-Vengerfeldt, 2015 | | | x | x | x | x | x | x | | | x | x |
| Hasani et al., 2023 | | | | x | x | x | x | | | x | x | |
| Agahari et al., 2021 | | | | x | | | | | | | | |
| Bamford, 2020 | | | | | | x | | | | | x | |
| Agrawal et al., 2021 | | | | | | | x | | | | | |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Scheibner et al., 2021 | | | | | | | | | | | x | |
| Helminger & Rechenberger, 2022 | | | | | | | | | | | x | |
| Witte et al., 2020 | | | | | | | | | | x | | |
| World Ecomic Forum, 2024 | | | | | | | | | x | | | |

## Appendix B Interview Questionnaire

**Interview - Masterthesis „CORPORATE DATA SHARING: A QUALITATIVE ANALYSIS OF FACTORS INFLUENCING DATA SHARING AND THE IMPACT OF PRIVACY ENHANCING TECHNOLOGIES“**

**Introduction:**

- Interviewer/interviewee introduces him/herself
- Presentation of the topic
- Note on anonymization for use in the Master's thesis and question about transcription of the interview

**Part 1: General Questions RQ1**

1) From your experience: With whom do companies share data and for what purpose? Are there contracts/agreements that regulate the exchange of data?
2) Would you consider trust between companies or control options to be a more important factor in data exchange decisions? Why?
3) Would you say that the decision of companies to share data is a subjective or objective decision-making process?
4) What dangers/risks do you see in sharing data? Do you think the risks of data sharing are overestimated/underestimated by companies?
5) What are the prerequisites for successful data exchange between companies?

6) Are there external/internal influences that encourage companies to share their data?
7) What are the reasons/barriers for companies not sharing data?

**Part 2: Questions regarding Privacy Enhancing Technologies (PETs) (especially Secure Multiparty Computation)**

1) For which use cases are privacy enhancing technologies such as SMPC useful?
2) Will privacy-enhancing technologies such as SMPC change the data sharing behavior of companies? If so, in what way?
3) Do the influencing factors of trust and control in data sharing decisions change with the use of PETs?
4) What conditions are needed for companies to invest in PETs? Which type of company is more willing to invest in PETs?
5) How do PETs change the risks of data exchange? Do you see any new risks that could arise from the use of PETs?
6) What do you see as the biggest challenges to the acceptance of PETs on the market? Or reasons against investing in PETs?

**Conclusion:**

- Now that you know more about my topic: Is there anything else I could have asked? Or do you have anything in mind that has not yet been mentioned and would be important for my topic?
- Can you recommend other interview partners?

## Appendix C Interview Analysis

| ID | Interview statement/paraphrase | Headings/Coding | Concept |
|---|---|---|---|
| B1 | Control is more important than trust, as many things can be precisely defined via contracts. For example, to which countries data may be sent (10). | Control through contracts | Control |
| B1 | Trust through long-term relationships can support, but control is more important (12). | Trust through relationships helpful, but control more important | Trust; Control |
| B1 | Top management can make the decision-making process more subjective, but mainly objective by weighing up costs and benefits (14). | Decision-making process largely based on cost-benefit considerations | Decision rationality; Internal factor (top management) |

| | | | |
|---|---|---|---|
| B1 | In the corporate environment, people are generally always risk-averse when it comes to sharing data. People always prefer to be on the safe side (16). | Companies risk-averse in data exchange | Risk |
| B1 | Legal risk of fines or reputational damage greatest. Data leaks can cause reputational damage that costs money or revenue. Fines due to non-compliance with the GDPR (20). | Reputational damage and fines as important legal risks | Risk |
| B1 | Risk when exchanging data where high investment sums or trade secrets are involved. This can destroy competitive advantages (22). | High risk for data in connection with high investment sums or trade secrets | Risk |
| B1 | High data quality through, for example, up-to-dateness is an important prerequisite for successful data exchange (24). | Data quality as a prerequisite for successful data exchange | Interoperability |
| B1 | Another important requirement is interoperability. If the Chinese government prevents the transfer of data, this naturally hinders data exchange (24). | Interoperability as a prerequisite for successful data exchange | Interoperability |
| B1 | Governments and their regulations are an external influence for data exchange (26). | Regulations as an external influence of data exchange | External factor (external pressure) |
| B1 | Internal influences are the willingness of employees, especially managers, to share data (26). | Influence of managers | Internal factor (top management) |
| B1 | People are initially more cautious about PETs because it is not yet clear what the tool will do. What benefit does it have for me? When you see use cases that work, such as benchmarking, you become more open to technology (36). | Clear use cases for the acceptance of PETs | Benefits (PETs) |

| | | | |
|---|---|---|---|
| B1 | If you have the certainty that technology works, you are more willing to share data. Of course, it must be ensured that you don't unintentionally share data with competitors or violate antitrust law (40). | After clarifying the uncertainties of these technologies, it could lead to increased data exchange | Risk (PETs) |
| B1 | It must be trusted that all actors involved contribute equally to the data exchange process. No data is withheld or unequal amounts of data are fed in (46). | Trust in all stakeholders involved in the data exchange with PETs | Trust (PETs) |
| B1 | Great benefits of PETs for pharmaceutical companies. This is because a large number of clinical studies, tests and analyses are required for new drug approvals. PETs could simplify approval procedures (50). | Drug approval as a use case for PETs | External factor PETs (sector) |
| B1 | The big challenge is to understand the technology, to know how it works and what benefits it brings. Only then are people prepared to invest in such technologies (54). | Technology understanding as a challenge | Acceptance of PETs |
| B1 | Depending on the type of data, there are different regulatory hurdles that need to be taken into account. These are primarily determined by the legal framework, such as the AI Act or the Digital Services Act (58). | Regulation of different types of data influences data exchange | Internal factor (data type) |

| **ID** | **Interview statement/paraphrase** | **Headings/Coding** | **Concept** |
|---|---|---|---|
| B2 | Data is exchanged with suppliers. It is important that suppliers see all | Encrypted data exchange reduces benefits | Benefits (PETs) |

| | | | |
|---|---|---|---|
| | data. Nothing can be done with encrypted data (4). | | |
| B2 | If competitors want to acquire another company, data about this company is shared with competitors. Non-disclosure agreements must be signed (4). | Data exchange with competitors with non-disclosure agreements | Control |
| B2 | Companies trust us as an intermediary for data exchange (6). | Trust in the central party | Trust |
| B2 | Contracts are absolutely standard. They are automatically concluded with every customer. This is a basic requirement before any data is exchanged (8). | Contracts as a basic prerequisite for data exchange | Control |
| B2 | Data exchange is very important, especially in the construction industry where many customers rely on data exchange. Greater risk of not exchanging data (16). | Risk if no data is exchanged | Risk |
| B2 | Company sales are a very sensitive area. You have to disclose sensitive information about companies to prospective buyers without knowing for sure whether they will actually buy them (18). | Risk in the sale of companies | Risk |
| B2 | No real use cases for PETs. If he were to offer this, he would not know who would be interested (30). | No use cases for PETs | Benefits (PETs) |
| B2 | Companies also have little knowledge that data can be exchanged without disclosing it (34). | Lack of knowledge about PETs | Internal factor PETs (resources) |
| B2 | If PETs ensures that data is not disclosed, trust naturally becomes easier for everyone (38). | PETs can increase confidence | Trust (PETs) |

| | | | |
|---|---|---|---|
| B2 | Implementation of PETs is trivial, anyone could introduce this (40). | No special requirements for the implementation of PETs | Internal factor PETs (resources) |

| **ID** | **Interview statement/paraphrase** | **Headings/Coding** | **Concept** |
|---|---|---|---|
| B3 | [Company] offers its own platform on which data can be exchanged (16). | Platform for cross-company data exchange | Interoperability |
| B3 | In any collaboration between companies, even if no data is exchanged, contracts are drawn up between companies, such as non-disclosure agreements (16). | Contracts | Control |
| B3 | Data sharing decisions are always reviewed by at least two people (4 eyes principle), but usually many more people are involved. Entire departments are even involved in supplier audits (18). | Involvement of many people in the decision on data exchange | Decision rationality |
| B3 | The decision-making process is very rational because everyone is very cautious and reluctant to take individual responsibility. Feelings are pushed to one side and people want to be completely sure that everything fits before data is shared. You look so objectively at how a person can do it before something is shared (20). | Maximum rationality in the decision-making process | Decision rationality |
| B3 | Of course, the things that are shared the most are those that have no business value. Nobody wants to share business secrets. The most sharing can be done in the area of cyber security. It's not about business secrets, it's about | Greatest potential of data exchange in the area of cyber security | Internal factor (data type) |

| | | | |
|---|---|---|---|
| | defending ourselves together against the “bad guys” out there. It's more about sharing things that benefit everyone without giving the competitor an advantage. This mainly concerns threat intelligence (24). | | |
| B3 | It's not like you can say a manager trusts a company more and therefore we share more. Many factors are taken into account, such as the size of the company, the country in which the company operates and how this country is assessed by the relevant department. In addition, whether or not they are competitors in certain areas, how many employees they have, when they were founded. All these factors come into play and determine how we work with the company. The decision-making process cannot therefore be reduced to the “relationships” factor alone (28). | More factors than just “relationships” are important for data exchange between companies | Trust |
| B3 | Wouldn't call it trust between companies, but rather risk assessment (30). | More risk assessment than trust | Trust |
| B3 | External pressure is a decisive factor in sharing data (34). | External pressure | External factor (external pressure) |
| B3 | You shouldn't just say, “OK, that's a cool concept, let's apply it somewhere” and then look for a use case. There are a thousand cool concepts out there, and if | Benefits of the technology are crucial | Benefits (PETs) |

| | | | |
|---|---|---|---|
| | you start like that, you'll never finish and never really make any money. You need a problem for which this technology is the solution, or a use case where you can apply it directly (38). | | |
| B3 | The benefits of PETs do not usually outweigh the costs. Large companies usually already have enough data. It is rare to need additional data from external sources. It is much more important to find ways to make better use of the data that you produce yourself, to store it and then work with it. If you share data with other companies, you always run the risk of other companies contributing less data or data that is not of the same high quality. It is perhaps easier to optimize your own data collection processes first (40). | For large companies, it makes more sense to make better use of their own data than to accept the risks of data exchange | Benefits (PETs); Risk (PETs) |
| B3 | SMPC is difficult because it often requires collaboration with competitors, which companies want to avoid in order to differentiate themselves. The current focus is on product and process optimization, automation, logging processes and predictive maintenance. In competitive areas, companies would rather invest in their own solutions than share data and make progress together (44). | Competition also prevents companies from exchanging data with PETs | Benefits (PETs) |

| | | | |
|---|---|---|---|
| B3 | Cyber security is something where we would also like to improve together. Could be an interesting use case for PETs (44). | Cyber security as a use case for PETs | Benefits (PETs) |
| B3 | The use of PETs to meet legal requirements could be useful. What is important here is that it is 100% certain that all legal requirements are met. Because any benefit that could be derived from it can never compensate for the penalties. One doubt would be enough to reject it. It would have to be approved by experts from your own company that technology fulfills its purpose. A sales person who wants to sell technology would never be believed. Confirmation is expected from internal experts, legislators or independent third parties (46; 48). | PETs are only introduced when it is 100% certain that the technology can be trusted. | Trust (PETs) |
| B3 | Situational circumstances can influence companies to exchange data. For example, the German automotive industry sharing data on specific topics in response to strong competition from abroad. Common goals are crucial for data exchange. If the point is leveraged that you want to stand out from each other and you want to stand out together in one aspect, such as from common enemies or legislation, then data | Common goals are crucial for the exchange of data | External factor (situational opportunity) |

| | | | |
|---|---|---|---|
| | exchange becomes meaningful (50). | | |

| ID | Interview statement/paraphrase | Headings/coding | Concept |
|---|---|---|---|
| B4 | Companies want to know how effective their advertisements are. Were they targeting the right audience. Traditionally it's done by third party cookies. So it's a piece of data that tracks you around. Now the usage is way more reduced. So they are looking for alternative ways to solve this problem of measuring how efficient was the campaign (6). | Due to external regulations gets data sharing more important. | External factor (external pressure) |
| B4 | Trust is less important, risks and commercial point of view more important when sharing data (8; 10). | Trust less important | Trust |
| B4 | Decision about sharing data is subjective, because you reach out typically to who you can reach out and depends whether you convince them or not (12). | Subjective decision process when sharing data | Decision rationality |
| B4 | Companies need to be open minded. They need to have some technical understanding. You need to be convincing in a sense, honest and talk in an understandable language when you want to convince them of SMPC (16) | Open minded and technological understanding important factor for adoption of PETs | Internal factor PETs (corporate culture) |
| B4 | Typically companies want some economical benefit, for example deeper insights, when sharing data (24). | Deeper insights as benefit of data sharing | Benefits |
| B4 | With PETs you can get accurate aggregated analytics of digital advertisements. This technology can replace third party cookies in a way. So one use case is marketing (26). | Marketing as use case for PETs | Benefits (PETs) |

| | | | |
|---|---|---|---|
| B4 | PETs have many use cases, but not many companies use them yet, because People don't understand them. They look for a simple solution, but don't recognize in PETs a clear solution. So reason is technical difficulty and lack of trust (32). | Challenges for PETs are technical difficulty and lack of trust | Trust (PETs) |
| B4 | There is this distinction between soft privacy and hard Privacy. Soft privacy for me is a signing, agreements and non-disclosures and trusting the other party if you share data with somebody. And hard privacy with PETs is removing the possibility for the other party to use data the wrong way. So they reduce risks (36; 40). | PETs create hard privacy and reduce risks | Risk (PETs) |
| B4 | But you need to trust the people enough who offer such new technologies (36). | Trust in supplier of PETs | Trust (PETs) |
| B4 | PETs more for big companies because it's a huge investment in development in technical skills (36). | Resources important for adoption of PETs | Internal factor PETs (resources) |

| **ID** | **Interview statement/paraphrase** | **Headings/coding** | **Concept** |
|---|---|---|---|
| B5 | The biggest current driver for data exchange between companies is the Supply Chain Act (4). | Supply chain law as an influencing factor for data exchange | External factor (external pressure) |
| B5 | Supply Chain Act as a welcome means for companies to better control suppliers, have better data on them and spur competition (8). | Supply chain law as an influencing factor for data exchange | External factor (external pressure) |
| B5 | Trust is a key factor when it comes to data sharing. The biggest discrepancy between where data could be shared and used to generate value and where it actually happens is in SMEs and small and medium-sized enterprises. They do not use data sharing offers out of fear. Laws | Lack of trust when sharing data, especially in SMEs | Trust |

| | | | |
|---|---|---|---|
| | such as the US Cloud Act unsettle companies and reduce trust (10). | | |
| B5 | Trust is an important factor. Control naturally depends on this (10). | Control depends on the level of trust | Trust; control |
| B5 | GDPR is a law that, according to the German interpretation, often initially leads to nothing being done. Especially for large companies. They prefer to be on the safe side and not share anything. Advantages that could arise are often ignored (10). | GDPR leads to caution when exchanging data | External factor (external pressure) |
| B5 | Trust is an increasingly important factor when sharing data, unless you use so-called zero trust solutions such as PETs, which use technology to replace this trust (10). | PETs replace trust with technology | Trust (PETs) |
| B5 | The caution generated by the GDPR leads to a certain subjectivity when deciding whether to share data or not. Initially cautious and negative attitude. Risks are overestimated. However, we are in the process of turning to a more objective view (12). | Subjective influence in decisions regarding data exchange | Decision rationality |
| B5 | New EU laws such as DataAct, which ensure that companies must share data in the event of a disaster (14). | Data Act as a driver of data exchange | External factor (external pressure) |
| B5 | Risks such as the disclosure of business secrets are present in data exchange, but are overestimated. The majority of data generated in general business operations is harmless (20). | Overestimating the risks in data exchange | Risk |
| B5 | Legal certainty is a reduction in legal risk. There are still few contracts that have actually been challenged in court. There are few reports on standard contracts | Legal certainty is a reduction in legal risk | Risk |

| | | | |
|---|---|---|---|
| | where you can be sure that you are legally protected (22). | | |
| B5 | Companies are trying to develop seals to ensure that data has a certain quality. A standard format for data is also an important prerequisite for data exchange (24). | Data quality as a prerequisite for data exchange | Interoperability |
| B5 | Data exchange easier when companies have a high level of digital infrastructure. Machines that have certain interfaces are also important (26). | Digital infrastructure and interfaces important for data exchange | Interoperability; digital infrastructure |
| B5 | Decision-making structures within a company are influencing factors. If only one person in the company makes decisions, this person must be convinced; if there is an extended management team, the situation is different (26). | Decision-making structures within a company | Internal factor (top management) |
| B5 | Sector is a very decisive influence in the exchange of data. The financial and healthcare sectors are highly regulated. Accordingly, data exchange is much more complicated (28). | Sector is a very decisive influence in the exchange of data | External factor (sector) |
| B5 | A use case for PETs would be fraud detection for money laundering offenses. This requires a large number of account movements to be tracked, which is only possible if banks exchange data with each other. Technically and legally it would have worked, but the compliance department was an obstacle (30; 38). | Use case for data exchange with PETs | Benefits (PETs) |
| B5 | So ultimately, and this is the funny thing about these zero trust technologies, we have to trust this technology. Even if we developed it ourselves, the others didn't develop it. So someone has to trust this technology and the providers (40). | Trust is still crucial for zero trust technologies | Trust (PETs) |

| | | | |
|---|---|---|---|
| B5 | The more data sharing becomes an issue, the more people talk about it and the more established PETs become, the greater the trust will be. This also increases the exchange of data (42). | Presence in society increases trust | Trust (PETs) |
| B5 | With the spread of PETs, trust is also increasing (46). | Spread of PETs increases confidence | Trust (PETs) |
| B5 | The more sensitive the data, the greater the need for PETs. Especially in the financial, healthcare and education sectors (48). | Need for PETs depends on the sensitivity of the data | Internal factor PETs (data type); External factor PETs (sector) |
| B5 | The more sensitive the data, the greater the need for PETs. Especially in the financial, healthcare and education sectors (48). | Demand for PETs depending on sector | External factor PETs (sector) |
| B5 | Reverse engineering and difficulty in verifying how safe the technology offered really is as risks of PETs (50). | Reverse engineering and difficult testing of technologies as risks of PETs | Risk (PETs) |

| **ID** | **Interview statement/paraphrase** | **Headings/coding** | **Concept** |
|---|---|---|---|
| B6 | When exchanging data, always use non-disclosure agreements (8). | Non disclosure agreements in data exchange | Control |
| B6 | Subjective decision-making process for data exchange. Hyperscalers frowned upon in German-speaking countries (10). | Subjective decision-making process for data exchange | Decision rationality |
| B6 | Data protection violations are subject to high penalties under legislation. Industrial espionage in connection with the Patriot Act is also a risk of data exchange (14; 16). | Industrial espionage and breaches of the law as risks | Risk |
| B6 | Everything is becoming more and more transparent due to publication laws. Companies have to publish data on certain things (18). | Laws for increased transparency increase data disclosure | External factor (external pressure) |

| | | | |
|---|---|---|---|
| B6 | German car manufacturers joining forces with regard to electromobility in order to gain a competitive advantage over competitors from other countries (20). | Association of companies against macroeconomic developments | External factor (situational opportunity) |
| B6 | PETs can be used for benchmarking purposes (26). | PETs for benchmarking | Benefits (PETs) |
| B6 | Legislation plays a major role in data exchange. Laws often formulated very vaguely. This creates uncertainty. Legislation is a much stronger influencing factor than the technology used (28). | Legislation plays a major role in data exchange | External factor (external pressure) |
| B6 | The risk with PETs is that the technology is not fully developed and there are safety leaks (32). | Risk in technology | Risks (PETs) |
| B6 | For the market, however, I actually see the risk that we could drift even further into this kind of price squeeze. With more opportunities for optimization, service providers could also be put under even more pressure (32). | Price pressure Risk | Risks (PETs) |
| B6 | In my view, it's not about the technology, but about the willingness to share data (40). | Type of technology secondary for data exchange | Acceptance of PETs |

| **ID** | **Interview statement/paraphrase** | **Headings/coding** | **Concept** |
|---|---|---|---|
| B7 | There was a use case for data sharing between banks with PETs. Didn't work in the end, although the technology for secure sharing was available. Banks had long discussions about correct data, data quality and could not agree. In the end, too much skepticism/fear (6). | Skepticism about data exchange despite technology | Trust (PETs) |
| B7 | There is no getting around the issue of control when exchanging data, it is essential (8). | Control as an important factor | Control |

| | | | |
|---|---|---|---|
| B7 | The banking example clearly shows that data sharing is a subjective decision-making process. There was nothing wrong with the technology and yet banks did not want to share the data (10). | Subjective decision-making process | Decision rationality |
| B7 | Regulations often still prescribe something to be done, even though it is outdated. Risk officers adhere to them, even though other solutions would be better. “you never get fired for buying IBM”. This also hinders data exchange (12; 14). | Safety before the risk of new technologies | Decision rationality |
| B7 | Trust through functioning use cases is a prerequisite for data exchange. When companies see that something works in small use cases over a certain period of time, they transfer it to larger cases (16). | Use cases generate trust | Trust |
| B7 | Fear of regulation, loss of reputation and initial adaptation as risks (18). | Regulatory risks, loss of reputation and initial adaptation | Risk |
| B7 | Top managers are very important when deciding on data sharing. If someone is more open to the topic or more knowledgeable, it is much more likely that data will be shared (20). | Openness of top management as an important influence | Internal factor (top management) |
| B7 | For example, the banking sector is highly regulated and there is a high risk of reputational damage. For this reason, there are probably greater concerns in this sector (22). | Influence of the sector on data sharing behavior | External factor (sector) |
| B7 | The sectors that are the most regulated, such as healthcare or the financial sector, would actually have the highest potential (24). | Regulation as an influencing factor | External factor (external pressure) |
| B7 | The most important thing is not to focus on PETs as a technology, but on the use | Use case must be the focus, not technology | Acceptance of PETs |

| | | | |
|---|---|---|---|
| | case. Because if you advertise as a provider with technology, you only have a very small circle of companies that can understand the topic at all. However, if you advertise a use case and prove that it can be implemented using PETs, you have a better chance (36). | | |
| B7 | So far, there is still too little evidence that MPC really works well because too few companies actually use it. With other technologies, the risk can be assessed much better. With new processes, all security people are very critical, especially if it is a black box process like MPC that we don't know much about (44). | Lack of proof of safety as a challenge | Acceptance of PETs |
| B7 | Financial size is not so decisive for the implementation of technology. More the expertise in the company that can evaluate the technology to some extent (50). | Company expertise important for evaluating the technology. Financial resources less important. | Internal factors PETs (resources) |
| B7 | PETs could be introduced via consulting companies. These are often well-known and generate the necessary trust that perhaps cannot be generated by PET providers (50; 52). | Consulting companies as technology brokers | Acceptance of PETs |

| **ID** | **Interview statement/paraphrase** | **Headings/coding** | **Concept** |
|---|---|---|---|
| B8 | Trust is a very important factor when companies exchange data. Of course, there are also instruments such as certifications that attempt to increase trust, but ultimately you have to trust | Trust is the basis for data exchange | Trust |

| | | | |
|---|---|---|---|
| | that companies will not misuse the data (7). | | |
| B8 | Trust is created at management level when it comes to working together as partners and through review processes such as contracts and questionnaires (9). | Trust through relationships at management level and through review processes | Trust |
| B8 | Both the subjective and objective decision-making process also depends on the size of the company. Larger groups require longer risk analyses (13). | Larger companies are usually more objective in their decision-making processes | Decision rationality |
| B8 | A core problem is that as soon as data is passed on, the other company must operationalize just as many security measures as you do, otherwise data is at risk. That's why we don't want to receive any data from others so that we don't have to bear this responsibility (11, 15). | Risk that other companies have fewer security precautions for data | Risk |
| B8 | The prerequisites for data sharing are contractual safeguards and the establishment of general control measures. The implementation and safeguarding of technical interfaces is also an important prerequisite (17). | Contractual safeguards, control measures and technical interfaces as a prerequisite | Control; interoperability |
| B8 | Sensitivity of data influences companies in their decision to share it (19). | Sensitivity of the data | Internal factor (data type) |
| B8 | Use case for PETs: Companies need to set up Security Operations Centers to continuously monitor their networks. This is incredibly expensive and know-how intensive, you need security specialists for this. So there are initiatives and ideas, for example, that | PETs for IT security optimizations | Benefits (PETs) |

| | | | |
|---|---|---|---|
| | one company does not do this alone, but that companies join forces and there is a network that provides this service, so to speak. But data is highly sensitive and no company wants to simply make it available. PETs can also be useful for thread intelligence. Sensitive data from typical attack patterns could be shared securely (21). | | |
| B8 | PETs as a confidence-building measure that ensures confidence by means of technical instruments (27). | PETs as a confidence-building measure | Trust (PETs) |
| B8 | PETs generally make sense in many sectors, but in most cases companies still lack the knowledge (29). | Most companies lack knowledge about PETs | Internal factor PETs (resources) |
| B8 | Risk that technical quality and implementation of PETs is insufficient (31). | Risk to the technical quality of PETs | Risk (PETs) |
| B8 | There must be evidence, proof that the technology works. This generates trust and only then will PETs be accepted. Could be described as a chicken and egg problem (33, 35). | Confidence in PETs through evidence that they work | Trust (PETs) |

| **ID** | **Interview statement/paraphrase** | **Headings/coding** | **Concept** |
|---|---|---|---|
| B9 | Trust is good, control is better (8). | Control more important than trust | Trust; control |
| B9 | The biggest risk of data sharing is not doing it. Very few companies have the data sets needed for AI topics available internally. They are therefore dependent on data cooperations (9). | The biggest risk of sharing data is not doing it | Risk |
| B9 | Market changes as an external influence. For example, if tracking with third-party cookies becomes more difficult. Or new | Market changes (new laws) as an external influence | External factor (external pressure) |

| | | | |
|---|---|---|---|
| | laws in the pharmaceutical industry regarding real-world evidence data (11). | | |
| B9 | Emotional traits, especially in companies that are not so technological. There, rational arguments for data exchange are useless. It's different with technology-driven companies, which would actually have a much higher risk, but they understand the technology and are more open (14). | Depending on their technological understanding, companies are open to data exchange in different ways | Internal factor (corporate culture) |
| B9 | PETs can enable data collaborations that were previously impossible. Where data collaboration is already taking place, it can be made more efficient. And, of course, sensitive data can be better utilized (19). | New or more efficient data collaborations with sensitive data become possible | Benefits (PETs) |
| B9 | Trust can, of course, be secured technologically. Emotional trust, however, cannot be secured (20). | PETs can generate trust | Trust (PETs) |
| B9 | Europe in particular is an exciting market because we simply have a more demanding regulatory market than America, where you can of course achieve more with PETs (22). | Europe as a very demanding regulatory environment, very interesting for PETs | External factor PETs (region) |
| B9 | If used incorrectly, PETs naturally pose risks (24). | Incorrect use of PETs leads to risks | Risk (PETs) |

| **ID** | **Interview statement/paraphrase** | **Headings/coding** | **Concept** |
|---|---|---|---|
| B10 | PETs especially useful for cybersecurity issues (8). | Cybersecurity issues with PETs | Benefits (PETs) |
| B10 | In healthcare, for instance, companies share a lot of data around effectiveness of their healthcare solutions or hospitals share a lot of KPIs with insurers, but also amongst each other to create benchmarks for instance (10). | Much data sharing in healthcare sector | External factor (sector) |

| | | | |
|---|---|---|---|
| B10 | Banks would like to share data about fraud. Here how similar banks are could be an influencing factor whether to share with other banks or not. In the Netherland banks share with some other banks from the same country, but they are more reluctant to share with external players. Also because of regulations like FISA. Similar companies (region, shared vision) helps (14). | Similarity between companies enhances data sharing | Trust |
| B10 | Different sectors have different degrees of cooperation, degrees of competition, different regulations. This is reason for different data sharing behavior (16). | Differences between sectors in data sharing behavior | External factor (sector) |
| B10 | When companies think about data sharing, it is a subjective decision making process. Even with the technology to fix privacy problems, there is still always some personal distrust. If companies collaborate, it must be right in every aspect and not just by solving one privacy issue. There is also some personal trust needed. People should like each other and if they don't like each other, it's much harder to get something going (18). | Trust is always needed, because it's a subjective decision making process | Trust |
| B10 | Biggest risks is getting penalties by privacy authorities, so regulatory risk. Then reputational risk and competitive risk (22; 24). | Regulatory risks, reputational risks, competitive risks | Risk |
| B10 | External factor is the law and company specific policies. So you have the GDPR which just says something in general about how you should do data processing and then a company would make its own policy based on the GDPR. Policy can say for example you need to do these 20 steps | Laws and company policies as external factors | External factor (external pressure) |

| | | | |
|---|---|---|---|
| | before sharing data and talk to these 50 people. So this can make it far more complex. Sometimes companies make these processes even a bit too strict because they really want to avoid those risks (28). | | |
| B10 | Culture of risk averseness. It's quite easy to say, well, there might be some risks here, so let's just not share anything (28). | Culture of risk averseness | Decision rationality |
| B10 | PETs can make data sharing possible where it was not possible before or very risky (30). | New possibilities in data sharing | Benefits (PETs) |
| B10 | One use case for PETs is to single out individuals who have certain characteristics without learning anything about all other individuals. You can use this for fraud for example (34). | New possibilities in data sharing | Benefits (PETs) |
| B10 | In the future more data will be shared due to new laws like the data act, but also very sector specific acts like the European health Data Space Act. So companies will have to open up more. And also GDPR has some clauses where it says that you should always strive for maximum privacy, given the current technological possibilities. So in the future with GDPR you will also be asked why you are not using MPC (48). | More data sharing with PETs in the future | Benefits (PETs) |
| B10 | With PETs you can put some of your trust in technology, and need less trust in people. But still you see that all MPC companies start off by trying to sell their technology and at some point they start talking more about data, collaboration and governance and those kind of terms. Also contracts are very important to create | Less trust needed with PETs, but still important | Trust (PETs) |

| | | | |
|---|---|---|---|
| | control. With smart data requests you can also with SMPC find out a lot more than intended (54). | | |
| B10 | You can manipulate inputs to infer sensitive information. By repeatedly entering slightly modified datasets and observing the changes in output, an attacker can deduce specific details about the original dataset. This technique allows an attacker to gradually compile or reconstruct sensitive data, making the system vulnerable to inference attacks despite the secure computation process (56). | New risks arising with PETs | Risk (PETs) |
| B10 | Another risk is when you use MPC it is basically a black box. So you can sort of get into sort of algorithmic decision making without understanding why the algorithm says something. And it's very complex technology. So it's hard to explain to people, management etc. (58). | New risks arising with PETs | Risk (PETs) |
| B10 | So there's a risk of people using it wrong or understanding it wrong, making the wrong assumptions. Companies also really don't like it when you are messing in their IT systems with all these new technology that they hardly understand (60). | New risks arising with PETs | Risk (PETs) |
| B10 | It's also quite easily to sort of just put in garbage data because you want to show that you're very good or you want to mess up. So it's a risk for data manipulation. Also you should always also be able to create a sort of replay of certain calculations (62). | New risks arising with PETs | Risk (PETs) |

| | | | |
|---|---|---|---|
| B10 | Especially insurances are interested to invest in PETs because they really have something to win if healthcare becomes more efficient and effective (66). | Companies willing to invest in PETs | External factor PETs (sector) |
| B10 | In the company typically people who are investing in PETs are more part of the innovation functions rather than sort of the operational part of the company (66). | People willing to invest in PETs | Internal factor PETs (corporate culture) |
| B10 | Biggest challenge for PETs is the fact that they're sort of collaboration technology. That means you don't sell your PET solution to one customer, you need to sell it to at least two at the same time. And then always one has less appetite or says yes, super interesting, but maybe in some months because now I'm full or busy or so that timing is hard (70). | PETs as collaboration technology is big challenge | Acceptance of PETs |
| B10 | Another challenge are the many angles of the technology. Tech angle, organizational angle, legal angle. So you not only have two companies on the table but also 5 departments of the company (70). | Many angles of technology as challenge | Acceptance of PETs |
| B10 | That's why top management is very important to manage this cross department cooperation (72). | Top management important factor | Internal factor PETs (top management) |

| **ID** | **Interview statement/paraphrase** | **Headings/coding** | **Concept** |
|---|---|---|---|
| B11 | Both trust and control are important for sharing data between companies (8). | Trust and control important for sharing data | Trust, control |
| B11 | Trust is a vague thing, it's hard to quantify, with PETs they only need to have trust in mathematics (8; 10). | With PETs only trust in mathematics is needed | Trust (PETs) |

| | | | |
|---|---|---|---|
| B11 | The problem currently is that people don't really trust the idea how PETs work, because they don’t understand it. They are sometimes just really confused. That’s also why we lose clients on really subjective grounds sometimes (14; 16). | No trust in idea of PETs | Trust |
| B11 | Risk in Data sharing that data can be leaked (18). | Data leakage as risk | Risk |
| B11 | Once the companies will understand that they can share certain sensitive data without significant leakage. Then I think it will just speed up collaborations between different companies. And also opens up certain data sets for companies to exchange. It will be shared more data in total (26; 28). | PETs will speed up data exchange | Benefits (PETs) |
| B11 | Trust will become less important. But there will be still the problem that results of SMPC collaboration might leak little bit more than one company wanted to leak to they're counteragent (30). | Trust less important, but still some information gets revealed with PETs | Trust (PETs) |
| B11 | Big companies with big network of suppliers would be most willing to use PETs to simplify the data sharing with their suppliers (38). | Big companies more willing to invest in PETs | Internal factor PETs (size) |
| B11 | Especially in food production sector there is a demand for sharing data. There are often many suppliers involved and companies need comparison of these. Also data about food quality is really important there (44). | Food production sector has high demand for data sharing | External factor (sector) |
| B11 | Technical issues and user friendliness are biggest challenges for acceptance of PETs on the market (50). | Technical issues and user friendliness are biggest challenges for acceptance of PETs | Acceptance of PETs |

| | | | |
|---|---|---|---|
| B11 | GDPR and removement of third party cookies were big external influences. With PETs companies can save time with legal topics (58). | Push through GDPR and third party cookies removement | External factor PETs (external pressure) |

| **ID** | **Interview statement/paraphrase** | **Headings/coding** | **Concept** |
|---|---|---|---|
| B12 | Trust is very important when sharing data (6). | Trust as important factor | Trust |
| B12 | The decision process when sharing data is both subjective and objective. Companies have a culture and some companies will be more willing to not just sharing data, but in general trying new things and some of it wouldn't be as much. The objective part is that you think about the financial side as well and how to benefit from data sharing (10). | Decision process both subjective depending on the culture of a company and objective depending on the benefits | Decision rationality |
| B12 | Reputational/regulatory risk is the biggest risk. But there is also competitive risk (14). | Risks of data sharing | Risk |
| B12 | There is no framework how to measure risks, so people overestimate or underestimate the risks of data sharing (16). | Hard to measure risks of data sharing | Risk |
| B12 | Interoperability is the way to do data sharing, it’s the tool, but it’s not the reason to do it (22). | Interoperability as the tool to share data | Interoperability |
| B12 | One reason for data exchange is cost sharing. It's basically saying, hey, I need to do something that is extremely complex. For example, the German automotive industry needs to compute the CO2 for the government and it's extremely expensive to compute CO2 by yourself. And so they do share the cost. | Cost sharing and joined innovation are reasons for data exchange | Benefits |

| | | | |
|---|---|---|---|
| | Another one is joined innovation you can really do a lot of things when you put data together because data by itself is useless (38). | | |
| B12 | Governmental influences are a very big external influence. Regulations like the data act forces companies to share data (40; 42). | Governmental influences as external influence | External factor (external pressure) |
| B12 | Data Mesh is main internal factor for not sharing their data with other companies, but only inside the company (42) | Data Mesh as internal factor | Internal factor (data governance) |
| B12 | PETs won't change the data sharing behavior in the future. We have already many technical tools which work perfect, but there are still cybercrimes. It is because usually problems come from the weakest link. So what I'm trying to say with that is that secure multiparty computation is probably a fantastic way, but does that eliminate every single human risk? The answer is no. SMPC is probably a fantastic technology, but also very complex to understand and so it is rarely applicable in practice. It's the same with blockchain (48). | No change through PETs because it takes more than just good technology to make a difference | Risk (PETs) |
| B12 | Trust is always needed. Look at self driving cars. Still people are scared because they are not used to it. Humans get used to a certain way of behavior and then changing that can be difficult (54). | Trust is always needed. | Trust (PETs) |
| B12 | You need to be ready to share data culturally and have a business use case, then you would use PETs (56). | Culture and Use case are important for accepting PETs | Acceptance of PETs |
| B12 | PETs can minimize risks, but also the amount of damage increases if something | Probability of damage due to data sharing decreases with PETs, | Risk (PETs) |

| | | | |
|---|---|---|---|
| | happens because it's all getting more sensitive (59). | but amount of damage increases | |
| B12 | The biggest challenge for acceptance of PETs on the market is the complexity. How much human resources do you need to get some benefits from it, how simple is technology to use (63). | High complexity of PETs require qualified people | Internal factor PETs (resources) |

| **ID** | **Interview statement/paraphrase** | **Headings/coding** | **Concept** |
|---|---|---|---|
| B13 | As a rule, various levels are involved in data exchange decisions. Depending on what data is involved and especially if it is sensitive data, the management level is usually always involved, often also a legal level and, of course, my business decision (6). | Several levels are usually involved in data exchange decisions | Decision rationality |
| B13 | Data Act means that companies are under real pressure to make data accessible in one form or another (12). | Data Act as leverage for companies to exchange data | External factor (external pressure) |
| B13 | Value creation is crucial for companies to share data (12). | Value creation as a factor for companies to share data | Benefits |
| B13 | Trust is very important. Companies are dependent on believing the statements of third parties. And trust simply plays a significant role here as a psychological factor (12). | Trust as an important factor | Trust |
| B13 | The most important influence when it comes to trust is to know: Others have already trusted this organization. If you know that others have had good experiences with it, trust is built up. Technical expertise, which can be proven by publications or titles, also helps (14). | Experiences of others and publications, titles build trust | Trust |

| | | | |
|---|---|---|---|
| B13 | The most important prerequisite for PETs is that there is a group of stakeholders who are interested in using the technology together, such as industry associations. The biggest challenge is to bring these many parties together (18, 30). | Group of stakeholders must be interested in technology | Acceptance of PETs |
| B13 | A high level of technical expertise is required for the implementation of PETs (20). | Technical expertise | Internal factor PETs (resources) |
| B13 | Incorrect use of PETs can pose a risk. PETs only guarantee so-called input privacy. Incorrect calculation queries can lead to sensitive information being disclosed despite the fact that security is technically guaranteed (24). | Incorrect use of PETs can pose a risk | Risk (PETs) |
| B13 | Agile, innovative companies, in particular, are dealing with PETs. The culture and attitude of companies is crucial (26, 28). | Agile, innovative companies get to grips with PETs | Internal factor PETs (corporate culture) |
| B13 | With PETs, trust in other organizations to behave correctly, to do the right thing with the data, is replaced by trust in algorithms and mathematics (32). | Trust in other organizations is losing importance | Trust (PETs) |

| **ID** | **Interview statement/paraphrase** | **Headings/coding** | **Concept** |
|---|---|---|---|
| B14 | If you share data on PETs with competitors in order to further develop your technology, you are of course also helping your competitors. But you also want to be first on the market. Companies with many data sets are reluctant to share them with others (18, 20). | Disadvantages of sharing data even with PETs | Risk (PETs) |
| B14 | Sharing data for cyber security issues also poses the risk that the sensitive data could be used to gain knowledge that the | Reputational damage due to data sharing | Risk |

| | | | |
|---|---|---|---|
| | company has not taken the right precautions (22). | | |
| B14 | PETs if then chance of success with long lead time and many positive use cases (24). | Chance of success for PETs with a long lead time and many positive use cases | Acceptance of PETs |
| B14 | PETs mainly applicable in the financial and medical sectors (26). | Financial and medical sector for PETs | External factor PETs (sector) |

| **ID** | **Interview statement/paraphrase** | **Headings/coding** | **Concept** |
|---|---|---|---|
| B15 | Sharing data with competitors is not easy, but this depends a bit on the competitiveness of the industry as well. For financial and insurance company it's a bit harder, that's why they got pushed a bit by regulations (7; 11). | Competitiveness of the industry as factor for data sharing | External factor (sector) |
| B15 | Trust is a very important factor. When I have a method to verify how my data is treated, this is very important and creates a lot of trust (17). | Trust through methods to verify data treatment | Trust |
| B15 | The decision process in a company whether to share data is both subjective and objective. Subjective because there are human in the loop. Objective a team is always deciding, everything has to be reported to the board, not only to the CEO (19). | Decision process about sharing data both subjective and objective | Decision rationality |
| B15 | Big companies less subjective. They do more risk analysis. That's why they are also sometimes a bit slower in adoption, because everything has to be tested before approval (27). | Big companies less subjective in decision process | Decision rationality |
| B15 | First risk in data sharing is when you share the data and then it contains some sensitive data, it could be like a personally identified information. Then | Regulatory risks and reputations risks when sharing data | Risk |

| | | | |
|---|---|---|---|
| | you can get a penalty. So it is a regulatory risk. Second risk when you share your data with another company and they are sloppy with their management of your data. And then you get a penality as well.Third risk is that your brand can get tainted. Especially bad for banks if their reputation is tainted through a data breach (29). | | |
| B15 | There has to be a business benefit for companies to share their data. It can be either to reduce risk, reduce cost or to increase revenue through collaboration. The used technology is just the how it is done, but whether it's done at all depends on the economic side (35). | Data sharing to reduce risk, reduce cost or to increase revenue | Benefits |
| B15 | One of the strongest external influences is regulation. If the law and regulations say you have to share data, then they will share (37). | Regulation as external factor | External factor (external pressure) |
| B15 | In the future there will be a need for industry and society to collaborate more. The value of data depends on the combination with other data. So PETs will become more important (43). | Combination of data increases value | Benefits (PETs) |
| B15 | Trust is still important with PETs, especially because it's a new technology (45). | PETs as new technology require trust | Trust (PETs) |
| B15 | Profitability very important. If there is an economic reason, you are willing to adopt the technology (49). | Economic reason for adopting PETs | Acceptance of PETs |
| B15 | Regulatory pressure is another reason. You need a technology to collaborate and preserve privacy (49). | Regulatory pressure as reason to adopt PETs | External factor PETs (external pressure) |
| B15 | For smaller companies a bit easier, if they have the money for it. They are | Smaller companies more willing to use PETs | Internal factor PETs (size) |

| | | | |
|---|---|---|---|
| | quicker making decisions and you only have to convince one or two stakeholder and not 20 (51). | | |
| B15 | PETs reduce risks of data sharing, but there are also some new risks arising like different types of attacks, for example data poisoning (53). | Risks are reduced with usage of PETs, some new risks arising | Risks (PETs) |

| **ID** | **Interview statement/paraphrase** | **Headings/coding** | **Concept** |
|---|---|---|---|
| B16 | In the healthcare sector organizations share their data already with NDA's and those things, but they look for other things like PETs (4). | In the healthcare sector much data is shared | External factor PETs (sector) |
| B16 | Decision process is both subjective and objective. A personal connection is always helpful, but there is also always an objective factor like complying to the law (9) | Decision process both objective and subjective | Decision rationality |
| B16 | When you send data out, it's gone and you don't know what is happening with the data. Even if the other party doesn't have any malicious intent, they can be hacked. So you are just not in control of your data anymore. This is one big risk (11). | Control loss when you share data is a big risk | Risk |
| B16 | People are getting older. More Healthcare is needed. This stimulates data sharing to optimize the costs (4, 13) | Demographic change as an external influence | External factor (situational opportunity) |
| B16 | It's a natural thing that more data will be shared in the future, but it's maybe not due to MPC. I think the MPC part is mainly doing it in the privacy preserving way (15). | Data sharing with PETs in a privacy preserving way | Benefits (PETs) |

| | | | |
|---|---|---|---|
| B16 | If you don't use MPC in the right way, you can still leak data. Some questions can disclose sensitive information. So trust is still needed (17). | Trust that other party doesn't leak data | Trust (PETs) |
| B16 | Companies which invest in PETs: MPC is still kind of new, so it's definitely the ones that are more innovative. So then they must have money to innovate and also they are already a bit mature with data (20). | Innovative, financial strong and data affine companies invest in PETs | Internal factor PETs (resources, corporate culture) |
| B16 | Also people very important when using PETs (22). | Human resources as factor | Internal factor PETs (resources) |
| B16 | PETs are like magic. People don't really understand it. That holds them back (26). | Lack of understanding about PETs | Acceptance of PETs |

| **ID** | **Interview statement/paraphrase** | **Headings/coding** | **Concept** |
|---|---|---|---|
| B17 | When it comes to trust, the crucial point is trust that interests will be safeguarded. It is about drafting contracts in such a way that the data remains secure even when it is passed on. Not trust on a personal level. Trust on the basis of an intensive risk assessment. (8, 23). | Trust that interests will be protected is an important factor in data exchange | Trust |
| B17 | Objective decision-making process when sharing data. Fixed parameters are used to decide whether data should be shared or not. Parameters: Is he a competitor? What data is involved? What needs to be cleansed? In which business area is the company active? The decision-making process always depends on the use case (10). | Objective decision-making process when sharing data based on defined parameters | Decision rationality |
| B17 | When sharing data, there is the so-called value risk dilemma. This means that the value of the data can only be quantified in very abstract terms, but the risk is very | Value risk Dilemma has an inhibiting effect on data exchange | Benefits, Risk |

| | | | |
|---|---|---|---|
| | concrete. Only at the end can you say how much added value the data has brought. However, the risks are very high. This makes companies reluctant to share data (10,12). | | |
| B17 | Regulations such as special use permits for micromobility providers if they share their data in return (18). | Micromobility providers share their data and receive approval in return | External factor (external pressure) |
| B17 | PETs can be used primarily within industry associations (20). | PETs can be used primarily within industry associations | Acceptance of PETs |
| B17 | Risk due to new link in the chain, which can potentially be leaked. Who is liable if something happens? How secure is the technology really? (25). | Risks of PETs | Risk (PETs) |
| B17 | For PETs to be accepted, early adopters must be convinced until a critical mass is reached (27). | Critical mass must be convinced to accept PETs | Acceptance of PETs |
| B17 | It is also important that the results calculated with PETs are also comprehensible. It must always be possible to explain how a result was arrived at, otherwise problems arise. Especially when large sums of money are involved, it is not enough just to present a result, you also want to know how the algorithm arrived at this result (27,29). | Traceability of the calculated results of PETs is an important point | Acceptance of PETs |

| **ID** | **Interview statement/paraphrase** | **Headings/coding** | **Concept** |
|---|---|---|---|
| B18 | The human side, the relationships, they matter a lot. If they have done previous business with each other, it helps (5). | The human side, the relationships, previous business together matters a lot | Trust |

| | | | |
|---|---|---|---|
| B18 | Both subjective and objective decision process. One needs to have the courage and the progressive mindset to do it. Then this person needs to go through the pipeline of all the compliance checks before it can be set up. So you need both (7). | Combination of progressive mindset and compliance topics | Decision rationality |
| B18 | Privacy penalties and competitive economic Disadvantage as biggest risks of data sharing (9). | Privacy penalties and competitive economic Disadvantage as biggest risks of data sharing | Risk |
| B18 | Use cases for PETs are mainly healthcare, advertising and security and defense topics (15, 17). | healthcare, advertising and security and defense as Use cases for PETs | Benefits (PETs) |
| B18 | With PETs more sensitive data is shared (19). | With PETs more sensitive data is shared | Benefits (PETs) |
| B18 | It helps to trust the technology if the technology is certified and the company is audited. And as soon as you can trust the technology and the technology enforces a lot of control, you need to trust less (21). | If you trust the technology, you need less trust | Trust (PETs) |
| B18 | In Netherlands people are more willing to use innovative technology to set up a data collaboration, whereas in Germany it's more about we need to protect and protection is critical. And collaboration is only an option. So I think the cultures per market differ a lot (25). | Cultural differences | External factor (region) |
| B18 | Communication is the biggest challenge on the market for PETs as the technology is so difficult to explain (29). | Communication is the biggest challenge on the market for PETs | Acceptance of PETs |

| **ID** | **Interview statement/paraphrase** | **Headings/coding** | **Concept** |
|---|---|---|---|

| | | | |
|---|---|---|---|
| B19 | SMPC mostly used for cross-company comparison of KPIs (6). | SMPC mostly used for cross-company<br>cross-company comparison of KPIs | Benefits (PETs) |
| B19 | There are also many challenges in cross-country data exchange. In the financial sector, for example, you have customer data in the EU and customer data in China. With statistical evaluations, there are often data protection problems if you want to exchange data across national borders (6). | Data protection issues with cross-country data exchange | Benefits (PETs) |
| B19 | Catena-X data exchange is based almost exclusively on trust. Because there is no technical enforcement. There are only contracts. Nobody can be prevented from copying data and, in principle, doing anything with it (8). | Data exchange at Catena-X is mainly based on trust | Trust |
| B19 | It's a subjective decision-making process. So I don't think there are objective criteria between those who are really classified. We share data with others under such and such circumstances. I think it's decided on an individual basis and I think there's a lot of subjectivity involved (10). | Subjective decision-making process without objective criteria | Decision rationality |
| B19 | Risks are disclosure of company trade secrets. Violation of data protection regulations. Otherwise always use case dependent (12). | Risks are disclosure of company trade secrets.<br>Violation of data protection regulations. | Risks |
| B19 | There are many non-digitalized processes between SMEs and | Digital infrastructure as a prerequisite for data exchange | Digital infrastructure |

| | | | |
|---|---|---|---|
| | automotive companies. First of all, there are completely different challenges. In other words, ensuring that data is in a standardized format, that you have a sensible interface and that you can retrieve data in real time. These are fundamental requirements (14). | | |
| B19 | The challenge is that you always have several stakeholders. Agreements are difficult. A lot of coordination required when exchanging data (16). | Stakeholder agreements, need for coordination as challenges | Acceptance of PETs |
| B19 | We didn't have a case where we approached a department from the area of data exchange with suppliers or other companies that even knew what multi-party communication was, because it's not a topic that comes up in most computer science courses. That is also a completely different challenge (18). | Lack of knowledge about SMPC as a challenge | Internal factor PETs (resources) |
| B19 | Internal influence: There are many people in departments who simply write off use cases because they don't even know that there is a technology for them (22). | Lack of knowledge about SMPC as a challenge | Internal factor PETs (resources) |
| B19 | Internal influence: Often rejection/mistrust of new technologies. In addition, there is often a lack of time to familiarize oneself with such new technologies and to understand | SMPC ultimately a question of trust due to a lack of time and technical capacity | Trust (PETs), internal factor PETs (resources) |

| | | | |
|---|---|---|---|
| | the topic. Therefore, the use of SMPC often ends up being a question of trust because there is a lack of time and technical capacity (22). | | |
| B19 | Internal factor: Large companies in particular have many issues to deal with. If use cases are implemented with PETs, you would have to familiarize yourself with the topic and it would be a considerable effort. More convenient to use old technologies (24). | Effort of familiarization with PETs | Internal factor PETs (size) |
| B19 | Internal factor: It is often easier for large companies to use familiar means such as contracts for data security. Departments do not necessarily share the interest in securing data technologically from a corporate perspective. As long as they have done everything right and the data is contractually secured, everything is fine for them (24). | Known methods are easier to implement than PETs | Internal factor PETs (size) |
| B19 | Mainly legal regulations as an external influence, especially for cross-border use cases. But this is also a challenge with PETs, as it is not yet 100% clear whether this technology really processes data in compliance with data protection regulations (24). | legal provisions as an external influence that cannot yet be resolved with certainty, even by PETs | External factor PETs (external pressure) |
| B19 | The classic use case is aggregation of CO2 emissions | Aggregation of CO2 emissions along the supply chain as a use case for PETs | Benefits (PETs) |

| | | | |
|---|---|---|---|
| | along the supply chain. Now also required by law (26). | | |
| B19 | Use case for PETs if a company wants to exchange data internally across countries. There is a concrete need there (26). | Cross-national intra-company data exchange as a use case for PETs | Benefits (PETs) |
| B19 | When it comes to the conditions that must be met for PETS, the technical aspect is secondary. It is more important that people get involved with the technology and that you manage to convince several stakeholders (28). | Technical conditions secondary, conviction of people from different parties more important | Acceptance of PETs |
| B19 | PETs have the potential to change data exchange. However, this would require pioneers who use the technology and clear rules from legislators (30). | Pioneers who use technology and have clear legal requirements | Acceptance of PETs |
| B19 | The larger and more digitalized companies are, the more likely they are to invest in PETs (34). | Size, degree of digitalization important for likelihood to invest in PETs | Internal factor PETs (size, corporate culture) |
| B19 | Particularly in the medical sector, areas of application for PETs such as joint statistical analyses (34). | Areas of application for PETs in the medical sector | External factor PETs (sector) |
| B19 | No new risks are added with SMPC. You just have to think about how to create output privacy. This could be solved with differential privacy, for example (36). | With SMPC only risk of output privacy | Risk (PETs) |
| B19 | The biggest challenge is awareness. So far, too few people are familiar with the topic (38). | Awareness as a challenge | Acceptance of PETs |

| **ID** | **Interview statement/paraphrase** | **Headings/coding** | **Concept** |
|---|---|---|---|

| | | | |
|---|---|---|---|
| B20 | Companies look more into the business failure business benefit of the data sharing than in trust. Trust and control options are nice to have, but not a main factor (6). | Business failure business benefits of data sharing are more important than trust and control options | Trust, control, benefits |
| B20 | You could also say companies don't trust each other. And then control options are there to make sure that with the conditions of no trust, business can still be run as usual (7). | Control replaces trust between companies | Control, trust |
| B20 | More objective decision process driven by benefits and costs of data sharing (9). | Objective decision process driven by benefits and costs | Decision rationality |
| B20 | Losing competitive advantage, misuse of consumer data and harm of reputation are risks of data sharing (11). | Risks of data sharing | Risk |

## Appendix D Interview Transcripts

1 Interview B1

2

| | |
|---|---|
| Interview-Nr. | 1 |
| Date of the interview | September 02, 2024 |
| Duration of the interview | 32:47 min |
| Interviewer | Felix Starnecker (I) |
| Interviewee | B1 (Germany) |
| Role | Consultant data governance |
| Sector | Automotive sector |
| Specialities | Interview in german |

3 [0:00:00] **I:** Und ok, dann wär meine erste Frage. Mit wem tauschen Unternehmen Daten aus? Bei dir im Unternehmen oder wie du es auch von anderen Unternehmen kennst? Und mit welchen Leuten und wofür genau?

4 [0:00:31] **B1:** Genau also ich würd jetzt mal aus einer generellen Unternehmenssicht sprechen, ohne jetzt einzelne Unternehmen einzugehen. Generell ist es natürlich so, dass man, wenn man es ganzheitlich betrachtet, dann macht man das natürlich mit Lieferanten, Kunden, wie du es da schon geschrieben hast, indem man halt einfach Vertragsbeziehungen zu verschiedenen Personen hat, sei es, (..) wenn jetzt ein Kunde irgendwas kauft oder wenn man vom Lieferanten natürlich irgendwelche (..) ja ich nenne jetzt mal Produktionsbestandteile wieder einkauft, die man für seine eigene Produktion benötigt. Da hat man dann in der Regel ja Vertragsbeziehungen (..) da werden dann ja Daten ausgetauscht über "was benötige ich", "wieviel benötige ich davon"? Und auf der anderen Seite, wenn du so an Systeme denkst, die beispielsweise von Unternehmen eingesetzt werden, dann hat man da natürlich auch bei ich sage es mal bei SARS Lösungen zum Beispiel, wenn ich Daten in die Cloud ziehe, für mein Personalsystem zum Beispiel, da hat man da natürlich auch ne Vertragsbeziehung und da natürlich auch ne Auftragsverarbeitungsvereinbarung (..) jetzt aus dem datenschutzrechtlichen Sinne, dass man sagt, man gibt Daten rüber für einen bestimmten Zweck, zum Beispiel für ja Personalcontrolling sag ich jetzt mal, dass man das auf einer Cloud hat und da hat man dann ja Daten. In der Sicht, dass man als Unternehmen eine Cloud Lösung zum Beispiel einsetzt, um dieses nutzen zu können und dann die eigenen Daten wieder zu haben. Der Anbieter kann die Daten da in der Regel dann nicht nutzen, sondern ist nur der Anbieter entsprechend eines Services. Und darf die dann auch gar nicht nutzen. Aber das wäre natürlich auch ein Datenaustausch in der Hinsicht, dass man auf eine externe Plattform beispielsweise geht.

5 [0:02:13] **I:** OK. Mhm und da gibt es dann auch immer Verträge, die das Regeln oder ist das einfach auf Vertrauensbasis dann?

6 [0:02:25] **B1:** Nee, also in der Regel gibt es ja Verträge, ich sag jetzt mal wenn du n Auto kaufst hast du ja n Vertrag, dass du deine Daten dann da zum Beispiel angibst wie Kontonummer, Name, Vorname, Adresse und sowas (..) dann bekommst du ja auch zum Beispiel vom Hersteller irgendwann zurück was ist denn deine Pin Nummer beispielsweise von deinem Auto? Wie weit ist das Auto im Produktionsprozess? Zum Beispiel vom Lieferanten bekommst du ja Infos mit oder schickst du Infos. Wohin soll der Lieferant die Sachen liefern? Was benötigst du zur Produktnummer? Produktanzahl und sowas und bei Systemdaten sind es natürlich solche Sachen, wie wenn ich jetzt zum Beispiel eine SAP Cloud gehe. Alles, was ich damit verbinden möchte, ist Lieferantenmanagement abdecken möchte ich mein Personalcontrolling oder meine Personaladministration machen? Das sind dann natürlich immer unterschiedliche Daten, die versendet werden, aber das passiert ja alles auf irgendeinem Vertrag und bei manchen müssen noch zusätzliche Verträge vereinbart werden. Nicht nur ich habe einen Kaufvertrag zum Beispiel, sondern ich habe auch einen im datenschutzrechtlichen Sinne (..) einen Vertrag zur Datenverarbeitung also, dass zum Beispiel die Landesbehörde der Landesdatenschutzbeauftragte falls der mal irgendwann das überprüfen sollte dann natürlich sagen kann ja zu welchem Zweck, auf welcher Rechtsgrundlage verarbeitet ihr denn die Daten zum Beispiel - ja zur Erfüllung des Arbeitsvertrages. Hast du mir deine Bankverbindung zum Beispiel gegeben, dass du monatlich dein Gehalt bekommst?

7 [0:03:50] **I:** Mhm OK und also in meiner Masterarbeit hab ich eine strukturierte Literaturrecherche gemacht und zwei Hauptpunkte, die rausgekommen sind, sind auch Vertrauen und Kontrolle als Faktoren. Würdest du die auch beide als wichtige Faktoren sehen oder würdest du einen als wichtiger ansehen?

8 [0:04:13] **B1:** Ja, also meinst du mit Vertrauen, das Vertrauen zwischen den Parteien? Und mit Kontrolle, dass eine Partei die andere kontrolliert oder die Kontrolle über die Daten selber?

9 [0:04:19] **I:** Ja, genau mit Kontrolle meine ich eher so Verträge und Sachen in die Richtung und mit Vertrauen eher so ja weiß nicht, dass jetzt vielleicht deutsche Unternehmen lieber an andere deutsche Unternehmen ihre Daten weitergeben, weil sie halt da irgendwie ne bessere Beziehung zu denen haben oder sowas in die Richtung ja.

10 [0:04:22] **B1**: Mhm ja, ich würd da schon wahrscheinlich eher den Faktor Kontrolle als ein oder als den wichtigeren Faktor sehen, weil einfach zum Beispiel über Verträge das wirklich definiert ist und ja ausverhandelt ist, zum Beispiel wo die Daten gespeichert werden, ob das in den USA ist, ob das in Deutschland ist, da gibt es ja dieses ganze rechtliche Regelwerk, was darum gebaut ist. Wo darf ich

zum Beispiel deutsche Daten hinschicken? Dürfen die in die USA transferiert werden? Aber ich würde da schon eher sagen, dass diese Verträge da das stärkere Mittel sind, anstatt das Vertrauen. Klar, das Vertrauen muss man auch in das Produkt haben, man lässt sich ja auch alles zeigen und so in der Regel wenn man jetzt zum Beispiel eine neue Cloud Lösung einkauft aber ich würde schon eher sagen, dass Kontrolle da ausschlaggebender ist.

11 [0:05:20] **I:** Würdest du dann eher das Vertrauen gar nicht mit einbeziehen? Oder ist Vertrauen schon relevant?

12 [0:05:44] **B1:** Ja. Ich würde es vielleicht in dem Sinne differenzieren (..) differenzieren, indem man zum Beispiel sagt, dass man schon ne langjährige Kundenbeziehung mit nem Anbieter hat, (..) dass man sagt ja, man arbeitet jetzt sehr, sehr lange schon zum Beispiel mit SAP zusammen und da ist dann ein entsprechendes Vertrauen da (..) dass man weiß ja, es wird sich um Probleme kümmern, man hat seine Ansprechpartner, das könnte man schon sagen. Auf der anderen Seite kann man natürlich aber auch sagen ja gut, alles, was ich jetzt so im 1 zu 1 Gespräch irgendwie verhandele klar, da ist ein Vertrauen, in dem ich mit den Leuten spreche, aber wirklich, um das Vertrauen zu haben gehe ich doch lieber auf diese kontrollschiene und sage ich mach jetzt nen Vertrag genau.

13 [0:06:21] **I:** Ja ja genau. OK und würdest du den Entscheidungsprozess, wenn jetzt so ein Unternehmen vor der Entscheidung steht? Tauschen Sie Daten mit irgendeinem anderen Unternehmen aus? Würdest du sagen das ist ne objektiver Prozess? Wo lange Risikoanalysen und so weiter gemacht werden oder würdest du sagen, dass das subjektiv ein Top Manager entscheidet, ob er, ob das strategisch jetzt gut ist?
Und dann wird das gemacht, oder?

14 [0:06:52] **B1:** Den kann man glaub ich auch wieder auf auf beide Arten betrachten wenn du jetzt sagst n Topmanager hat jetzt sich irgendeine Idee in den Kopf gesetzt und möchte das mit dem und dem Unternehmen oder sowas Daten ausgetauscht werden, weil die sich kennen, zum Beispiel blöd gesagt muss natürlich alles immer die ganzen Kartellrechtsregelungen und sowas im Kopf haben aber es gibt natürlich Szenarien, wo dann zum Beispiel ein Manager sagt ja, ich hab mir das in den Kopf gesetzt. Ich möchte, dass er jetzt es durchzieht und schaut, wie ihr das umsetzen könnt, das durch diese subjektive Entscheidungsprozess auf der anderen Seite würde ich aber eher sagen, dass es eher ein ein objektiver Entscheidungsprozess ist, dass man in der Regel wirklich schaut ja was bringt es uns?

15 [0:07:22] **I:** Mhm danke. Ja.

16 [0:07:33] **B1:** Was sind die Kosten nutzen von diesen ganzen Punkten?
Und ja, welche Risiken stecken dann auch dahinter, indem man wirklich analysiert, ja was dürfen

wir, was dürfen wir nicht? Und da würde ich auch schon sagen ist man im Unternehmensumfeld in der Regel immer ja risikoscheu, indem man sagt ja, ich Reiz jetzt die Grenzen nicht komplett aus, sondern geh lieber n bisschen drunter um da rechtlich auf der sicheren Seite einfach zu sein.

17 [0:07:59] **I:** Und? Risikoscheuer, als es eigentlich vielleicht nötig wär. Also würdest du sagen bisschen übertrieben, dann sogar?

18 [0:08:09] **B1:** Ja, ich würd schon sagen, dass man da zurückhaltender in der Regel ist und Daten eher ungerne, dann auf der einen Seite preisgibt, als auch dann natürlich schaut ja das Risiko zu vermeiden im Sinne von ich Reiz es jetzt nicht voll aus, wie weit ich rechtlich gehen darf, sondern bleibt vielleicht sogar ein bisschen darunter. Ich kenn das zum Beispiel von amerikanischen Unternehmen, die ja per se erstmal in der Regel eigentlich alles machen dürfen, sag ich mal so salopp und dann eher sagen, bis sie eingebremst eingebremst werden und ich glaube, deutsche Unternehmen, da ist das genau auf der anderen Seite, dass man eher ein bisschen risikoscheuer ist und sagt ja, ich Reiz das jetzt nicht komplett aus.

19 [0:08:46] **I:** Aber du glaubst, dass das rechtliche das größte Risiko ist, was gesehen wird, so dass man irgendwas aus Versehen vielleicht da falsch macht und dann verklagt wird oder sowas in die Richtung?

20 [0:08:57] **B1:** Ja, das auf der einen Seite auf jeden Fall, dass man rechtlich, wenn da zum Beispiel Strafzahlungen oder Reputationsschäden natürlich hinter stehen, wenn man jetzt n Datenleck oder sowas hat, dann steht das gleich große Ich sag jetzt mal in der Bild Zeitung in der in der FAZ oder sowas das ist natürlich n riesen Reputationsschaden, der auch Geld oder Umsatz kosten kann, dann entsprechend als auch dann natürlich Strafzahlungen. Wir jetzt an die Datenschutzgrundverordnung denken, dass da entsprechend bei personenbezogenen Daten, wenn die geleakt werden und ich hab als Unternehmen natürlich irgendwie Mist gebaut, bei der Implementierung oder so, dass da dann natürlich Strafzahlungen entstehen, das ist natürlich schon nicht zu vernachlässigender Faktor.

21 [0:09:32] **I:** OK. Mhm, das hab ich jetzt eh schon, das wär eigentlich eh meine nächste Frage gewesen, welche Risiken du siehst beim Datenaustausch, das ist dann genau was du, wenn ich gerade gesagt hast oder also rechtliches Risiko und.

22 [0:09:41] **B1:** Mhm. Genau rechtliches Risiko und da, wo vielleicht noch wo viele oder ne ne hohe Investitionssumme hintersteht, wo Betriebsgeheimnisse hinter stecken, da ist man dann natürlich auch sehr, sehr vorsichtig erstmal in welche Systeme gehen diese Daten rein? Wer hat Zugang zu diesen Daten oder Zugriff zu diesen Daten? Und die teile ich dann natürlich auch nicht, weil ich mir dadurch theoretisch irgendeinen Wettbewerbsvorteil, ja zunichte machen kann, indem ich die irgendwie preisgebe.

23 [0:10:11] **I:** Ja. Gibt es irgendwelche Voraussetzungen, damit ein Datenaustausch erfolgreich läuft zwischen Unternehmen? Oder?

24 [0:10:28] **B1:**Mhm, Mhm ja also ich würd schon sagen, dass man natürlich auf beiden Seiten dann ne hohe Datenqualität haben muss. Also die Daten, die ich teile, die sollten ne hohe Qualität haben, damit beide Seiten damit arbeiten können also. Entsprechende Format zum Beispiel die Aktualität und. Auf der anderen Seite auch so ne Interoperabilität, dass man Daten dann halt auch nutzen kann gleich und ja, die Bereitschaft entsprechend zum Teilen, also wenn man nicht bereit ist, die Daten zu teilen oder wenn da irgendwelche Restriktionen hinter stecken. Ich denk jetzt auch noch mal an das an den Bereich China die chinesische Regierung versagt ja zum Beispiel den Transfer von bestimmten Daten außerhalb ihres Hoheitsgebiets und von daher, das ist auch so ne Voraussetzung also rechtlich als auch Unternehmensintern dazu sagen ja, man hat jetzt die entsprechenden Bedingungen und Voraussetzungen geschaffen wie Interoperabilität, Interoperabilität, Datenqualität und die Bereitschaft zum Teilen.

25 [0:11:35] **I:** Okay, okay und ja, das wäre jetzt auch schon irgendwelche externen oder internen Einflüsse, die Unternehmen dazu beeinflussen, Daten nicht zu teilen, wäre dann auch wie Du meintest, zum Beispiel Regierungen in China oder?

26 [0:11:49] **B1:** Zum Beispiel das wäre jetzt so ein externer Einfluss, wenn ja, da Systeme zum Beispiel so macht es glaube ich die wenn ich das richtig im Kopf habe. Die chinesische Regierung, die bewertet dann welche Daten sind zum Beispiel im System drin und welche dürfen davon geteilt werden? Also außerhalb des Landes fließen? Dann kriegst du eine Liste zum Beispiel wo? Dann steht ja 320 Felder wurden approved 5 Felder wurden jetzt nicht approved und das wären natürlich externe, regulatorische Einflüsse und interne dann natürlich auch. Also wenn ich jetzt so aus der Unternehmenssicht wieder schaue, dann gibt es in der Regel natürlich jemanden, der verantwortlich ist für die Daten das Unternehmen einerseits natürlich selbst, aber meistens auch noch bestimmte Mitarbeiter, sei es Führungskräfte oder andere und da muss natürlich dann auch entsprechend ja die Bereitschaft da sein. Datenqualität Hochzutreiben und dann natürlich diese Daten ja zu teilen, natürlich auch intern zu teilen. Das ist natürlich auch schon mal manchmal ein Problem, Daten intern zu teilen und danach extern zu geben, ist dann natürlich noch mal ein ganz andere Hausnummer.

27 [0:12:57] **I:** Und würdest du dann eher sagen, dass das abhängig ist? Von jetzt zum Beispiel Leuten wie dir du bist ja in der Data Governance oder also Leute, die das übergeordnet Regeln oder würdest du sagen, dass da auch alle Mitarbeiter dann selber mitziehen müssen, damit sowas funktioniert?

28 [0:13:15] **B1:** Ich glaub, da braucht es auf beiden Seiten, die das Test Commitment einfach auf der einen Seite, dass es Richtlinien gibt, die so einen ja datengetriebenes denken und n Datenaustausch ermöglichen, dass man halt wie viele Unternehmen das jetzt auch machen zu einer Datengetrieben, um dann zu einem datengetriebenen Unternehmen wird. Wir kennen es von Google, die werden, glaube ich, alles aus, was wir, was wir in die Suchmaschine eintippen, also hinsichtlich der Daten getrieben datengetriebenen Unternehmen und auf der anderen Seite natürlich dann dem Mitarbeiter selbst, die dann ja dafür verantwortlich sind, natürlich auch Daten oder ihr wissen dann natürlich auch in die Systeme zum Beispiel rein zu packen und diese Daten, die dann in den Systemen sind, auch entsprechend über Schnittstellen natürlich anzubinden und da wo notwendig, dann auch zu teilen zum Beispiel, dass man ich sag. Analysen fahren kann für für Advanced Analytics oder sowas und wenn man das intern, glaube ich etabliert hat, dann ist der Schritt nicht mehr so weit oder nicht mehr so groß, dass man dann auch theoretisch extern natürlich auch je nach Datenlage oder nach Use Case dann mal abklopft. Mit welchen Unternehmen kann man denn sprechen? Wie sieht das denn bei den anderen aus?

29 [0:14:32] **I:** OK, alles klar, jetzt würden wir zum zweiten Teil kommen, und zwar dann jetzt mit einbezogen mit Secure Multiparty Computation. Nachdem wie du es jetzt verstanden hast, wo würdest du da Anwendungsfälle für die Technologie? Sehen kannst du dir überhaupt vorstellen, dass du was gebraucht werden könnte in Unternehmen oder glaubst du, dass es eher ja vielleicht nur ne gute Technologie, aber praktisch einfach nicht anwendbar?

30 [0:15:06] **B1:** Mhm also ich find die Technologie auf jeden Fall interessant ich glaub, da gibt es auch anwend Anwendungsfälle auf der einen Seite muss natürlich die Voraussetzung dafür im Unternehmen geschafft werden und welche Anwendungsfälle da sinnvoll sein können, ist natürlich generell würde ich erstmal sagen da wo man ja Daten irgendwie gemeinsam analysieren möchte vielleicht ich sag jetzt mal in einem Joint Venture oder sowas wo jetzt 5050 2 Unternehmen dieses dieses Unternehmen an dem Unternehmen beteiligt sind und man da vielleicht einfach sagt ja gut, wir teilen jetzt unsere Daten möchten aber nicht, dass die Daten, die wir aus dem Einen Unternehmen haben 100% auch an das andere an den Wettbewerber zum Beispiel fließen. Aber wir haben gemeinsam den Joint Venture und haben da dann eine gemeinsame Datenanalyse bei diesem Joint Venture eben, weil es uns 50 zu 50 teilen entsprechend gehört, dass man da dann entsprechend was macht was macht? Ich glaub in der in der Wissenschaft generell ist es schon. Einfacher zum Beispiel zu sagen, dass ich als Universität oder als Wissenschaftler da entsprechend Daten von Unternehmen bekomme, die entsprechend verschlüsselt sind, und ich kann damit irgendwelche Auswertungen machen sei es für die medizinische Forschung oder sowas in die Richtung jetzt gedacht also überall, wo hohe Investitionskosten, glaub ich hinter stecken mit einem hohen Aufwand

an Rechenkapazität, vielleicht auch, weil ich kann mir vorstellen, dass die nicht gerade rechenarm sind. Diese Verfahren und das wären so Anwendungsfälle, die ich jetzt gerade mal so spontan im Kopf hätte.

31 [0:16:46] **I:** Ja ja. Und? Würdest du es bei deinem in deinem Unternehmen sehen oder glaubst du da könnte irgendwas nicht passen? Oder wieso es nicht?

32 [0:17:03] **B1:** Ich also, ich bin jetzt nicht nicht so nah an der Entwicklung dran, deshalb ist es da immer für mich schwierig zu beantworten, ich komme ja mehr aus dem HR Bereich, da müsste ich jetzt mal kurz überlegen ja.

33 [0:17:08] **I:** Ja ja. Würdest du sagen, würdest du sagen? In dem Unternehmen sucht man Möglichkeiten, um Daten vielleicht noch sicherer zu teilen. Oder ist es gerade eh keine kein Bedarf wirklich, weil es auch so funktioniert, wie sie eben gerade funktioniert?

34 [0:17:30] **B1:** Also ich glaub, wir sind auf jeden Fall auf einem sehr, sehr sicheren Niveau, was das Thema Informationssicherheit Cybersecurity und so angeht. Die Daten, die zwischen Systemen zum Beispiel geteilt werden. Das läuft dann natürlich über irgendwelche Schnittstellen und sowas und da darf dann natürlich auch die Daten, die angefragt werden oder die verteilt werden, die sind dann natürlich auch verifiziert und auf der anderen Seite genehmigt, dass die Daten ausgetauscht werden können. Im Unternehmen ist es da natürlich ein bisschen angenehmer, würde ich sagen, aber sobald es halt nach extern geht, ist es da schwieriger und ich glaube, da sind generell einfach von der Mentalität her wieder. Die deutschen Unternehmen also wenn ich jetzt an uns zum Beispiel denke etwas zurückhaltender was. Angeht Daten jetzt einfach so mal extern zu teilen mit dieser Privacy Enhancing Technologie könnte ich mir schon durchaus vorstellen, dass man da eher bereit wäre zu sagen ja, es könnte keine Rückschlüsse auf die Daten gezogen werden. Wir können das jetzt auch mal in in irgendeinen Kreis reingehen und damit dann entsprechend Benchmarks machen zum Beispiel. Ist natürlich alles wieder, ich glaub damit spiel ich jetzt wieder auf der ich OK ich wart erstmal deine zweite Frage ab weil da ist mir gerade noch aus dem Kopf gekommen.

35 [0:18:39] **I:** Alles. Nicht wohl ehrlich, also würdest du jetzt sagen deutsche Unternehmen? Würden die Technologie eher anwenden oder eher nicht anwenden, weil sie da allgemeine gegenüber Daten und so bisschen zurückhaltender sind? Und?

36 [0:19:03] **B1:** Ich glaub erst mal ist man da zurückhaltend, weil man schauen möchte was bringt mir das Tool? Denn was ist es halt, was hast du für N für n nutzen für mich als auch welche Kosten und welcher Aufwand steckt dahinter? Aber ich denke schon, dass das Vorteil bringen kann, vor allem halt, wenn du große Mengen an Daten irgendwie analysieren möchtest und die Benchmarken möchtest zum Beispiel, dass du da dann sagst, ich geb so meine Daten nicht preis kann aber

trotzdem den Benchmark ziehen und hab dafür für meine interne Entwicklung den Vorteil, dass ich halt diesen Benchmark habe und zum Beispiel wie du es eben gezeigt hattest mit dem mit dem Durchschnitt zum Beispiel bin ich über unterm Durchschnitt und das kann ich so tatsächlich dann analysieren. Ich glaube, wenn man den Vorteil erkannt hat, dann gibt es da definitiv Anwendungsfälle, aber man ist am Anfang erstmal zurückhaltend, bis man da ja Use Cases gesehen hat.

37 [0:19:55] **I:** Ja gut ja, das ist halt sicher nen Nutzen irgendwie hat dann quasi ja Mhm und ich hoffe, ich hab gerade nämlich die Mitteilung bekommen, dass in 5 Minuten. Des Teamsmeeting vorbei ist. Ich bin mir nicht sicher, ob das dann abbricht automatisch oder nicht also OK gut.

38 [0:20:12] **B1:** Eigentlich soll es weiterlaufen, bis wir auflegen.

39 [0:20:17] **I:** Dann. Wie glaubst du, würde Secure Multiparty Computation des Datenaustauschverhalten von Unternehmen verändern oder glaubst du, es wird verändert werden?

40 [0:20:33] **B1:** Ja also ich, ich könnt mir vorstellen, dass man bereitwilliger ist. Daten auszutauschen, wenn man die Sicherheit hat, dass die Daten wirklich verschlüsselt, nicht auslesbar von jemandem dritten, sondern nur von mir selbst. Ja gelesen werden können oder verändert werden können ich glaub dann wär man da auf jeden Fall offener was sowas angeht und und sagt ja, man kann natürlich unter den rechtlichen Bedingungen, wie zum Beispiel im Kartellrecht den Wettbewerbsrecht, dass man da nicht unabsichtlicherweise vielleicht wenn man die Daten mit dem Wettbewerber teilt in irgendein Kartell reingerät oder sowas also Preisabsprachen oder sowas das ist natürlich teufelsküche dann, wenn wir sowas machen würden, aber wenn man da natürlich unter den gegebenen Umständen und den verschiedenen Use Case die man dann betrachten müsste, kann ich mir schon vorstellen, dass das Datenaustauschverhalten höher. Würde im Vergleich zu jetzt auch mit zum Beispiel mit mit Anbietern oder sowas klar hat man jetzt aktuell auch schon Verträge, aber ich glaub da gibt es dann noch ganz andere Möglichkeiten, wenn man sagt, man überträgt nicht mehr unverschlüsselt, sondern hat wirklich einfach irgendwie ein Verfahren, einen Schlüssel dahinter und dann wird das irgendwie ja gemixt.

41 [0:21:52] **I:** Und auf die Faktoren vertrauen und Kontrolle, die ich ja vorhin schon erwähnt hab glaubst du, da hat es irgendwie Auswirkungen, dass man vielleicht dann noch mal bereitwilliger auch mit Leuten, die man eben dem man nicht vertraut Unternehmen Daten austauschen würde oder glaubst du so ein Grundvertrauen braucht man immer egal, zu welcher mit welcher Technologie?

42 [0:22:15] **B1:** Also ich weiß nicht, wie du dich fühlen würdest, wenn du jemandem deine Daten schicken würdest, dem du nicht vertraust, ich würde es nicht machen. Also von daher würde ich das

auch aus Unternehmenssicht so sehen, dass man sagt ja wenn da wenn da kein Vertrauen da ist, wenn ich den Anbieter nicht kenne, wenn der nicht ich sage jetzt mal zertifiziert ist und sowas dann. Merk ich dem lieber keine Daten, bevor da doch irgendwie was passiert? Und ja, das würd ich dann auf Unternehmenssicht entsprechend so bewerten zu sagen ja n Grundvertrauen muss schon da sein.

43 [0:22:38] **I:** Mhm. Ja ja und siehst du neue Risiken, die entstehen?

44 [0:22:57] **B1:** Uh gute Frage, neue Risiken? Müsste ich jetzt tatsächlich einmal überlegen?

45 [0:23:04] **I:** Paar Beispiele Risiken geben ob, ob du die sinnvoll siehst oder nicht zum Beispiel also, es ist ja alles verschlüsselt und man kann sich ja vertrauen, aber es kann natürlich auch sein, dass. Dass irgendwie jetzt das andere Unternehmen Daten von schlechter Qualität einspeist in die Technologie oder irgendwelche Daten zur Manipulation von Ergebnissen. Ja siehst du sowas als Gefahr oder glaubst du sowas wird dann eh nicht gemacht? Gemacht. Wenn es, wenn es dann natürlich auch noch andere Verträge und alles gibt.

46 [0:23:36] **B1:** Ja. Ja, würde ich schon so als berechtigt ansehen, ist da natürlich dieses Thema Vertrauen wieder vielleicht nicht in den Anbieter, sondern in den Vertrag, das Vertrauen in alle Teilnehmer dieses Verfahrens, wie du es gerade sagtest, dass man da schon vertrauen muss, ist ja jeder auch dann. Seine Daten teilt und nicht irgendwie was zurückhält, nur um die besten Ergebnisse für sich selber abzuziehen, denn ich glaube, von diesem Verfahren leben. Oder das Verfahren lebt davon, dass alle gleich beitragen oder vielleicht der anderen ein bisschen mehr als der als der der, der die andere Partei aber schon, dass alle auf einem ähnlichen Level Daten da auch mit Reinspielen zumindest wenn man jetzt ein Wettbewerber ist. Ansonsten klar hat es auch Vorteile für die Unternehmen, die vielleicht gerade weniger Daten zur Verfügung haben als andere Unternehmen und daraus mehr nutzen ziehen als die anderen und deshalb, das ist der Punkt, den ich eben meinte der Nutzen muss halt für die Unternehmen da sein, für die eine ist er vielleicht größer als für die anderen, weil Google zum Beispiel hat jetzt Milliarden von Daten und ich sag jetzt mal die kleine Klitsche aus was weiß ich? Die hat vielleicht ganz wenig Daten und die hat natürlich einen viel größeren Nutzen davon als vielleicht Google, die das jetzt schon machen. Deshalb kommt es immer so ein bisschen darauf an würde ich jetzt auch mal wieder behaupten? Aber natürlich steht halt das Thema Kosten sehr, sehr stark, dann auch und nutzen dann entsprechend in einem Fokus würde ich jetzt mal sagen das wäre ja natürlich genau und und ein Risiko vielleicht, wenn wir jetzt von einer riesen Datenmenge haben natürlich das Thema Performance, also wenn es jetzt auf meinem System stabil läuft und ich hab dann auf einem anderen System auf auf einer Cloud, wo alle, die die Daten einspielen performanceverluste und kann nicht mehr so richtig meine Arbeit ausführen, dann ist es natürlich da deutlich schwieriger das dann jetzt einzusetzen.

47 [0:25:13] **I:** Ja, wie im ich kostenlos ne ja.

48 [0:25:38] **B1:** Kein Risiko?

49 [0:25:40] **I:** Okay, dann hab ich jetzt eigentlich eh nur noch 2 letzte Fragen die eine siehst du irgendwelche? Unternehmen.
Die vielleicht die Technologien mehr nutzen würden oder eher bereit wären, da zu investieren, von so Größe, Sektor oder Region, wo die herkommen oder glaubst du es kommt einfach nur auf den Use Case an und durch die Bank? Alle Unternehmen könnten die Technologie nutzen.

50 [0:26:09] **B1:** Also ich könnt mir auf jeden Fall vorstellen, dass so die ja US amerikanischen Tech Firmen da also wie wie Google, Microsoft, Apple, Nvidia zum Beispiel in die Richtung oder meta da vielleicht auch schon sehr, sehr stark unterwegs sind, das weiß ich jetzt nicht, aber natürlich, so Technologieunternehmen, die vielleicht auch gerade etwas weniger reguliert sind, würde ich jetzt mal sagen in die Richtung da vielleicht eher. Bereit wären oder ja die, die ich eher im Fokus da sehen würde und einen großen nutzen würde ich glaube ich echt für die Pharmaunternehmen oder Medizinhersteller sehen, weil ich weiß nicht ob du es kennst, wenn man so ein neues Medikament zulässt, dann sind natürlich sehr, sehr viele, viele klinische Studien zum Beispiel notwendig. Sehr viele Tests, sehr viele Analysen und ich glaube, dadurch könnte sich sehr, sehr stark, dann natürlich auch dieser dieses Zulassungsverfahren oder dieses Testverfahren der Medikamente verkürzen, einfach durch den Austausch durch den Datenaustausch und und den Benchmark mit anderen Unternehmen (..) Und da dann entsprechenden Vorteil haben.

51 [0:27:18] **I:** Ja interessant, ja, ja.

52 [0:27:20] **B1:** Genau und von der Größe her der Unternehmen ist es glaube ich eher schwieriger einzuschätzen. Das können die großen Unternehmen sein, wie natürlich hier eine Microsoft ein Apple oder sowas, aber auch kleinere Unternehmen, die entsprechend da aufspringen können auf so eine Cloud oder auf so einen Verfahren. Und dann Daten von anderen Unternehmen mit nutzen können, die profitieren natürlich auch und ich denke mal, man muss halt ein bisschen affin sein, was diese Sache angeht und das Hintergrundwissen haben und und natürlich auch für sich für sein Unternehmen, den Use Case abstecken.

53 [0:27:47] **I:** Ja, ein. Was sind Deiner Meinung nach die größten Herausforderungen für die Akzeptanz von Secure Multiparty Computation auf dem Markt?

54 [0:28:09] **B1:** Auf jeden Fall muss ich wissen, worum es sich handelt. Ich muss die Technologie verstehen, ich muss den Algorithmus vielleicht verstehen oder zumindest wissen darüber haben, wie ich es einsetze.

Und je größer dieses Wissen ist, desto größer ist glaub ich auch die Akzeptanz von den entsprechenden Verfahren und Technologien. Weil ich mach ungerne Sachen, von denen ich nicht weiß, wie sie funktionieren. Und was ihr mir für nen Nutzen dann bringen und für welche Kosten dann natürlich entstehen und ich glaub das ist so ne Grundvoraussetzung für die Akzeptanz einfach zu sagen ja, ich muss verstehen oder? Die Unternehmen müssen verstehen. Was kann ich damit machen, was bringt es mir? Und dann bin ich natürlich auch bereit, da rein zu investieren und das Thema anzugehen. Und super sind natürlich irgendwie Experten, die man dann im eigenen Unternehmen hat, die dieses Thema kennen, die das Thema vorantreiben. Und die dann entsprechend natürlich auch sagen ja, das passt, das passt nicht und dann natürlich klar, man muss das Management von sich überzeugen, wenn man solche Sachen macht, denn das ist natürlich die Ebene, die die Punkte dann entscheidet, ob man da rein investiert, ob man die Sachen und die Verfahren nutzt.

55 [0:29:09] I: Ja. Ja. Mhm, Mhm, alles klar, dann vielen, vielen Dank gibt es noch irgendwas, was du, was ich jetzt nicht gefragt hätte, was du vielleicht noch wichtig an dem Thema siehst oder so?

56 [0:29:32] **B1:** Sehr gerne.

57 [0:29:40] **I:** Oder ob ich alles ganz gut.

58 [0:29:44] **B1:** Ja, also vielleicht auch der Punkt, dass es klar so wenn man sich die rechtlichen Rahmenbedingungen natürlich anschaut, dass es da glaube ich, je nach Daten unterschiedliche Hürden gibt, zum Beispiel bei personenbezogenen Daten werden die Hürden wahrscheinlich schwieriger sein oder höher liegen als jetzt bei anderen Daten, die man vielleicht so im Unternehmen hat, die jetzt keine personenbezogene Daten sind und dann natürlich ja welches wissen möchte ich preisgeben? Wie geheim sind meine Daten und kann ich darauf vertrauen? Aber ich glaub wenn man wenn man aus der aus der Brille schaut, möchte ich es anwenden, dann ist es stärker, als wenn man da schaut. Was für rechtliche Verfahren oder rechtliche Rahmenbedingungen gibt es denn? Die müssen vielleicht wahrscheinlich zum Teil noch gelegt werden, wenn das Verfahren in einer größeren Masse angewendet wird. Wir sehen es die aktuelle Reg. Regularien ändern sich gerade wir haben den AI Act, wir haben ja gut die Datenschutzgrundverordnung, die jetzt schon ein bisschen länger gilt. Jedes Land hat vielleicht noch andere rechtliche Rahmenbedingungen, wie wir das Bundesdatenschutzgesetz oder das TDDDG für Telekommunikationsanbieter zum Beispiel oder Internetdienstleister. Dann gibt es in der EU ja gerade viele viele Punkte, die wir da berücksichtigen und diese Digitalmarkets Digital Services Act, die alle darauf einspielen können oder einen Einfluss haben können natürlich in der in der Nutzung von diesen Verfahren und Technologien. Und ja, da ist klar je nach Region ist es dann ein bisschen unterschiedlich, wie stark die Verfahren aktuell schon zum Einsatz kommen können.

59 [0:31:21] **I:** Was braucht es denn dafür, dass man dann mal sicher weiß, ob Secure Multiple Party computation rechtlich erlaubt ist oder nicht? Gibt es da irgendwie einen Fall, mal der dann.

60 [0:31:34] **B1:** Das müsste man ja, das müsste man anhand des des jeweiligen Use Cases durchspielen.
Würde ich da sagen natürlich zu welchem Zweck werden die Daten einmal verarbeitet, dann ausgetauscht und weiterverarbeitet, habe ich dann noch einen Personenbezug? Zum Beispiel bei drin in den Daten, denn selbst bei einer Anonymisierung musst du ja zum Beispiel das datenschutzrechtlicher Sicht sagen ich anonymisiere oder ich verarbeite die Daten. Zum Zwecke der Anonymisierung und danach habe ich halt das Risiko sehr, sehr minimal, dass ich da wieder Personenbezug herstellen kann, aber das muss natürlich alles berücksichtigt werden und am einfachsten ist es da wirklich über einen Use Case zu gehen und zu sagen, so sieht es aus das ist meine Idee, das möchte ich damit verwirklichen und da dann zu sagen ja. Wie sehen die rechtlichen Rahmenbedingungen aus? Wieviel KI steckt zum Beispiel in diesem Verfahren oder wo wird KI genutzt und welche Daten werden dann natürlich auch genutzt? Also da kann man dann genauer abschätzen, welche ja Kontrollen man zum Beispiel auch dafür machen muss oder welche Parteien man einbinden muss.

61 [0:32:40] **I:** Mhm
Ja, dann vielen Dank.

62 [0:32:47] **B1:** Gerne.

1 [0:00:00.0] Interview B2

2

| | |
|---|---|
| Interview-Nr. | 2 |
| Date of the interview | August 30, 2024 |
| Duration of the interview | 22:27 min |
| Interviewer | Felix Starnecker (I) |
| Interviewee | B2 (Germany) |
| Role | Founder of IT-security company |
| Sector | IT - Security |
| Specialities | Interview in german |

3 **I:** Was würdest du erfahrungsgemäß behaupten? Mit wem tauschen Unternehmen Daten aus? (.) in deinem Unternehmen. Sind dies eher Wettbewerber, Lieferanten, Kunden (.) in welcher Beziehung stehen die dazu?

4 [0:00:17.4] **B2:** Also, da gibt es mehrere Möglichkeiten. Also zum Einen, Lieferanten ist ganz klar, die tauschen Dokumenten mit ihren Lieferanten aus - ob dies Aufträge sind, Beilagen zu den Aufträgen. Aber hier ist natürlich wichtig, dass die Lieferanten (.) das alles sehen, das heißt, wenn dies verschlüsselt wäre, kann ich nichts anfangen damit. Das heißt, die nutzen dann unsere Plattform und tauschen darüber das aus. Sie wissen, dass kein Fremder hinkommt, aber sie wissen, dass der, dem sie das Zugriffsrecht geben, da auch hinkommt. (.) Also der Lieferant. Das Gleiche in die andere Richtung - also mit den Kunden. Ist natürlich genau das Gleiche. Das heißt, natürlich tauschen unsere Kunden, die Daten mit ihren Kunden aus. (..) Und wir haben sogar die Situation, dass sie die Daten mit Wettbewerbern austauschen, weil wir ja eine Plattform zur Verfügung stellen. Das heißt, wenn jetzt ein Unternehmen "A" einen Wettberwerber kaufen möchte und dann gibt es so eine Analyse - sprich hier wird alles angeschaut (.) die ganzen rechtlichen Sachen, die finanziellen Sachen (.) Was gibt es für Prozesse? Wie waren die finanziellen Ergebnisse? Wie viele Mitarbeiter haben sie? Was verdienen die Mitarbeiter? Dies sind natürlich Unmengen an Daten. Aber die müssen natürlich auch wieder im Klartext an den Wettbewerber gehen, weil das bringt ihm gar nichts, wenn er mit den Daten nichts anfangen kann. Natürlich unterschreiben sie dann irgendwelche Geheimhaltungsvereinbarungen und geben dies nur "pö a pö" frei. Das heißt, sie laden dann in den Datenraum ein (.) vielleicht drei verschiedene Interessenten, die das Unternehmen kaufen wollen und dann schauen sie sich immer den ersten Satz von Daten an und dann geht man mit einem weiter und dann bekommt dieser eben noch mehr Informationen (.) und noch mehr Informationen, bis er

eben alle Informationen hat. Weil er wird den Vertrag nicht unterschreiben, die Firma zu kaufen, wenn nicht alle Informationen dargelegt wurden.

5 [0:02:09.8] **I:** Aber sie vetrauen euch, dass ihr nicht

6 [0:02:12.5] **B2:** Sie vetrauen uns, dass kein Fremder (.) also es ist praktisch so, dass wir den Datenraum zur Verfügung stellen und unser Kunde bestimmt, wer darauf zugreifen kann. Er richtet dann die User ein über eine Email Adresse, dann legen sie ihr Kennwort an und dann darf er darauf zugreifen. Er bestimmt zum Beispiel, diese Datei darfst du sehen (.) die darfst du nicht sehen oder die Datei darfst du auch im Original sehen, die Datei darfst du nur als PDF mit einem Wasserzeichen sehen.

7 [0:02:40.4] **I:** Aber beruht dies nur auf Vertrauen? Oder müsst ihr auch Verträge unterschreiben, dass ihr alles geheim haltet?

8 [0:02:45.5] **B2:** Das machen wir sowieso. Das ist absoluter Standard. Das heißt, es gibt da so einen Auftragsverarbeitungsvertrag. Diesen schließen wir mit jedem Kunden automatisch ab. Und da steht genau drinnen, dass erstens wir nicht darauf zugreifen, dass unsere Mitarbeiter erst gar nicht darauf zugreifen können und bla bla bla. Des ist einfach eine Grundvoraussetzung bevor irgendwelche Daten ausgetauscht werden.

9 [0:03:07.7] **I:** Würdest du daraufhin sagen, dass Vertrauen oder Kontrolle wichtiger ist mit Blick auf Datenaustauschaktionen bei Unternehmen?

10 [0:03:15.0] **B2:** Also, dass was wir machen (..) also mit den Leuten, denen müssen sie ja die Daten unverschlüsselt zur Verfügung stellen. Wir sind ja bloß der Intermediär, der die Plattform zur Verfügung stellt. Aber der Empfänger sieht immer die vollständigen Daten im Klartext - komplett unverschlüsselt. Also jetzt für uns wäre dies überhaupt gar keine Anwendung. Dies würde jetzt unseren Kunden genau gar nichts bringen.

11 [0:03:49.9] **I:** Wenn sich jetzt Unternehmen dazu entscheiden würden, bei euch ihre Daten auszutauschen. Würdest du sagen, ist dies dann eher ein subjektiver oder objektiver Entscheidungsprozess? Also machen sie erstmal irgendwelche Risikoanalysen (.)

12 [0:04:08.2] **B2:** Nein, die haben den Bedarf ihre Daten auszutauschen. Und dann kommen sie zu uns. Und wie gesagt, immer nur Datenaustausch (.) Ich sage jetzt mal, dass Einzige, was man nicht

verschlüsselt (.) wie es eingeschränkt werden kann, dass man praktisch Sachen nur im Browser anschauen kann. Und du kannst es vielleicht nur als PDF anschauen und das hat auch noch ein Wasserzeichen, aber du kannst es nicht downloaden. Und dann gibt es halt sämtliche Abstufungen bis "ich kann die Originaldatei" downloaden.

13 [0:04:43.4] **I:** Mhm. Und ganz allgemein gefragt, welche Risiken siehst du beim Austausch von Daten? Egal, ob in eurem Unternehmen oder ganz allgemein betrachtet.

14 [0:04:56.3] **B2:** Wie gesagt, mit Lieferanten, Kunden, da muss ich Daten austauschen, weil sonst kann der Lieferant seine Aufgaben nicht erfüllen.

15 [0:05:05.7] **I:** Also besteht eher das Risiko, wenn man keine Daten austauscht quasi?

16 [0:05:08.6] **B1:** Genau! Wichtig ist (.) oder ich sag jetzt mal im Bauprozess. Wir haben viele Kunden vom Bau - da werden Protokolle erstellt, da gibt es ständig neue Pläne. Die werden in ein System eingestellt. Dann werden alle automatisch benachrichtigt "hey, da ist ein neuer Plan. Bitte berücksichtigt den neuen Plan und nicht den Alten." Weil sonst macht der irgendwas, was nicht mehr dazu passt. Also es ist eher wichtig, dass sie immer die aktuellsten Daten haben. Und so eine Anwendung, dass mehrere Parteien gemeinsame Berechnungen auf irgendetwas anstellen wollen ohne dass die anderen genau wissen, was das ist (.) gibt es bei uns überhaupt gar nicht.

17 [0:05:52.7] **I:** Aber wenn es ganz klar besser ist, Daten auszutauschen (.) deutsche Unternehmen tauschen ja immer noch relativ wenig Daten aus im Vergleich zu anderen Unternehmen.

18 [0:06:08.7] **B2:** Weiß ich jetzt gar nicht. (.) also man muss wirklich unterscheiden, Lieferanten und Kunden. Mit denen muss man sich natürlich austauschen. Da brauch ich irgendeine Plattform. Und dieser Prozess ist natürlich ein sehr sensibler Bereich. Wenn ich jetzt zum Beispiel sagen würde, ich möchte jetzt diese Firma kaufen - die legen mir alle ihre Daten offen und ganz zum Schluss sage ich "Nein, ich kaufe euch doch nicht", habe ich alle Informationen über diese Firma. (..) alles. Und dies ist in vielen Fällen aber ein Wettbewerber und dann weiß der alles über mich (.) der weiß, wie viele Mitarbeiter, was die können, wie die heißen, was die verdienen, wie viel Umsatz ich mache, in welchen Bereichen ich meine Umsätze mache, wie mache ich meinen Vertrieb, wie mache ich mein Marketing. Die kennen alles Interne aus der Firma, sie kennen mehr als die Mitarbeiter. Weil die Mitarbeiter sehen ja immer nur einen Teil und sie wissen alles über die ganze Firma. (..) Und wenn das Geschäft nicht zustande kommt, ist dies natürlich ein riesen Nachteil.

19 [0:07:12.4] **I:** Aber wie kann dies verhindert werden?

20 [0:07:15.7] **B2:** Gar nicht. Ich muss ja dies denen geben zum Lesen für ihren Entscheidungsprozess. Also ich muss ihnen vertrauen, dass es ernst ist. Natürlich unterschreiben sie einen Vertrag "Ich beabsichtige, bla bla bla (..) das zu kaufen (.) zu einem Preis von X Euro". Aber dies ist nicht rechtlich bindend. Sie können irgendwann sagen "Nein, wir wollen es doch nicht machen."

21 [0:07:39.7] **I:** Theoretisch können sie sich also nur als Käufer ausgeben, um Informationen zu bekommen? (Kann man machen!) Hindert sie ja eigentlich auch nichts daran?

22 [0:07:47.1] **B2:** Nein, natürlich sind hier auch immer Berater im Spiel, die kennen die natürlich dann auch untereinander und (.) Also das wird schon (..) das macht man nicht einfach so, dass du die Daten hergibst. Und was natürlich auch ist, da gibt es unterschiedliche Teams bei den Käufern, Die einen dürfen des sehen, die anderen dürfen des sehen. Das heißt, dass nicht eine Person alles sehen darf, sondern der, der für Finanzen zuständig ist, der schaut sich die Finanzen an; der für Rechtliches zuständig ist, schaut sich alle rechtlichen Belange an usw.. ((Unterbrechung durch Klingel))

23 [0:08:51.5] **I:** Kannst du dir irgendwelche externen oder internen Einflüsse vorstellen, wieso Unternehmen ihre Daten teilen oder eben nicht teilen? Also was sie daran hindern könnte?

24 [0:09:04.5] **B2:** Also was mir ehrlich gesagt nicht klar ist, ist die Anwendung. Also wenn ich mir jetzt mich anschaue, was hätte ich für ein Interesse dran, mit anderen Daten zu teilen? Also was wäre jetzt in meinem Interesse und in dem des Mitbewerbers oder irgendwelchen anderen Akteuren - es können ja auch Verbände sein, die irgendeine Statistik machen wollen über einen bestimmten Firmenbereich und fragen von hundert Firmen das ab und da könnte jeder dann anonymisiert seine Daten abgeben. Und dann hätten alle einen Überblick über das Ganze (..) Ja, das könnte ich mir eventuell vorstellen.

25 [0:10:03.7] **I:** Aber ich meine jetzt, bei externen oder internen Einflüssen eher so (.) ob Unternehmen bereitwilliger sind, ihre Daten zu teilen, wenn die Unternehmenskultur offener ist oder wenn makroökonomische Dinge dafür sprechen, ihre Daten zu teilen. Der letzte Interviewpartner hat mir gesagt, dass die deutschen Autohersteller jetzt vermehrt gezwungen sind, ihre Daten miteinander auszutauschen, weil sowohl die Chinesen als auch die Amerikaner im Bereich des Autoherstellung zur Konkurrenz werden. Aus diesem Grund müssen sie nun verstärkt zusammenarbeiten.

26 [0:10:48.2] **B2:** Ich glaube, dass Entscheidende ist, was die Anwendung ist. Wenn die deutschen

Autobauer sich alle gegenseitig unterstützen wollen gegen die Chinesen, dann macht es natürlich Sinn, Daten auszutauschen. Es geht immer um die Anwendung. Und natürlich hilft es, wenn du eine Kultur hast, wo du offener bist und mehr Sachen nach außen trägst.

27 [0:11:29.9] **I:** Oder würdest du sagen, manche Sektoren teilen eher Daten als Andere?

28 [0:11:38.3] **B2:** Unsere Kunden teilen natürlich die Sachen im Klartext, weil sie es wollen, weil sie es müssen, weil sie einen Bedarf haben. Also es ist ja nicht so, dass wir versuchen, es jemanden zu verkaufen, der es nicht braucht, sondern die Leute kommen zu uns, weil sie einen Bedarf haben.

29 [0:11:55.0] **I:** Und zu PETs wie SMPC siehst du eher weniger Anwendungsfälle? Oder wie würdest du spontan die Technologie einschätzen?

30 [0:12:09.7] **B2:** Also ich müsste vielleicht nochmal in Ruhe darüber nachdenken, aber ad hoc fällt mir wenig ein (.) oder nichts. Also wenn ich jetzt dies anbieten würde, dann wüsste ich nicht wen ich ansprechen sollte.

31 [0:12:30.0] **I:** Also besteht in diesem Bereich einfach keine Nachfrage?

32 [0:12:33.4] **B2**: Also ich wüsste keine.

33 [0:14:07.5] **I:** Glaubst du, wenn PETs wie SMPC mehr genutzt werden, dass dies das Datenaustauschverhalten von Unternehmen verändern würde? (.) dass mehr ausgetauscht wird? dass sensiblere Daten ausgetauscht werden? Dass mehr Konkurrenten zusammenarbeiten?

34 [0:14:32.9] **B2:** Also ich kann mir vorstellen, dass bei den Unternehmen auch kein Wissen da ist, dass man auch Daten austauschen könnte, ohne dass die Beteiligten das Genaue wissen (.) also praktisch die genauen Daten kennen. Das würde es sicher erleichtern. Aber wie gesagt, mir fehlt der Anwendungscase. Was sollen jetzt zum Beispiel die Autobauer austauschen? Also welche Art von Informationen sollten sie austauschen, wo sie gemeinsam etwas berechnen und wo sie dann alle etwas davon haben?

35 [0:15:16.1] **I:** Vielleicht so etwas wie Prozessvergleiche? Wo man sieht, ob eventuell die Prozesse bei dem einen Autobauer besser funktionieren als bei einem Anderem?

36 [0:15:22.5] **B2:** Dies sieht man ja nicht. Weil hierfür müsste man ja die Daten wirklich hergeben. Du

kannst praktisch nur einen Durchschnitt machen, woraus du schließen kannst, ob du über oder unter dem Schnitt liegst. Ich meine, das hilft vielleicht auch etwas, aber dann müsste man eindeutige Messgrößen definieren.

37 [0:17:15.6] **I:** Wie siehst du die Auswirkungen von PETs wie SMPC auf Vertrauen und Kontrolle?

38 [0:17:56.5] **B2:** Also wenn die Technologie sicherstellt, dass es sozusagen egal ist (..) dann fällt das Vertrauen natürlich jedem leichter.

39 [0:18:13.8] **I:** Gibt es irgendwelche Voraussetzungen, dass Unternehmen in PETs wie SMPC investieren? Kostet dies viele Ressourcen? (.) also muss man groß genug sein, um dies machen zu können?

40 [0:18:30.1] **B2:** Also ich würde mal davon ausgehen, dass das banal ist. Da muss halt jeder ein Stück Software installieren und dann muss man wahrscheinlich überhaupt definieren, was man austauschen will und ich gehe mal davon aus, den Rest macht die Software.

41 [0:18:43.2] **I:** Also jeder kann dies machen, wenn es Anwendungsfälle dafür gibt.

42 [0:18:46.9] **B2:** Ja, also ich glaube nicht, dass das Problem ist, dies einzuführen.

43 [0:19:09.1] **I:** Und siehst du neue Risiken, die daraus entstehen (.) wenn man so etwas nutzt?

44 [0:19:17.1] **B2:** Nein, würde ich jetzt keine sehen. ((Unterbrechung))

45 [0:22:18.1] **I:** Jetzt sind wir eigentlich auch schon am Ende. Danke für das Interview.

46 [0:22:24.3] **B2:** Gerne. Tut mir leid für die Unterbrechungen.

1 Interview B3

2

| | |
|---|---|
| Interview-Nr. | 3 |
| Date of the interview | august 24, 24 |
| Duration of the interview | 36:12 min |
| Interviewer | Felix Starnecker (I) |
| Interviewee | B3 (Germany) |
| Role | Cybersecurity consultant |
| Sector | IT Security |
| Specialities | Interview in german |

3 [0:00:00] **I:** Also ich bin Student an der Universität Passau und schreibe zur Zeit meine Masterarbeit über das Thema, welche Faktoren die Entscheidung von Unternehmen, ob sie ihre Daten mit anderen Unternehmen teilen, beeinflussen und wie sich die Nutzung von Privacy enhancing Technologien darauf auswirkt und Ziel von dem Interview ist es, Unternehmen und ihre Abwägung bezüglich Datenaustausch besser nachvollziehen zu können. Ich werde natürlich alle Antworten nur in anonymisierter Form in meiner Masterarbeit verwenden. Passt das für Sie?

4 [0:00:58] **B3:** Ja, das ist OK.

5 [0:01:00] **I:** Ähm gut, dann meine erste Frage wer könnten sie mir bitte kurz, Ihr Unternehmen und Ihre Rolle innerhalb des Unternehmens beschreiben?

6 [0:01:11] **B3:** Mhm klar also, ich bin Mitarbeiter bei [Unternehmen]. Das ist der deutsche Mutterkonzern von [Unternehmen]. Wir machen vor allem 3 verschiedene Sachen. Das ist zum einen SI, das ist Smart Infrastructure, die machen Gebäude und Gebäudeautomatisierung. Dann haben wir DI, das Digital Industries, die machen Automatisierung für die Industrie sowohl für Prozesse als auch für generelle Abläufe. Und dann haben wir noch die dritte Business Sparte, die machen [unternehemnsspezifische Informationen]. Abgesehen davon haben wir dann noch [unternehemnsspezifische Informationen]. Aber ich glaube, das geht zu weit das alles zu erklären. Zu meiner Person ich bin in der Corporate, also in der zentralen Cyber Security Abteilung, da habe ich mehrere Hüte auf neben dem normalen Cyber Security Consulting mache ich auch die, wie wir es nennen können Global Security Transformation Method Methodology, das heißt, ich entwerfe die zentrale Methodik, wie wir in allen [Unternehmen] Werken IT Security gestalten vorantreiben und verbessern das ist der erste Hut, der zweite Hut ist das Training. Wir haben ein Trainingsprogramm

für IT Security oder auch Cyber Security, das wir in der kompletten Firma versuchen aufzubauen und auszurollen und abgesehen davon mache ich noch Consulting für die internen Siedlungsfabriken als auch für die externen Kunden, die auf uns zukommen. Für alle Bereiche von IT Security.

7 [0:02:49] **I:** Mhm OK alles klar, sehr interessant. Wie wichtig würden Sie allgemein das Teilen von Daten mit anderen Unternehmen für die Strategie und den Erfolg von [Unternehmen] einschätzen?

8 [0:03:04] **B3:** Würde ich allgemein als essentiell einschätzen. Wir haben da zum Beispiel das Chart of Trust, das ist bei uns das Größte, wenn man jetzt speziell auf Cyber Security schaut, wo wir Informationen austauschen. Aber darüber hinaus haben wir noch mehrere andere ich sag mal Prozesse, wo wir Daten austauschen. Da gibt es sowas wie den Dax Ruf und dann gibt es noch Vereine, wie zum Beispiel den CSSA wo ich Mitglied bin, wo Unternehmen vertraulich Daten austauschen können. Das ist jetzt speziell in der Cyber Security. In anderen Geschäftsbereichen ist das ein bisschen anders und da habe ich nicht so viel Einblick, aber bei Security speziell sind das so die Hauptaspekte, wo ich betonen will, dass vor allem Thread Intelligence immer wichtiger wird. Das heißt Informationen Austausch bezüglich möglichen Angriffsvektoren und was ist gerade so da draußen? Was könnte demnächst kommen, wo kann man von anderen Unternehmen lernen, bevor es ein selbst trifft? Das sind so die die Hauptaspekte in der Cybersecurity und ich glaube auch, dass es jetzt essentiell ist und in der Zukunft immer wichtiger wird.

9 [0:04:09] **I:** OK und teilt [Unternehmen] im Moment Daten mit anderen Unternehmen, wovon Sie wissen?

10 [0:04:16] **B3:** Ja ja genau also es gibt den Chart of Trust, der wird genutzt, wenn wir mit anderen Unternehmen Daten teilen. Dann haben wir den CSSA, wo das dann über die Experten läuft, wie zum Beispiel bei mir. Ich bin dann im IT Security Workstream vom vom CSSA drin der Working Group. Dann gibt es da noch mehrere andere Working Groups für Thread Intelligence oder für IT Security und dann auch beim Dax Ruf da kann man dann eine Frage stellen oder einen Fragebogen R stellen, der geht dann an alle anderen Unternehmen raus eben auch in diesem Dax.Ruf teilnehmen und so kann man Informationen austauschen, also das findet sehr oft und sehr, sehr rege statt.

11 [0:04:54] **I:** OK und der Hauptaustausch ist wegen Cyber Security Aspekten oder ist das jetzt nur, weil sie in dem Bereich arbeiten?

12 [0:04:59] **B3:** Ich denke, es ist vor allem in der Cybersecurity wichtig. In anderen Bereichen findet das sicher auch statt, ist aber da von dem, was ich gehört habe, begrenzter. Aber da muss man eben auch sagen, dass ich in der IT Security angestellt bin, deswegen hab ich bei anderen wahrscheinlich nicht den größten Einblick.

13 [0:05:17] **I:** Und sind es dann Partner, die mit ihnen konkurrieren auch oder was für Partner sind das, mit denen Sie Informationen teilen?

14 [0:05:28] **B3:** Eine gute Frage. Das ist immer unterschiedlich, also in den verschiedenen Prozessen zum Informationsaustausch sitzen verschiedene Partner mit drinnen. Und natürlich tauschen Unternehmen lieber nicht mit Konkurrenten Informationen aus. Und meistens gibt es halt irgendeine Verbindung, also beim CSR zum Beispiel ist es halt so, dass das deutsche Unternehmen sind und dann dementsprechend, da der Austausch darauf fokussiert ist. Bei anderen Austausch Programmen gibt's dann immer eine Gemeinsamkeit, die die Unternehmen zusammenführt. Und bei Konkurrenten ist es halt dann immer eine Frage wo kann man was austauschen? In Cyber Security ist es halt so, dass es da nicht so eng gesehen wird, weil die Cybersecurity an sich ein sehr offenes Feld ist. Deswegen würd ich zusammenfassend sagen, es gibt immer einen gemeinschaftlichen Faktor und es sind vorzugsweise keine Konkurrenten, aber auch mit Konkurrenten werden in gewissen Bereichen Daten ausgetauscht.

15 [0:06:30] **I:** OK und gibt es dann irgendwelche Verträge, mit denen oder wie werden die Daten ausgetauscht? Wie läuft sowas ab?

16 [0:06:40] **B3:** Mhm also, da gibt es im Unternehmen von unserer Größe mehrere Wege, die man nutzen kann. Da gibt es zum einen sozusagen eine Plattform selber, die wir anbieten, auf der man sicher Daten austauschen kann, wo das alles mit Verschlüsselung und so weiter 2 Faktor Authentifizierung mit integriert ist. Das ist der offizielle und der beste Weg, aber bei jeglichen Datenaustausch gibt es immer einen Vertrag vorher, selbst wenn wir nicht mal über Datenaustausch reden, sondern einfach nur über Zusammenarbeit ist eigentlich immer ein, ich sag mal Unternehmensprozess dahinter, der sicherstellt, dass es da ein NDA gibt, dass da die Rechtsabteilung ihren Segen gegeben hat. Das heißt selbst wenn wir nicht über Datenaustausch reden, was schon ein kritisches Thema ist, sondern nur über zum Beispiel ein Zulieferer, dann gibt's da schon für ein Supply Assessment Tool, wo es intern dann eine Liste gibt. Dann werden auch verschiedene Sachen geprüft. Es müssen verschiedene Verträge unterschreiben werden, bevor die überhaupt was zuliefern dürfen. Und beim Informations und Datenaustausch ist es dann meistens sogar noch strenger.

17 [0:08:29] **I:** Wer entscheidet da, ob das jetzt ausgetauscht wird oder nicht, ist das eine Person, ist das eher so ein subjektiver oder objektiver Entscheidungsprozess? Würden sie sagen, dass ja läuft über viele Personen?

18 [0:08:56] **B3:** Gute Frage. Also, wir bei [Unternehmen] verfolgen allgemein immer das 4 Augen Prinzip. Das heißt, solche Sachen werden immer von mindestens 2 Personen abgesegnet. Das ist ne unternehmensvorgabe und in allen Prozessen würde ich sogar sagen es geht darüber hinaus, weil

solche Sachen sehr vorsichtig gehandhabt werden. Also das geht normalerweise über mehrere Tische. Da schaut der Chef drüber. Der Chef vom Chef und wenn wir über definiertere Prozesse sprechen, wie zum Beispiel die Supply Assessments, die ich vorher erwähnt habe, dann gibt es da sogar Abteilungen. Also wir haben eine Abteilung, die nur für Supplier (.) Management Supply Chain Management SCM heißt die (.) die kümmert sich all Day Everyday nur um das. Dann die Experten, wo man dann als normaler Mitarbeiter sagt: „Hey, ich habe hier einen Supplier, mit dem würde ich gerne zusammenarbeiten, oder mit dem würde ich gerne was austauschen." Dann sagen die: „okay pass auf, folgende Prozess Schritte musst du befolgen, ich helfe dir bei jedem Schritt, wir machen den Kontakt mit den Leuten von der Rechtsabteilung." Also, da gibt es sehr viele Ladevorgaben und Leute, die das sich drum kümmern.

19 [0:10:00] **I:** OK, du würdest dann sagen, dass es also zum Beispiel jetzt, wenn man jetzt von der Individuen Ebene ausgeht, dann ist ja oft so, dass man bei Daten Preisgabe Entscheidungen relativ spontan entscheidet, da sind auch Gefühle involviert und alles. Würdest du sagen, dass das auf Unternehmensebene viel rationaler ist und eher mehr Risikoanalyse und alles oder würdest du sagen, dass da auch die Unternehmenskultur und die allgemeine Einstellung von den Topmanagern bisschen mitschwingt? In dem wie viel Daten geteilt werden oder auch nicht geteilt werden.

20 [0:10:40] **B3:** Das ist eine komplexe Frage. Wo ich aus meiner Erfahrung sagen würde, dass es sehr rational ist, weil jeder sehr vorsichtig ist, und man lieber 23 Unterschriften mehr einholt, bevor man etwas teilt, weil jeder ungern die individuelle Verantwortung übernimmt. Also da werden eher Gefühle beiseite geschoben, und man ist sich wirklich zu 110 % sicher, bevor man etwas teilt, und holt sich die Absicherung von allen Seiten. Man schaut so objektiv darauf, wie ein Mensch das machen kann, bevor etwas geteilt wird. Ich würde schon sagen, es kommt aber auch auf die Unternehmenskultur an.

21 [0:11:23] **I:** Ja, ich wollte gerade sagen, weil vorsichtig muss ja nicht objektiv sein. Zu viel Vorsicht ist ja auch subjektiv. Es kann ja gut sein, dass deutsche Unternehmen einfach vorsichtiger mit dem Teilen von Daten sind als amerikanische Unternehmen

22 [0:11:35] **B3:** Das kann sein, wobei ich sagen kann, dass das bei uns unternehmensweit sehr gepusht wird. Also diese Idee von "Pass darauf auf". Du hast in jedem großen Unternehmen eine Klassifizierung für alle Daten, die an jeden Mitarbeiter sehr streng kommuniziert wird. Alles, was du im Arbeitskontext behandelst, musst du klassifizieren, und jedes Unternehmen hat sein eigenes System. Die meisten nutzen entweder ein 1-4-System oder ein Ampelsystem, wo du dann definierst, okay, dieses Dokument ist auf Level 1 – das könnte man theoretisch öffentlich machen, sollte man aber nicht. Level 2 ist dann intern, und weit verbreitet ist dann "Confidential", das darf nicht mal im Unternehmen geteilt werden, außer die Leute sind dafür berechtigt. Und dann gibt es meistens noch

die letzte Stufe "Strictly Confidential", die dann wirklich niemand sehen darf und sehr speziell behandelt werden muss. Jedes Unternehmen ab einer gewissen Größe hat so etwas, und das hat nichts mit kulturellen Unterschieden zwischen Ländern zu tun. Ab einer bestimmten Unternehmensgröße musst du das einfach einführen, um zu funktionieren.

23 [0:12:44] **I:** Und gibt es irgendwie spezielle Datenarten, die mehr geteilt werden und andere Datenarten, die weniger geteilt werden?

24 [0:12:54] **B3:** Ja, auf jeden Fall. Das, was am meisten geteilt wird, ist natürlich das, was keinen Business-Wert hat, sondern eher der Verteidigung und dem Best Practice Sharing, vor allem im Bereich Cyber Security. Das sehe ich am meisten: Da geht es nicht um Geschäftsgeheimnisse, sondern darum, dass wir uns zusammen gegen die „Bösen“ da draußen verteidigen. Das wird sehr anders gehandhabt als beispielsweise Industrieprozesse. Solche Informationen wird, glaube ich, kein Unternehmen teilen – niemand redet darüber, seine Geschäftsgeheimnisse zu teilen. Es geht eher darum, Dinge zu teilen, von denen alle profitieren, ohne dem Konkurrenten einen Vorteil zu verschaffen. Das betrifft vor allem Threat Intelligence.

25 [0:14:16] **I:** OK, also Industrieprozessdaten werden auf jeden Fall eher nicht geteilt?

26 [0:14:22] **B3:** Ja Ja, genau. Es gibt natürlich auch Ausnahmen, wie bei allen Sachen, aber im Allgemeinen kann man das so sagen.

27 [0:14:29] **I:** Ja, ja. Und jetzt noch mal zurück zu den Partnerunternehmen. Würdest du sagen, dass das eher Partner sind, denen mehr vertraut wird, oder ist das egal? Kommt es einfach auf den Mehrwert der Daten an und nicht auf die Beziehungen zwischen den Unternehmen?

28 [0:14:50] **B3:** Mhm, also wie gesagt, es steckt immer ein sehr komplexer Prozess dahinter, der verschiedene Faktoren einbezieht. Es ist nicht so, dass man sagen könnte, ein Manager vertraut einem Unternehmen mehr, und deshalb teilen wir mehr. Es gibt einen festgelegten Prozess, der Faktoren wie die Unternehmensgröße, das Land, in dem das Unternehmen tätig ist, und wie dieses Land von der entsprechenden Abteilung eingeschätzt wird, berücksichtigt. Dazu kommt, ob sie in bestimmten Bereichen Konkurrenten sind oder nicht, wie viele Mitarbeiter sie haben, wann sie gegründet wurden. All diese Faktoren spielen mit hinein und entscheiden, wie wir mit dem Unternehmen arbeiten. Es geht ja nicht nur um das Teilen von Daten, sondern auch um Aufträge und Compliance-Prozesse, bei denen man aufpassen muss, in welchen Bereichen man mit dem Unternehmen zusammenarbeitet und was das für andere Projekte bedeutet. Bei internationalen Unternehmen ist das sehr komplex. Ich glaube, man kann das nicht einfach auf einen Faktor wie „Beziehungen“ reduzieren.

29 [0:16:01] **I:** Aber Vertrauen könnte eventuell mit einspielen, weil ja die Faktoren wie Land, Industrie oder Mitarbeiteranzahl auch Einfluss darauf haben, wie man das Unternehmen einschätzt, oder?

30 [0:16:18] **B3:** Ja, das kann man am Ende als Vertrauen beschreiben, aber eigentlich handelt es sich eher um eine Risikobewertung. Es heißt, wir haben den Risikofaktor X und den Risikofaktor Y, die beide mit hineinspielen. Basierend darauf treffen wir eine Entscheidung. Das könnte man natürlich auch als Vertrauen beschreiben.

31 [0:16:52] **I:** Und eine letzte Frage zu dem Teil würden Sie sagen, dass [Unternehmen] mehr Daten teilt, wenn es mehr menschliche und finanzielle Ressourcen hat? Oder würde [Unternehmen] mehr Daten teilen, wenn sie da irgendwie noch kompetentere Fachkräfte hätten, weil das ja auch ein komplexeres Thema oder mehr finanzielle Ressourcen, um irgendwie die Plattform besser zu gestalten?

32 [0:17:14] **B3:** Mhm, ich würde das ehrlich gesagt ganz klar verneinen, weil wir, wie vorher schon beschrieben, bereits mehrere Prozesse haben, um Daten zu teilen, die sehr gut sind und sehr gut funktionieren. Da liegt es absolut nicht am Know-how oder daran, dass irgendwelche Funktionalitäten fehlen. Was geteilt werden soll, wird geteilt, und dann wird einer der vielen Wege genutzt, die wir haben – und die funktionieren auch gut. Da sehe ich absolut keine Lücke, ehrlich gesagt.

33 [0:17:52] **I:** OK, gibt es externe/interne Einflüsse, die Unternehmen beeinflussen Daten zu teilen/ nicht zu teilen?

34 [0:18:22] **B3:** Ich würde sagen, da ist externer Druck ein entscheidender Faktor. Das ist das, was mir sofort einfällt, weil wenn zum Beispiel neue Gesetze kommen und man nicht weiß, wie man sie umsetzt oder wie andere Unternehmen damit umgehen, dann kommt schnell die Situation auf, dass die Leute sagen: „Hey, wir müssen uns alle daran halten, lasst uns darüber reden, wie wir das schaffen."

35 [0:19:48] **I:** Gut, dann war das der erste Teil, und im zweiten Teil geht es jetzt um Privacy-enhancing Technologien und wie diese das Datenteilen beeinflussen. Ich fokussiere mich in meiner Arbeit auf PETs wie Secure Multiparty Computation. Das ist eine von vielen Privacy-enhancing Technologien. Es gibt ja zum Beispiel Pseudonymisierung, Anonymisierung und die Nutzung von künstlichen Daten für Analysen. Secure Multiparty Computation ist eine Technologie, bei der mehrere Parteien gemeinsam Berechnungen auf Daten durchführen können, ohne dass die Rohdaten geteilt werden müssen. Oder sie sind verschlüsselt, und es braucht keine zentrale Partei, der man vertraut, um die Daten hochzuladen – das läuft alles dezentral ab. Es gibt also kein Risiko, dass die Rohdaten von der anderen Partei gesehen werden. Am Ende können alle nur die Analyseergebnisse sehen. Diese

Technologie gibt es eigentlich schon relativ lange, aber bisher konnte sie nur für relativ einfache Berechnungen genutzt werden, weil es viel Rechenkapazität braucht. Mittlerweile kann sie jedoch auch für komplexere Aufgaben wie künstliche Intelligenz genutzt werden, indem Rohdaten von vielen Parteien verarbeitet werden. Jetzt wäre die Frage: Nutzt Ihr Unternehmen Multiparty Computation oder andere Privacy-enhancing Technologien auf einem ähnlichen Niveau? Vielleicht nicht nur Anonymisierung, sondern eher so etwas in die Richtung wie Secure Multiparty Computation?

36 [0:21:52] **B3:** Ja, das ist schwierig zu beantworten, weil ich mich frage, was man auf das Level von Secure Multiparty Computation setzt. Natürlich machen wir viele verschiedene Dinge, die irgendwie Privacy-enhancing sind, aber ob sie auf diesem Komplexitätslevel sind, weiß ich nicht.

37 [0:22:17] **I:** Also Secure Multiparty Computation auf jeden Fall nicht, und auch nicht beispielsweise Homomorphic Encryption oder Differential Privacy?

38 [0:22:30] **B3:** Ja, wir machen das auch. Das wird in die Produkte eingebettet, soweit ich weiß. Es gibt ein spezielles Department, das solche Technologien für die Produkte entwickelt – im Grunde eine Sicherheitsabteilung für die Produktentwicklung. Allerdings muss man sagen, dass das im OT-Bereich (Operational Technology) für Industrieumgebungen nicht so entscheidend ist wie im IT-Bereich, weil dort die Sicherheit oder der Datenschutz oft an anderen Stellen vorgeschoben wird. Für das eigentliche Produkt ist es dann nicht mehr so relevant wie normalerweise in der IT. Aber allgemein wird das Thema immer größer, und die Kunden fragen immer mehr danach. Deshalb wird es, glaube ich, noch kommen. Man braucht allerdings immer einen klaren Nutzen. Das ist für mich auch der Punkt bei Secure Multiparty Computation: Man sollte nicht einfach sagen, „OK, das ist ein cooles Konzept, wir wenden es jetzt irgendwo an“ und dann nach einem Anwendungsfall suchen. Es gibt tausend coole Konzepte da draußen, und wenn man so anfängt, wird man nie fertig und macht nie wirklich Geld. Man braucht ein Problem, für das diese Technologie die Lösung ist, oder einen Use Case, wo man sie direkt einsetzen kann.

39 [0:23:57] **I:** Mhm, okay. Zum Beispiel kann man Secure Multiparty Computation hauptsächlich dafür einsetzen, um mit Konkurrenzunternehmen, denen man eigentlich keine Daten geben würde – wie Sie vorhin meinten, wichtigere und sensiblere Daten wie Industrieprozessdaten – gemeinsame Analysen durchzuführen oder vielleicht auch eine künstliche Intelligenz zu trainieren, ohne die Daten wirklich preiszugeben. Man könnte so vielleicht die Supply-Chain-Prozesse verbessern. Könnten Sie sich vorstellen, dass [Unternehmen] das nutzen würde? Finden Sie diesen Anwendungsfall interessant?

40 [0:24:50] **B3:** Interessant auf jeden Fall. Ich würde nur dagegenhalten, dass der Nutzen

wahrscheinlich die Kosten nicht überwiegt, weil wir als Unternehmen unserer Größe selbst genug Daten haben. Es kommt selten vor, dass wir sagen, „OK, wir brauchen das.“ Ich glaube halt, dass man da selten sagt OK, ich hab da so viel Mehrwert daraus, wenn ich da noch ne andere Datenquelle einbinde, als wenn ich einfach mit dem arbeite, was ich schon habe und damit halt den Prozess optimiere Automatisiere oder mir irgendeine KI trainiere. Also das ist eher seltenes Problem, dass wir nicht genug Daten hätten, sondern es fängt viel früher an. Also ich glaub, was da eher gebraucht werden würde, ist irgendein Ansatz um Unternehmen datenlastiger laufen zu lassen. Also, dass man im Unternehmen irgendwie Wege findet die Daten, die man selbst produziert, besser zu nutzen, zu speichern und dann damit zu arbeiten. Das wäre schon mal das erste. Wenn ich dann überlege OK, ich arbeite dann mit einem anderen Unternehmen zusammen und teil die Daten und will dann zusammen darauf Analysen machen, dann frage ich mich auch: „OK, wie greift das andere Unternehmen die Daten ab und machen die das genauso gut wie wir und werden dann da quasi gleichwertige Daten mit eingebracht“. Oder ist es dann so, dass das eine Unternehmen sagt: „OK, wir geben uns Mühe und haben da die richtigen Schnittstellen und geben dann quasi mehr in den Pool als das andere. Ist es nicht vielleicht einfach leichter die eigenen Datenerfassungsprozesse zu optimieren, um dann damit zu arbeiten.“ Weil es auch so ist, dass so riesen Unternehmen ein eigenes Biest ist. Also da sind alle Prozesse intern sehr komplex und sehr aufeinander abgestimmt und fast immer sagt man, so wie wir das machen, ist es besonders und jemand, der mit unseren Prozessen nicht vertraut ist, der muss erstmal in unsere Prozesse lernen, damit ihr überhaupt versteht, wie das bei uns abläuft und dann kann man das erst gut machen. Und da ja ja für mich gibt es da um das zusammenzufassen eher um die Datenerhebung als um das Data Sharing, das wäre eher der Ansatz, den ich da wählen würde.

41 [0:27:35] **I:** Also einfach im eigenen Unternehmen, die Daten richtig strukturieren, erheben und wenn man dann mit anderen teilt, dann ist immer noch die Gefahr, dass die das irgendwie nicht so gut machen, wie man selber oder man selber mehr reinsteckt oder vielleicht die Daten sogar manipulieren, also von Konkurrenzunternehmen ja.

42 [0:27:56] **B3:** Alles so Problemansätze, die man dann glaub ich im Detail durchdenken muss (.) also ich glaube, das ist jetzt halt auch alles sehr generisch, was wir hier reden man müsste da n speziellen Use Case raussuchen und sagen: „OK wie wär's mit dem speziellen Use Case was wäre da benefit was wie würde das ablaufen, wenn da mehrere Unternehmen zusammenarbeiten würden? Welche Daten würden da genutzt werden und was wäre der Outcome?“ Und dann kann man das im Detail diskutieren.

43 [0:28:25] **I:** Das wär eigentlich indirekt meine nächste Frage und zwar wo sie sich vorstellen könnten, in Ihrem Unternehmen SMPC vielleicht zu nutzen. Wäre das dann, wenn man den guten

Use Case hat, der dann halt auch entsprechend Benefits bringt, für den Aufwand? Dann nur dann würde man es quasi versuchen zu nutzen?

44 [0:28:50] **B3:** Ja ja, würd ich schon sagen. Für mich jetzt speziell SMPC ist schwierig, weil die die meisten Sachen, die man damit ja machen würde, wären dann, wenn ich so spontan drüber nachdenke, Sachen, wo wir auf den ersten Punkt zurückkommen, wo ich eigentlich nicht will als Unternehmen, das ich und der Konkurrent zusammen arbeiten. Wo ich mich eigentlich differenzieren will vom Konkurrenten und das sind aber die Einsatzbereiche, wo ich spontan Secure multiparty computation sehe. Also was ich heute machen würde, ist halt so Sachen wie Produktoptimierung Prozessoptimierung für die Automatisierungsprozesse, für Protokollabläufe und für generelle Maintenance, Verfahren zum Beispiel, so Predictive Maintenance sind so Sachen, die bei uns jetzt sehr viel hochkommen. Da kann man nie genug Daten haben. Wenn es dann nicht um Sachen wie Cyber Security oder Security oder solche allgemeingültige Sachen, wo sich jeder zusammen verbessern will geht, sondern um Sachen, wo ich mich abheben will, dann steck ich das Geld wahrscheinlich lieber in eigene Verfahren anstatt Daten zu teilen und quasi zusammen voranzuschreiten.

45 [0:30:33] **I:** OK macht Sinn, äh, noch ein Use Case wäre der rechtliche Aspekt, also manche Daten darf man ja einfach rechtlich nicht teilen oder an andere Unternehmen weitergeben, weil sie personenbezogene Daten sind. Mit SMPC würde man den Aspekt quasi umgehen in der Theorie, weil man die Daten ja nicht wirklich teilt. Vielleicht wär es für sowas sinnvoller.

46 [0:31:70] **B3:** Ja, das hört sich sehr gut an. Ehrlich gesagt, da muss man halt nur beachten, das muss wirklich hundertzehn Prozent wasserfest sein, weil da jeder Benefit, den man daraus ziehen könnte, niemals die Strafen ausgleichen kann. Also damit da ein Unternehmen mitmacht, glaube ich muss das wirklich wissenschaftlich komplett durchgeleuchtet sein und da bin ich nicht vertraut genug mit SMPC, um das zu bestätigen. Wenn es so ist, dann glaube ich, wäre das auf jeden Fall der größte Use Case jetzt so aus dem Bauchgefühl raus, aber wenn das nicht so ist oder da auch nur der geringste Zweifel besteht, dann würde ich das sehr kritisch sehen. Und wäre dann auch wahrscheinlich, dass in so einem Unternehmen dann die Leute zur Cyber Security Abteilung kommen würden und fragen würden: „hey, könnten wir das machen? Ist das abgesegnet von euch?“ Wenn es dann Zweifel gibt, würde ich direkt sagen: „Nein.“ Ein Zweifel würde mir reichen, um das abzulehnen.

47 [0:31:59] **I:** Okay und angenommen, es wird jetzt Use Cases in ihrem Unternehmen geben, wo Secure Multiparty computation angewendet werden könnte, um Daten mit konkurrierten Unternehmen oder allgemeinen Unternehmen zu teilen. Äh, hätten Sie ein besseres Gefühl, dass da keine (.) ,dass die Daten nicht missbraucht werden oder würde das ihr Vertrauen erhöhen

beziehungsweise die Risikoeinschätzung senken von ihnen?

48 [0:32:51] **B3:** Ja, wenn da die Experten von unserem Unternehmen das auch abgesegnet hätten, dann ja. Also, da gibt es auch wieder beim Unternehmen unserer Größe eine extra Abteilung für Risk Management, die beschäftigen sich jeden Tag nur mit solchen Fragen. Und die dann zum Beispiel sagen: „ja, dieses Risiko schätzen wir aufgrund dieser Technologie als akzeptabel oder als gering oder als am besten nicht existent ein.“ Dann wäre das ein riesiger Bonus. Also das Wichtigste ist, dass die internen Experten das bestätigen und sagen ja, wir sehen es genauso, wie das jetzt zum Beispiel das Unternehmen, das das anbietet. Weil wenn mir der Sales Mensch sagt, dann ist mir das egal, also dem Glaube ich kein Wort, das muss schon eine unabhängige dritte Instanz bestätigen am besten sowas wie das BSI oder noch besser halt die internen Experten in Kombination mit dem BSI und am besten auch der Gesetzgeber. Dann dann ja.

49 [0:33:49] **I:** OK OK. Ja gut. Eine Frage hätte ich noch, die ist wieder vielleicht bisschen rückläufig, weil wir davor über das Makroökonomik Thema gesprochen haben gibt's denn auch irgendwelche situativen Umstände oder makroökonomischen Umstände, die data sharing vielleicht verändern?

50 [0:34:05] **B3:** Mhm. Ja, glaub ich auf jeden Fall, also da gibt es einfach Aspekte wie zum Beispiel gewisse Regionen, die von anderen Regionen abhängen und dann sich Unternehmen dieser Region zusammenschließen und sagen: „OK, wir haben da ein gemeinsames Gap, das wir füllen müssen.“ Wie zum Beispiel jetzt aus meiner nicht beruflichen, aber aus der privaten Ansicht, wo man ja immer wieder liest: Die deutschen Autobauer sind in der Produktion extrem hinterher. Und wenn die dann sagen würden OK, wir schließen uns zusammen, machen da irgendwie sicheres Data Sharing, um zusammen voranzugehen, dann wäre es auf jeden Fall glaube ich ein makroökonomischer Aspekt, der Sinn macht und wo das dann Benefit für alle diese Unternehmen bringt. Und solche kann ich mir auch in anderen Bereichen vorstellen, wo man sagt okay, zum Beispiel ein Land hängt uns komplett ab, oder es gibt Sanktionen oder sonst irgendwas, die dazu führen, dass mehrere Unternehmen halt einfach ein gemeinsames Ziel haben. Es den Punkt aushebelt, den ich vorher genannt hab, wo man sich ja eigentlich abheben will, also wenn man sagt: „OK, wir wollen uns gemeinsam abheben“ anstatt „wir wollen uns voneinander abheben.“ Im Vergleich zu einem gemeinsamen Feind sozusagen oder sich einem gemeinsamen Ziel zu verschreiben. Dann, glaube ich, wäre das sehr wichtig und da dann wieder zurück auf das, was ich vorher gemeint habe. Externe Faktoren. Wenn jetzt halt irgendeine Gesetzgebung kommt oder irgendeine andere Sache, die Unternehmen zwingt, irgendwas zu machen, zusammen, dann sehe ich dann einen großen Einsatzpunkt.

51 [0:35:56] **I:** OK, interessant ja. Okay, das war auch schon meine letzte Frage, vielen Dank für das Interview

52 [0:36:12] **B3:** Kein Problem, war ein sehr interessantes Gespräch.

1 [0:00:00.0] Interview B4:

2

| | |
|---|---|
| Interview-Nr. | 4 |
| Date of the interview | august 31, 2024 |
| Duration of the interview | 33:32 min |
| Interviewer | Felix Starnecker (I) |
| Interviewee | B4 (Spain) |
| Role | Lead data scientist |
| Sector | IT-Security (Privacy Enhancing Technologies) |
| Specialities | No specialities |

3 **I:** So, my first question would be like, with whom do companies out of your experience share their data and for what purpose?

4 [0:07:09.5] **B4:** Yeah. Well, at least in my field that we are working right now, it's between publishers and advertisers. So if you are dealing with, let's say, digital advertisement then of course online shops would like to know typically purchase, campaigns, they buy space for their advertisement on publisher. I don't know if you're familiar with this or this topic.

5 [0:07:48.9] [0:07:51.1] **I:** A bit. Yeah.

6 [0:07:53.9] **B4:** Okay, so they buy, of course advertising space. So they pay for it for their advertising of their product. And, you know, these typical banners that are on the sides of your browser or whatever page you're visiting, you have advertisements and then, yes, those are paid by the companies that you see in the advertisements. And of course, companies would like to know how effective were those, you know, they were targeting the right audience, whether you clicked, and you went to the visit of the page or you actually put something in the cart, or you actually even bought something and they would like to know, where are you coming from? So this traditionally is done by third party cookies. So it's a piece of data that tracks you around. Now the usage is way more reduced. So anyway, so in my experience. So the latest change would be in principle between advertisers and publishers. Why? Because the third party cookies are either not working or very few browsers allow it. So? So they are looking for alternative ways to solve this problem of measuring how efficient was the campaign? So a shop, an online shop would like to know, which publishers should I pay for to wear to put my advertisement. Contract agreement? Yes. Of course. Yeah. They make agreements either just for actually starting the campaign, you know, putting the advertiser on a

publisher and of course they have an agreement in terms of, in case they use some alternative tools. They need the agreements for how they can share their data and what measurements.

7 [0:10:01.1] **I:** Would you say that trust is an important factor when companies think about sharing their data with other companies?

8 [0:10:10.5] **B4:** Yeah, I mean, it's a complex topic in the sense that yes, they need to trust each other, but, mostly I think they are afraid of fines, you know, so they want to comply to with the existing regulations, the GDPR and so on. (..)

9 [0:10:30.3] **I:** Okay. So but like that they like maybe that some countries, for example, European companies share their data like preferring to share the data with European companies because they just have a deeper trust.

10 [0:10:51.4] **B4:** In that case, I mean, it's more a commercial point of view. If there is business value in sharing the data, they do, right? Otherwise they don't. And if you want to, if you're in Europe and you want to share, data in the US is more, let's say the legal part is more cumbersome, right? Because in the US suddenly each state has its own regulation and you need to really make they do it. I mean, it's not that, but it's just less natural.

11 [0:11:31.8] **I:** Okay. And would you say when companies think about sharing data that it's more subjective or objective process like is it one person who just decides like a top manager or, I don't know. (...)

12 [0:11:52.7] **B4:** (..) That's kind of a philosophical question. Yeah, I would say, yeah. You reach out typically to who you can reach out and depends yes whether you convince them or not. So it's I would say subjective. Yes. Okay. It takes a decision, if they are in the position to take the decision. Yes.

13 [0:12:23.0] **I:** And it's more a feeling, about sharing data then.

14 [0:12:28.4] **B4:** Yeah. I mean, if I understand correctly your question is whether it's obvious that you get some value in sharing data or is a part of.

15 [0:12:51.9] **I:** Actually the question is more about if it's a subjective or objective process because I have the feeling that often maybe data sharing is seen as a too high risk because many companies just don't share their data as they're just afraid or overestimate the risk sometimes. And then it would be a bit more subjective.

16 [0:13:21.9] **B4:** Yeah, I would say that the default would be just in case not share, you know and it depends on the company. There are some companies, they are actually more open minded if you want. Because when you approach companies, publishers, advertisers, whoever, at least in the marketing area, they need to have also some technical understanding of what you are trying to do, especially if you start talking about secure multi-party computation. Whatever technique. Right. So, they see it as some sort of magic. And you need to be convincing in a sense, honest and talk in an understandable language. They have to specific subjects that actually take the decision and they push your (..) so what we are working on, we produce some pieces of software. Probably similar to the company you are dealing with. And that you want to sell it to clients.

17 [0:14:47.6] **I:** Yes. And what would you see as the biggest risk when companies share their data? (..)

18 [0:14:55.1] **B4:** The biggest risks.

19 [0:14:56.3] **I:** Yeah.

20 [0:14:56.8] **B4:** Like well here I would say it depends on what you are sharing because there are many techniques. Right. So if you use secure multi-party computation, you are probably sharing only a specific part of information needed only to compute whatever you agree to compute. So the risk is lower. If you say, okay, let's put all together everything in this trusted third party, a clean room, and then you trust the clean room and everything is there.

21 [0:15:47.2] **I:** Yeah. Of course.

22 [0:15:47.8] **B4:** So anything can happen. I mean, you have to trust. Suddenly there's a third party in between and an accident of, like, a data leak or some attack can happen. So I would say that depends on what? What you mean by data. If the data is the whole data set with everything that. But you might only need one column of that. If you think about a data frame but reduce the amount of data that you share.

23 [0:16:23.6] **I:** Yes. And what would you see as external or internal influences that encourage companies to share data. Like, I don't know, is the management for companies important that as an internal factor, for example, company share the data or I don't know, like external factors could be as an example, maybe, regulations or something like this. (..)

24 [0:16:58.9] **B4:** Yeah. Okay. Typically they like looking forward to some benefit, economical benefit, meaning either deeper insights, whether, by sharing data in, let's say network with collaborators, competitors even they are guaranteed to have a better targeting, better analysis of their

data or some kind of amplification of the information they have. Because especially smaller companies, right? I'm not talking about Facebook. Facebook doesn't need all of this because they have already half of the world. Yeah. Or Google. Right. Or Apple. But, smaller companies that might have just not significant amount of data. So that would be benefit. Okay. I share your data, but I have some benefit in the consortium. Yeah.

25 [0:18:15.0] **I:** So especially the benefits are important for companies to share the data. Yeah. And now in the second part, regarding privacy enhancing technology and special focus on secure multi-party computation, for which use cases would you see secure multi-party computation? Or would you say it's like a useful method or is it only in theory?

26 [0:18:58.3] **B4:** Well, in my experience from the theory point of view these techniques have been used in the last decades for academic purposes, mostly, but lately there are more commercial applications. In our case, we are applying them to marketing. So marketing where (..) There are two fronts, right? There's the front of being compliant with legal regulations. So you want to protect user privacy. You are a company and you have users of your web page or of your product, and you collect their data and you need to be compliant and protect their data. And then there is the need for protecting your competitive advantage. So not disclosing too much information to a potential collaborator that can be also a competitor. Right? So I cannot tell you in general in life where these techniques are but (..) In the last, let's say five years, I saw more and more application of these techniques. But in my field it is marketing. So digital advertisement. And that's definitely useful you can provide because what you want to substitute what the third party cookie was doing and the third party was, freely, happily linking people all around, the internet. And now it's over. Mostly not completely, but so secure Multi-Party computation is a way to enable those techniques, but providing users still privacy protection. So not disclosing, let's say, who actually visited your page, but how many people from your campaign visited your page. Of course, you don't need to know names or name and where are they from. But more accurate aggregated analytics, for example, metrics that still are useful to the market, but, yeah, but I'll provide the details of single individuals.

27 [0:21:52.0] **I:** And do you think that like in the future privacy enhancing technologies will change the data sharing behavior of companies in a big way?

28 [0:22:03.8] [0:22:09.0] **B4:** I hope so. You know, at least in my experience, there's like, last three, five years even, you see this popping up of weekly papers on new, distributed, computation technique among which secure multi-party computation. But you can talk about federated learning. So an explosion of in research and in, I don't know, company providing these services. So everybody's talking about an escalation of adoption of these techniques in the market. However, it

doesn't happen yet, so. Yeah.

29 [0:22:54.7] **I:** So what are the reasons for this?

30 [0:22:58.2] **B4:** The reason for this, I would say that there are some, there's a great variability in use cases, because what I described to you earlier is one specific case of measurement of campaigns. But then there are so many other use cases, for example, targeting, analytics in general, just in marketing, a lot of variability. Then healthcare is another sector then. I don't know. But why is another public sector so, so there's a lot big range of applications. And it's difficult to find the silver bullet. Even software product that solves all the issues you have, all this customization that you need to perform. And I didn't see at least dominating solution. I'm talking about the software in the market yet. So it might be in the future. It will.

31 [0:24:02.5] **I:** But why? I didn't understand why there are so many use cases for these kind of technologies, but why is it not? Because actually most of the people don't know something about stuff like secure multiparty computation. And it's not much used in industry yet. So what would you say is the reason for actually there are use cases and it's a nice technology, but nobody really uses it.

32 [0:24:35.0] **B4:** Because it's let's say it's a longer process and to adopt an available tool to solve your specific problem if you're a healthcare company or an advertiser and you also need to be open minded. And so people I feel they are just looking for the simpler solution is the hassle. Probably don't understand, as you say. And they don't see a clear solution. Like there's not a, I don't know, a Google tool or a Facebook tool. That solvent is well established and everybody uses start using to solve these kind of problems. So it's in between technical difficulty and lack of trust.

33 [0:25:31.0] **I:** Yeah. And yeah that would be the next question. Do the influencing factors of like, is trust still needed when, companies share their data with secure multi-party computation, for example, because, yeah, I don't know, people who offer secure multi-party computation say that you only have to trust the technology anymore. But in some ways, maybe there is still trust needed. I don't know.

34 [0:26:04.1] **B4:** But, if you apply it properly, yes, you can trust the technology, in my opinion. Yeah, but it is a concept that is probably difficult to explain. People think it's just magical, right?

35 [0:26:22.4] **I:** So they don't trust.

36 [0:26:24.2] **B4:** But it's cryptography is a cryptography technique. So you don't need to, there is this distinction between soft privacy and hard Privacy. Soft privacy for me is a signing, agreements and

non-disclosures and trusting the other party if you share data with somebody. Right. And our privacy is, you know, you just remove the possibility for the other party to, I don't know, cheat or lose the data, or it doesn't have to be cheating. You know, you can just, I don't know, lose the data because somebody else, attacked the data set and stole your data. So it reduces the risks. Yeah, but as you said, like still people don't trust the people enough who offer such new technologies because it's definitely because you need to trust the company that produces this piece of software, for example. Right. And like, Look. For what companies or what conditions are needed for companies to invest in privacy enhancing technologies, or better, which type of companies are more willing to invest in such technologies from like from the sector point, you already said in marketing. Well, here, yes. But big companies, right. Because it's a huge investment in development in technical skills. So that anybody there's at least in my experience, very few people with this profile in.

37 [0:28:14.3] **I:** Then why is it such a big investment?

38 [0:28:19.0] **B4:** Because you need, I don't know, a team of engineers developing a software tool during months. And it's basically just research so far. I mean development, research and development. You have some open source tools available but are still not production ready. And it's still very niche if you want.

39 [0:28:45.8] **I:** So and are there some like new risks you see who come up with secure multi-party computation like or how are the risks changing maybe? (..) So risks for how do PETs change the risk of data change.

40 [0:29:09.1] **B4:** Well they reduce the risk right. In my opinion.

41 [0:29:13.0] **I:** But are there maybe also some other like some new risks coming from PETs?

42 [0:29:19.0] **B4:** Well that if I understand the question correctly that you never know because if you're talking about secure multi-party computation there's, you know, several protocols that in the years have been found some on from their abilities afterwards. After some years publishing so that can happen. Quantum computing is yet another technology that might, you know, oblige us to rethink many of these protocols or security assumptions that we are using currently. So, but that, that you never know.

43 [0:30:12.0] **I:** Okay. And, what do you see like last question, what do you see as the biggest challenge for the acceptance of privacy enhancing technologies on the market? (..)

44 [0:30:42.5] **B4:** I don't know. I would say that, establishment of, so if, let's say, well, already

established company would adopt such techniques, it would make it easier for the rest of the market to adopt. It's a matter of trust, I would say, and also the availability of the tools, technical tools. So there are some dominant solutions that are production ready. And we're not there yet. So right now I see only prototypes or some side projects of Facebook or Google. Yeah. So it's a matter of I would say hopefully time and trust.

[0:31:37.3] **I:** Yeah, hopefully. Yeah. Okay. That was it. Thank you very much for your time.

[0:31:48.2] **B4:** You're welcome. Hope it was useful.

[0:31:50.1] **I:** Yeah, it was very useful.

[0:31:53.1] **B4:** And good luck for your thesis.

[0:31:56.7] **I:** Thank you.

1 [0:00:00.0] Interview B5:

2

| | |
|---|---|
| Interview-Nr. | 5 |
| Date of the interview | September 19, 2024 |
| Duration of the interview | 47:16 min |
| Interviewer | Felix Starnecker (I) |
| Interviewee | B5 (Germany) |
| Role | Data Space Expert |
| Sector | IT-sector |
| Specialties | Interview in german |

3 [0:07:03.4] **I:** Okay, gut, dann wäre meine erste Frage aus deiner Erfahrung. Mit wem tauschen Unternehmen denn Daten aus? Sind es Wettbewerber, Zulieferer oder Kunden? Und wozu? Also Hauptzweck.

4 [0:07:22.8] **B5:** Also ja, das ist schon eine große Frage, weil da gibt es eine Menge Antworten drauf. Also was wir sehen, was quasi eigentlich der größte Treiber aktuell oder naja, sagen wir vor zwei Jahren dazu ist, dass man Daten austauscht ist, dass es für das Lieferkettengesetz vorgegeben ist, dass man Daten austauscht und zusammenführt. Also nach dem Lieferkettengesetz müssen wir beispielsweise über die ganze Lieferkette beispielsweise bei einem Automobil quasi einmal die CO2 Werte zusammenrechnen und brauchen da eine gewisse Transparenz. Und genau dafür müssen eben diese Daten ausgetauscht werden, weil es muss von dem Rohstofflieferanten, der am Anfang vielleicht tatsächlich Eisenerz aus der Erde gräbt und was dann wieder in der Raffinerie kommt und bla bla und dann machen wir daraus einen Kolben und dann kommt der in den Motor und dann kommt der Motor in das Auto und dann müssen wir diesen am Ende letztendlich berechnen können. Und das war ein ganz großer Treiber für ein ganz großes Datenprojekt. AGs zum Beispiel, bei dem es eigentlich um die Lieferketten in der Automobilindustrie geht. Getrieben dadurch, dass sie für diese Lieferketten Transparenz schaffen müssen. Das ist jetzt der vordergründige Use Case, mit dem das erstmal vorangetrieben wird, weil sie müssen halt erstmal Compliance in diesem Gesetz werden. Allerdings kann man dann natürlich ganz viele andere Sachen machen.

5 [0:08:50.2] **I:** Aber mit dem Gesetz, das ist dann jetzt eine ganz neue Entwicklung, oder? Weil das gibt es ja noch nicht so lange oder?

6 [0:08:58.4] **B5:** Also ich weiß jetzt nicht, wann das Lieferkettengesetz in Kraft getreten ist. Ja, aber

das ist eine relativ neue Sache. Ich meine, letztendlich haben die Unternehmen das natürlich schon auf dem Schirm, bevor das Gesetz final abgeschlossen wird und tatsächlich in Kraft tritt. Also das heißt in dem Fall, um an deiner Frage zu bleiben, tauschen sie tatsächlich quasi Daten mit, vor allem in dem Fall, Lieferanten und Kunden aus. Also in dem Fall ist es vor allem so, dass es quasi nach oben in der Lieferkette geht bis zu den OEMs und die das dann quasi alles zusammenrechnen. Das ist allerdings auch nicht der einzige Fall. Ich meine, andere Use Case, den wir auch so mehr industriell haben, ist auch wieder eher in der Lieferkette, dass wenn wir Bauteile haben, die eine bestimmte Länge haben und dann zum Beispiel zwei Millimeter kürzer sind, was innerhalb der Richtwerte ist, wie lang es sein muss. Aber wenn wir sehr genau sein wollen, dann müssen wir vielleicht ein anderes Bauteil machen, das auch zwei Millimeter zu kurz ist, damit unser Ding zu 100 % passt oder zu 99,9 % und nicht nur zu 99,2 %. Also da können wir, müssen wir auch irgendwelche Daten dafür zusammenführen. (..) Und dann gibt es Gesundheitsprojekte, in denen tauscht man quasi Daten mit allen möglichen aus. Also das ist sehr vielfältig.

7 [0:10:33.7] **I:** Ich hätte tatsächlich eine kurze Frage. Das bezieht sich gar nicht so direkt auf meine Masterarbeit, nur weil ich eben auch die Masterarbeit zusammen mit einem Startup schreibe, die auch in dem Bereich arbeiten. Die sind eben Anbieter von Secure Multi Party Computation. Und da war jetzt, da schauen wir gerade, ob das sinnvoll wäre mit Secure Multi Party Computation wegen diesem Lieferkettengesetz, dass vielleicht dadurch auch Konkurrenten ihre Daten miteinander austauschen, die vielleicht denselben Zulieferer haben und dann damit quasi ein bisschen weniger Aufwand haben, um diese ganzen rechtlichen Anforderungen zu erfüllen. Würdest du das als sinnvoll sehen? (..)

8 [0:11:20.5] **B5:** Das ist eine gute Frage. (..) Das darfst du jetzt nicht zitieren, weil letztendlich ist es aber so, dass in  Container für die OEMs eigentlich auch ein willkommenes Mittel ist, um ihre ganzen Zulieferer. (..) Das müssen wir jetzt politisch korrekt ausdrücken. (..) Noch besser zu kontrollieren und bessere Daten über sie zu haben und den Wettbewerb noch besser anzustacheln. Wenn das heißt, es könnte sein, dass diese Anwendungsidee von dem Startup quasi nicht gut dem entspricht, was die OEMs eigentlich wollen, auch wenn sie es nicht offiziell so sagen. Aber. (..) Das weiß ich natürlich auch nicht, weil ich habe da nicht mit irgendeinem CEO von VW darüber gesprochen, sondern man bekommt so den Eindruck.

9 [0:12:15.9] **I:** Okay, okay. Und würdest du Vertrauen als wichtigen Faktor sehen, wenn Unternehmen ihre Daten austauschen?

10 [0:12:25.2] **B5:** Das ist eigentlich der wichtigste Faktor. Also das größte Ding. Die größte Diskrepanz zwischen dem, wo man Daten teilen könnte und daraus quasi Werte generieren könnte

und dem, wo es auch passiert, haben wir im Mittelstand bzw in kleinen mittelständischen Unternehmen bei den großen auch. Da ändert sich jetzt vielleicht ein bisschen mehr. Aber wenn wir jetzt gerade zum Beispiel den mittelständischen Maschinenbauer genauer anschauen. Der ist zum Teil vielleicht noch gar nicht in der Cloud, weil er vollkommen zu Recht Angst hat, dass ihm seine Daten da abhandenkommen, weil es natürlich auf chinesischer Cloud manchmal recht offensichtlich passiert. Aber auch auf amerikanischen Cloudanbietern gibt es Gesetze wie den US Cloud Act, die dafür sorgen, dass wenn die NSA diese Daten will, sie die bekommt. Wenn das jetzt irgendwie wichtig genug wäre, dann ist natürlich die Frage, wann wäre es vielleicht wichtig genug? Aber mit der aktuellen politischen Situation in den USA wäre da ja einiges denkbar, dass man seine Marktmacht da dann auch politisch ausspielt. Das heißt also, es gibt da dann schon kein Vertrauen in cloudbasierte Anwendungen und dann noch weniger oder ähnlich wenig. Ich weiß jetzt nicht, wie Sie es nicht werten darin, die Daten mit anderen Leuten auszutauschen. Da gibt es viel Angst, sodass man daraus potenziell Geschäftsgeheimnisse so reverse engineeren kann letztendlich. Oder dass man das könnte. Weswegen? Deswegen haben wir ja Privacy Enhancing Technologies. Viel Angst, dass andere irgendwas rausfinden könnten. Und deswegen wird lieber mal nichts gemacht. Auch bei den Großen zum Beispiel. DSGVO ist also so eine Sache, in der auch in der deutschen Auslegung besonders bei den großen Unternehmen lieber mal nichts gemacht wird, weil das auch viel an den Corporate Structures liegt, weil dann die Rechtsabteilungen sagen Na gut, also auf der sicheren Seite sind wir, wenn wir einfach nichts teilen, wenn wir nichts machen. Und was man dann natürlich daraus dann quasi für naja Vorteile haben kann, wird dann da quasi nicht gut genug gegengerechnet. Sagen wir so, aber letztendlich ja. Vertrauen ist der größte Faktor. Fehlendes Vertrauen. Kontrolle hängt davon natürlich ab. Also wir kommen da jetzt ja dann gleich drauf, aber letztendlich gibt es dafür zwei verschiedene Ansätze. Wir können quasi Akteure haben, denen man vertrauen kann. Also dass man sagt okay, wir haben hier jemanden, der ist nicht kommerziell, nicht gewinnorientiert oder ähnliches. Die sind im Austausch mit Behörden oder ähnlichem, die sind so was, was wir vielleicht einen Datentreuhänder nennen. Und bei denen geben wir das und bei denen darf das stattfinden und die haben vielleicht noch einen besonders guten vielleicht noch einen technologischen Ansatz, dass sie quasi selber in die Daten nicht reinschauen können oder irgendwas in der Art. Letztendlich müssen wir aber jemandem vertrauen. Oder wir nehmen halt quasi Zero Trust Solutions, mit denen wir sagen okay, wir brauchen gar kein Vertrauen, sondern es lässt sich alles technologisch regeln, dass wir da gar keine Probleme haben. Was dann? Worauf manche Privacy Enhancing Technologies auch rauslaufen.

11 [0:15:36.8] **I:** Ja okay, dazu hätte ich noch eine Frage, aber die Frage stelle ich erst später. Würdest du sagen, dass Entscheidungen von Unternehmen Daten auszutauschen subjektive oder objektive Prozesses oder auch beides sind?

12 [0:15:53.3] **B5:** Naja, also erstmal beides natürlich. (...) Ich würde sagen, dass man, was ich gerade beschrieben habe, quasi auch mit der DSGVO, das viel gemrkt hat. Da kam dann quasi auf, dass man bei großen Verstößen bis zu 2 % des Jahresumsatzes zahlen müsste, was jetzt glaube ich noch nicht passiert ist und schon gar keine Mittelständler. Allerdings ist da viel Angst im Raum, dass man sich, naja, das Geschäft ruiniert, indem man da irgendwelche Sachen ein bisschen zu leichtfertig betrachtet. Deswegen hat man lieber nichts gemacht. Das heißt, ich glaube, man ist in einem sehr auch immer noch in einem sehr subjektiven Modus, dass man lieber erstmal ablehnend ist und sehr, sehr vorsichtig. Und ich glaube, dass das jetzt so langsam, quasi ein bisschen objektiver aufgearbeitet wird, dass man das halt an vielen Stellen immer mal wieder jemand vorstellt. Ja, aber wenn wir das mit denen zusammenbringen könnten, dann könnten wir hier unsere Lieferkette optimieren. Oder wenn wir dieses und jenes zusammenbringen würden. Das wäre ja auch nicht schlecht, oder? Diese Daten haben wir wirklich nur rumliegen, die könnten damit nix machen. Das wäre natürlich eigentlich eine gute Sache und uns schadet das nicht. Ja, also ich würde naja, ich würde hoffen blöd gesagt, dass wir uns eigentlich gerade jetzt in der Wende dazu bewegen, dass es immer immer objektiver wird, weil so aus dem Subjektiven raus ist macht man es vielleicht eher gar nicht.

13 [0:17:34.5] **I:** Also du würdest sagen, dass das Risiko eigentlich immer noch überschätzt wird. Zurzeit noch im Datenaustausch, zumindest in Deutschland. Wahrscheinlich.

14 [0:17:43.6] **B5:** Ja, würde ich so sagen. Es ist auch so, dass die Risikobewertung ist. Wir bewegen uns jetzt langsam auch mit neuen EU Gesetzen. Vielleicht auch da noch wichtig erwähnenswert der DataAccess, der letztendlich, sobald er also ist in Kraft ist, aber dafür sorgen wird, dass auch Unternehmen Daten teilen müssen, zum Beispiel für Katastrophenfälle oder ähnliches, was wahrscheinlich für mehr relevant ist, als man im ersten Moment glaubt. Aber das ist nochmal ein ganz eigenes Thema. Das ist ein Akt, der jetzt erst in Kraft ist, glaube ich. Der ist in Kraft, noch nicht so lang, aber der ist in Kraft. Es ist aber auch so, wenn der in Kraft ist, dann wird er auch noch nicht durchgesetzt. So richtig. Das heißt, es ist jetzt alles gerade so in der Mache, da wird sich gerade langsam angepasst usw.

15 [0:18:34.6] **I:** Ist es dann so was, wo dann der Staat sagen kann, wenn er irgendwelche Daten braucht, dann muss das Unternehmen die quasi rausgeben.

16 [0:18:42.8] **B5:** Ist einfach gesagt. Ja, also das ist jetzt auch nicht einfach nur willkürlich. Und irgendjemand hat sich jetzt gedacht okay, wäre ja interessant, wenn wir das hätten. Ihr müsst uns das jetzt geben. Aber man muss im Prinzip halt auch eine Infrastruktur schaffen, dass man im Katastrophenfall, Überflutungen und Ähnliches usw haben wir jetzt immer öfter seine Daten auch

zusammen bekommen kann, dass man da für Prävention und für Bekämpfung Sachen mitmachen kann.

17 [0:19:15.2] **I:** Ja, ja, so was gibt es ja glaube ich in Amerika auch schon, oder? Deswegen ist es ja auch so, ich dachte, das hätte irgendjemand auch schon gesagt in einem Interview, dass deswegen das ja auch so kritisch ist, wenn man da Daten irgendwie an Microsoft oder so gibt, weil der Staat, also Amerika, eigentlich immer diese Daten anfordern kann im Krisenfall oder so.

18 [0:19:38.6] **B5:** Das ist das, was ich vorhin meinte mit dem US Cloud Act, Das ist genau das aber. Also das ist dann eben die Definition von Krisenfall, die mit immer weiter ausgehebelter Gewaltenteilung in Amerika vielleicht noch ein bisschen kritischer ist als bei uns ja.

19 [0:19:55.5] **I:** Und welche Hauptrisiken siehst du im Datenaustausch?

20 [0:20:02.3] **B5:** Letztendlich gibt es viele Risiken von Datenaustausch, gerade für Unternehmen. Ich meine, wenn man zum Beispiel seine Buchführung nicht richtig macht, dann kann einem jemand dabei auf die Schliche kommen, wenn man seine Daten teilt. Das wäre so eine Sache. Theoretisch kann man natürlich auch seine Geschäftsgeheimnisse aus Versehen weitergeben oder auf eine Art und Weise weitergeben, wie man sie tatsächlich reverse engineering kann. Das ist natürlich auch nicht wünschenswert. Da ist man allerdings auch sehr vorsichtig und verständlicherweise sehr vorsichtig. Der Großteil der Daten, die so generiert werden im allgemeinen Geschäftsbetrieb, sind allerdings ziemlich unbedenklich, denke ich. Also ich denke, die Risiken werden überschätzt. Ja, aber sie sind natürlich da. Also es ist schon sinnvoll, da auch drüber nachzudenken, gerade wenn man halt, naja, das alleine verantwortet oder wenn man halt ein relativ kleines Unternehmen vielleicht hat, da kann einem schon auch was passieren. Ja okay.

21 [0:21:11.0] **I:** Und Voraussetzungen für einen erfolgreichen Datenaustausch? Braucht es da welche? Oder kann theoretisch jedes Unternehmen einfach anfangen Daten zu teilen?

22 [0:21:21.0] **B5:** Na ja, also das ist auch immer die Sache. Also ich sage den Leuten immer, rein theoretisch brauchen wir dieses Ganze gar nicht. Wenn wir wollen, können wir auch uns PDFs per Email schicken, sondern das können alle jetzt machen, wenn sie möchten. Also wenn ich jetzt ein Maschinenbauunternehmen habe, kann ich jemanden treffen auf irgendeiner Konferenz und dann sagen wir Hey, ich gebe dir meine Datent, ich kriege dafür welche von dir. Oder du gibst mir 100 € dafür oder irgendwas. Und dann gebe ich ihnen einen USB Stick. Das ist alles theoretisch natürlich möglich. Woran es aber eigentlich hapert, weil das so kaum jemand macht, ist einerseits, dass es einfach ist, dass es skalierbar ist, dass es rechtssicher ist. Das ist noch so eine Sache.

Rechtssicherheit ist eigentlich eher eine Reduzierung des Rechtsrisikos. Also wir haben aktuell zum Beispiel noch wenig Verträge, die schon mal vor Gericht irgendwie dann tatsächlich beklagt wurden und quasi irgendwas standhalten mussten, ob die so auch gelten. Weil rein prinzipiell kann jemand natürlich meine Daten veruntreuen und dann könnte ich die Leute, wenn wir einen Vertrag darüber haben, dafür verklagen. Wie das Ganze dann wirklich gut funktioniert, so effektiv gibt es jetzt noch nicht so viele Erfahrungsberichte drüber. Das heißt aber so was wie Standardverträge, auf die man sich verlassen kann, dass es rechtlich passt. Wir brauchen erstmal naja, eigentlich einen guten Grund, warum wir das überhaupt machen wollen. Also irgendeine Art von Businessmodel. Im besten Fall oder zumindest für diese eine Transaktion, irgendein Incentive. Also ich kriege irgendwas zurück, irgendjemand kann was damit machen. Vielleicht ist es dann altruistisch. Also ich mach das tatsächlich ohne Gegenleistung. Einfach weil ich glaube, dass es für eine gute Sache ist. Oder halt natürlich einfach ein finanzieller Gegenwert. Und all diese Sachen. So in der rechtlichen, der ökonomischen, auch in der technischen Infrastruktur quasi, dass das halt möglichst einfach und sicher und so funktioniert und skalierbar funktioniert. Das ist das, woran es auch ganz krass scheitert. Also das ist alles noch nicht so groß skaliert. Da gibt es jetzt schon ganz, ganz viele Sachen. Also wir sind hier ja da auch schon eine Weile dran.

23 [0:23:32.1] **I:** Ich hätte auch an so eine Datenqualität oder so gedacht oder halt einfach an Standardisierung, technische Standardisierung.

24 [0:23:38.9] **B5:** Genau das ist dann noch ein weiteres Thema. Das ist dann quasi mehr so in der speziellen Domäne. Aber da gibt es auch schon Unternehmen, die quasi zumindest versuchen zu entwickeln, dass sie quasi ein Siegel dafür haben, dass diese Daten bestimmte Qualität haben. Oder wir. An vielen Orten kann man sich auf ein Standardformat einigen, einfach auf ein offenes Dateiformat. Dann sind die Daten schon mal zumindest in einem ja offenen, standardisierten Dateiformat. Welche Qualität die dann haben, das muss dann quasi noch mal jemand extern prüfen. Das wäre dann eigentlich auch ein Dienst, für den man Geld bezahlen müsste in den meisten Fällen, damit die dann sagen okay, hier setzen wir unser Siegel drauf. Hier sind nicht nur Nullen und Einsen drin, sondern das sind Daten, mit denen kann man was anfangen, nach welchen Kriterien auch immer.

25 [0:24:22.6] **I:** Und was ist mit externen und internen Einflüssen? Also würden dir welche einfallen, die Unternehmen dabei beeinflussen, wenn sie ihre Daten teilen? Also manche hast du ja schon gesagt wie rechtliche Sachen.

26 [0:24:41.0] **B5:** Ja genau, also der rechtliche Rahmen sowohl auf EU Ebene als auch auf nationaler bzw in Deutschland dann auch noch auf der Bundeslandebene an vielen Stellen. Dann haben wir wie

gesagt die Businessmodels natürlich. Was sind die überhaupt? Womit verdient dieses Unternehmen Geld? Was sind die Geschäftsgeheimnisse? Was will man davon oder was will man vielleicht nicht rausgeben? Womit möchte man vielleicht Geld verdienen? Wen möchte man mit sonst was unterstützen? Ist man bereit für mehr digitalen Datenaustausch und damit vielleicht irgendwas zu gewinnen? Oder ist man das nicht? Also wie offen ist man für solche Sachen? Wie digitalisiert ist man generell schon? Also sowas ist natürlich wesentlich einfacher, wenn wir schon eine hohes Maß an digitaler Infrastruktur haben und nicht noch manche Sachen quasi händisch auf Papier aufgeschrieben werden und irgendwo rübergereicht. Dann natürlich, wenn wir zum Beispiel Maschinen haben. Und das muss nicht nur in Fabriken so sein oder in oder sonstigen Manufakturen, sondern auch in einem Krankenhaus, ist das ein großer Fall. Naja, wenn die 20, 30 Jahre alt sind, dann hat man Glück. Wenn die eine serielle Schnittstelle haben, dann braucht man dafür erstmal eine extra Software, die das ganze in irgendwas modernes quasi umrechnen kann, mit dem man arbeiten kann. Dann natürlich die Entscheidungsstrukturen, intern, in so einem Unternehmen. Vielleicht ist, wenn das nur eine Person ist, die alles entscheidet, na ja, dann muss die von irgendjemandem überzeugt werden. Wenn es sowas wie einen Aufsichtsrat gibt oder eine erweiterte Geschäftsführung, dann sieht das natürlich alles ganz anders aus, von der Governance her bzw gibt es vielleicht auch eine Abteilungsleitung, die sowas eigenmächtig entscheiden kann, dass man dieses und jenes jetzt mal anfängt. Also das hängt natürlich auch alles mit drin. Jetzt haben wir rechtlich, technisch, ja technisch so mehr oder weniger, digital oder wie digital man bereits ist.

27 [0:26:37.2] **I:** Data Governance dann auch quasi oder wie das Management eingestellt ist. Und würdest du sagen, der Sektor spielt eine Rolle? Also ob das jetzt ich weiß nicht, Autoindustrie ist hoher Wettbewerb, ob so was eine Rolle spielt oder wo es viele Regulierungen gibt.

28 [0:26:57.5] **B5:** Also generell habe ich jetzt viele Beispiele mit der Automobilindustrie und mit Maschinen gemacht, weil es ist alles noch wesentlich einfacher, als wenn wir jetzt über Gesundheitsdaten reden oder Finanzdaten. Die sind zu Recht hoch reguliert. Jetzt gerade im Gesundheitsbereich ist das Ganze auch viel komplizierter. Also da haben wir weniger diesen freien Markt von die einen produzieren was, geben es den anderen kriegen Geld dafür und dann am Ende kommt ein Produkt raus, das man dann an Konsumenten verkauft, sondern da haben wir dann ja eigentlich Krankenkassen, die für bestimmte Leistungen, die irgendwo festgelegt werden, dann Geld an andere auszahlen nach bestimmten Schlüsseln. Und woher kommt eigentlich der Incentive, Digitaler zu werden. Was haben wir davon, wenn wir diese Daten zusammenkriegen? Wir haben ganz viel davon. (..) Das ist ganz klar. Wenn wir das irgendwie ein bisschen gesammelt kriegen, dann können wir im größeren Stil mal tatsächlich an so was wie Demenz forschen. Das ist nämlich super multifaktoriell. Und dann kriegt man normalerweise nicht genug Daten zusammen, dass es wirklich viel bringt. Ja, aber wie kann man das Ganze angehen? Das heißt ja, der Sektor ist ein sehr,

sehr großer Bestandteil. Ich habe das jetzt ein bisschen unterschlagen, weil aus der Perspektive von dem einen Unternehmen, das ist halt nun mal in seinem Sektor, aber gerade auf der Skala zwischen sehr personenbezogen und gar nicht personenbezogen, da ist es sehr, sehr unterschiedlich.

29 [0:28:21.4] **I:** Okay, und jetzt zum zweiten Teil: Für welche Anwendungsfälle würdest du jetzt Privacy Enhancing Technologien wie zum Beispiel Secure Multiparty Computation am sinnvollsten sehen? Oder glaubst du die Technologie ist überhaupt praktikabel anwendbar oder zu aufwendig in der Umsetzung?

30 [0:28:49.4] **B5:** Ich habe da eng mit einem Projekt gearbeitet das naja, also das hat das Ganze jetzt so nicht dezentral gemacht, aber wir bewegen uns glaube ich in ein ähnliches Feld von Anwendungsfällen. Das Projekt heißt Eurodad, also Eurodad europäischer Daten, Treuhänder euronat.org. Da geht es ging es zunächst erstmal um Finanzdaten, beispielsweise weil wir um da einen großen Anwendungsfall, wo wir ein großes Problem haben, zu beschreiben, wenn wir Geldwäsche verfolgen möchten, wenn Leute an verschiedenen Banken quasi ein Konto haben und es dann immer wieder weiterschicken und irgendwann verliert sich die Spur. Das ist sehr schwer nachzuverfolgen, wenn man nicht quasi über alle Kontobewegungen, die existieren, draufschauen kann. Wir wollen allerdings nicht, dass irgendjemand all diese Daten hat, weil damit würde eine viel zu große Macht kommen. Das wäre nicht gut. Das heißt, die haben quasi eine Lösung entwickelt, die nach dem, was ich gerade vorhin beschrieben habe, eigentlich nach, wenn man so will Federated Learning Computer Data Bauende, sagen wir mal eine algorithmische Blackbox, in der ist ein Algorithmus drin und die schicken wir zu Bank eins und die dann, der nimmt quasi die Rohdaten auf, bearbeitet die, dann schicken wir den weiter zu Bank zwei, aber ohne die Rohdaten mitzunehmen, sondern nur quasi mit dem Eindruck, den wir da gewonnen haben, mit dem Ergebnis, dann machen wir das da immer so weiter und dann gehen wir alle Banken durch und am Ende spuckt quasi nur ein PDF raus. Mit diesen 15 Konten sollte man sich mal genauer anschauen beispielsweise. Das ist jetzt die Theorie. Das hat am Ende nicht 100 % funktioniert, weil naja, weil größtenteils zum Beispiel die Banken Angst davor haben, dass sie das am Ende doch nicht dürfen. Auch wenn die BaFin, die Finanzaufsicht, sagt, dass das okay wäre. Für diese algorithmische Blackbox brauchen wir einen ganzen Haufen an Privacy Enhancing Technologien. Damit auch der Datentreuhänder selbst zum Beispiel nicht reinschauen kann, was denn, wo denn die Rohdaten sind. Also er darf unter gar keinen Umständen jemals irgendwie an diese Rohdaten rankommen können.

31 [0:31:07.4] **I:** Aber seid ihr an die Rohdaten rangekommen als Partei quasi, die den Algorithmus weitergegeben hat, oder?

32 [0:31:15.7] **B5:** Also wir gar nicht. Die Barriere eigentlich am Ende auch nicht. Man muss auch

sagen quasi, wenn wir diesen diese stellt quasi ein, dass es am Ende ein kubernetes Cluster ist. Das ist so etwas wie eine virtuelle Maschine bei einem Computer. Und da geben wir die Daten ein und dann kommt das PDF raus und dann machen wir das platt und es kann niemand reinschauen, auch wir selbst nicht. Also auch nicht die, die es quasi produziert haben. Man gibt es dann aus der Hand und im besten Fall kommt eigentlich der Algorithmus, der das Ganze da drin arbeitet und das Ganze bearbeitet noch von jemand anderem. Also wir können uns das vielleicht in einem anderen Beispiel auch so vorstellen: Wenn ich jetzt herausfinden möchte, ob für meine Schriftproduktion Maschine A oder Maschine B besser ist, die ich mir am Ende angeschaut habe, dann könnte ich, wenn wir Datenräume haben, quasi 100 Hersteller, die mit Maschine A arbeiten anschreiben und 100, die mit Maschine B anschreiben. Dann geben die mir jeweils für von mir aus 0,10 € ihre Maschinendaten. Die kommen dann alle in diesen Datentreuhänder, den bezahle ich auch dafür. Dann bezahle ich jemanden dafür, dass die mir einen Algorithmus geben und einen Algorithmus haben, der quasi gut diese Maschinen miteinander vergleichen kann über diese Daten drüber hinweg. Und dann arbeitet der Algorithmus mit diesen Daten durch und gibt mir am Ende mein PDF, auf dem verschiedene Graphen sind, welche Maschine wie viel besser ist. Und ich lasse mich das vielleicht durchaus um irgendeine Zahl zu sagen. Also das ist mit den monetären Gegenwert, das ist alles nicht so klar, aber vielleicht lasse ich mir das 100 € kosten und tausche die Teile dann quasi zwischen allen aus, weil die müssen ja auch alle dafür bezahlt werden. Es muss ja auch auf irgendeinem Server laufen. Irgendwelche CPUs müssen das auch tatsächlich berechnen. (..) Und da könnte es natürlich so sein, dass die Leute in dem Ganzen nur zustimmen, wenn ich denen sage Schau mal, ich habe hier diesen zertifizierten Partner, die sind unabhängig, nicht gewinnorientiert, haben das, machen das schon eine ganze Weile mit diesen Datentreuhänder. Ihr gebt eure Daten nicht mir, sondern ihr gebt die denen und die legen die da rein und das wird da berechnet. Weil sonst haben die vielleicht Angst, dass ich aus deren Maschinendaten was sonst was machen will. Vielleicht wollen die ihre Daten nicht mir geben, sondern nur denen.

33 [0:33:33.7] **I:** Ja. (..) Also würdest du sagen, der Grund, wieso es nicht geklappt hat, war hauptsächlich aber das Rechtliche eben, dass das Unternehmen immer noch Angst hatte oder die Banken, dass sie gegen irgendwas Rechtliches verstoßen.

34 [0:33:49.8] **B5:** Das rechtliche war tatsächlich okay am Ende. Sie haben quasi, um das jetzt nicht zu weit auszuführen, aber sie haben quasi den Use Case dann noch mal ein bisschen kleiner gemacht und haben jetzt mit drei Banken tatsächlich zum Laufen bekommen. (..) Aber halt nur mit drei. Ich meine natürlich, am besten wäre es, wenn man alle hätte oder halt die zehn größten in Deutschland beispielsweise. Das wäre natürlich schon sehr hilfreich.

35 [0:34:15.1] **I:** Aber das Vertrauen letztendlich scheitert.

36 [0:34:17.4] **B5:** Genau das scheitert mehr an der Compliance Abteilung von so einer Bank. Rein rechtlich hat tatsächlich die Finanzaufsicht gesagt, dass das okay ist. Gut, dass das okay ist. Ich meine, da müsste man dann natürlich eigentlich quasi eine sehr genaue rechtliche Beschreibung dessen machen, damit man sich dann wirklich sicher sein kann.

37 [0:34:35.4] **I:** Okay. Aber sie haben einfach, wenn zu viele Akteure quasi im Netzwerk sind, dann wird es einfach zu unüberschaubar, zu wenig Vertrauen dann.

38 [0:34:46.9] **B5:** Ja in dem Fall wären sogar wenig genug Akteure im Netzwerk gewesen. Also es liegt an vielen Stellen hier, eben auch viel an der Struktur. Dass die Compliance im Unternehmen eher erstmal sagt, machen wir lieber nicht, weil könnte schief gehen. Und dann? Ich glaube, das ist der Grund, warum die meisten es erstmal abgesagt haben. Zum Beispiel.

39 [0:35:10.3] **I:** Und gibt es irgendwelche Bedingungen, die erfüllt sein müssen, damit so ein Daten, damit man überhaupt Daten mit Secure Multi Party Computation austauschen kann?

40 [0:35:27.1] **B5:** Also letztendlich und das ist das Lustige an diesen Zero Trust Technologien, müssen wir dieser Technologie vertrauen. Also wenn wir die nicht selber entwickelt haben, na gut, selbst wenn wir die selber entwickelt haben, dann haben die anderen die nicht entwickelt. Also irgendjemand muss dieser Technologie vertrauen, dass das passt. Das ist so die eine Sache. Und sonst müssen natürlich quasi alle anderen Bedingungen, die wir vorhin hatten, trotzdem erfüllt sein. Wir müssen halt trotzdem schon mal irgendeinen Gegenwert darin sehen, überhaupt Daten auszutauschen. Vielleicht ist quasi das Risiko verringert. Also quasi in der Kosten - Nutzen - Rechnung sind die Kosten oder das Risiko ist niedriger, wenn wir bestimmte PCs haben, aber unsere, die nutzen müssen immer noch da sein und vielleicht sind ganz andere Kosten die viel größere treibenden Faktoren.

41 [0:36:17.4] **I:** Und glaubst du, dass PETs wie Secure Multiparty Computation das Datenaustauschverhalten von Unternehmen in der Zukunft verändern werden oder nicht?

42 [0:36:35.1] **B5:** Ich denke ja. Also wie gesagt, ich hoffe bzw gehe eigentlich auch davon aus, dass quasi je mehr man darüber redet und je mehr das so über mehr und mehr Stellen in der Breite einsickern, dass man diese Sachen hat und dass die vertrauenswürdig sind und dass die funktionieren und dass ein paar Vorreiter, die vielleicht schon mal das ein oder andere ausprobiert haben, nicht direkt ihre Existenz verloren haben oder halt auch nichts groß schiefgegangen ist. Dann wird da halt mehr und mehr Vertrauen einerseits quasi in den Datenaustausch an sich, aber auch in Privacy Technologien stattfinden. Und das wird einfach dafür sorgen, dass der Datenaustausch an sich mehr

und mehr eine valide Option wird für ganz viele verschiedene Sachen. Also ich glaube wir sind noch nicht an dem Punkt, wo sich genug Leute darüber Gedanken machen, was man mit mehr Datenaustausch nicht alles erreichen könnte. Also ich glaube, das ist ein bisschen wie bevor wir Smartphones hatten, konnte sich auch keiner vorstellen, was man nicht alles für Apps machen kann. Und wenn wir jetzt mal ein bisschen mehr dahin kommen, dass wir so eine Art von Infrastruktur haben und das alles ein bisschen einfacher und sicherer und vor allem einfacher, weniger komplex ist. Man muss sich nicht über alle Sachen Gedanken machen, dann wird da, glaube ich, ein großes Umdenken passieren.

43 [0:37:59.6] **I:** Die Themen haben wir eh schon angeschnitten, mit Vertrauen und Kontrolle. Ist ja beides gefordert bei einem. Oder braucht es beides für den Datenaustausch zwischen Unternehmen? Wie werden denn diese Faktoren Vertrauen und Kontrolle verändert? Durch Privacy Enhancing Technologien? Oder werden sie verändert? Werden sie weniger wichtig, wichtiger?

44 [0:38:20.4] **B5:** Naja, also wenn wir jetzt mal mal zehn Jahre in die Zukunft springen und wir uns vorstellen, dass wir gewisse PCs haben, die sehr breit bereits akzeptiert sind und viel genutzt werden. Dann sorgen die, glaube ich dafür, dass man quasi weniger Vertrauen in jedem einzelnen Datenaustausch hat oder zu jedem einzelnen Use Case, für den man Daten austauschen würde, muss man weniger Vertrauen fassen, um den dann auch durchzuführen. Kontrolle. Das kommt natürlich dann sehr darauf an, wie diese Technologie ist, weil die kann letztendlich dafür sorgen, dass ich mehr Kontrolle über meine Daten habe. Oder auch nicht. Also das kommt eben auf die explizite Technologie drauf an, also es gibt auch so Ideen, dass man quasi weil letztendlich, wenn ich jemandem meine Daten zur Verfügung stelle, dann kann ich natürlich einen Vertrag unterschreiben lassen, dass die nicht weitergeben dürfen. Aber ich kann auch nicht verhindern, dass sie die weitergeben. Also es gibt quasi dann auch die Idee, dass ich den anderen meine Daten, also nicht tatsächlich quasi mein PDF schicke, sondern ich lade die in einen virtuellen Raum ein, indem sie sich mein PDF durchlesen können, aber die PDF nicht mitnehmen dürfen. Okay, gibt es auch als Idee.

45 [0:39:47.8] **I:** Aber wieso hast du jetzt gesagt in zehn Jahren? Glaubst du, dass jetzt Privacy Enhancing Technologien schon noch Vertrauen erfordern?

46 [0:39:54.3] **B5:** Jetzt ist der Faktor noch größer, dass ich meine, dass man in die Privacy Enhancing Technologie vertrauen muss. Da ist es dann viel weniger ein Faktor, wenn die eine oder andere schon schon sehr verbreitet ist. Ich glaube, letztendlich entwickelt sich das bei ganz vielen Sachen so. Ich meine, wir haben jetzt einen elektronischen Personalausweis, der ausnahmsweise mal deutlich besser ist, als er Anwendungsfälle hat. Wenn das quasi mehr publik wird, dann funktioniert

das auch alles deutlich besser. Und natürlich ist es auch jetzt schon so, dass PCs bei sowas helfen. Ich meine, wenn wir keine Anonymisierung und Pseudonymisierung hätten, dann könnten wir noch viel weniger Sachen mit personenbezogenen Daten machen. Also das ist ja schon die einfachste. Aggregieren und Zusammenfassung von irgendwelchen Sachen ist ja letztendlich Privacy Enhancing und deswegen hat es jetzt schon ganz große Auswirkungen auf den Datenaustausch.

47 [0:40:56.1] **I:** Und in welchem Sektor siehst du am meisten Anwendungsgebiete für Privacy Enhancing Technologien oder auch Regionen oder irgendwelche Faktoren von Unternehmen?

48 [0:41:11.0] **B5:** Also je sensibler die Daten, desto mehr Bedarf an PTS, das heißt die Finanz, Gesundheit und Bildung sind wahrscheinlich die sensibelsten Sachen. Ja, größer glaube ich schon eher naheliegend, gerade in diesen Bereichen, dass man, wenn man ein großes Unternehmen ist kommt natürlich auch. Also es kommt wieder darauf an, was man daraus bekommen kann. Also wenn ich ein großes Unternehmen bin und quasi daraus viel Mehrwert generieren kann, dann lohnt es sich vielleicht eher, weil es einfach besser skaliert, dass ich mir, dass ich eine PC entwickle, eine bestimmte oder mehr entwickeln lasse. Wenn ich ein kleines Unternehmen bin, dann skaliert das halt nicht so gut. Ja, ja.

49 [0:42:04.5] **I:** Okay. Und wie würdest du sagen verändern Privacy Enhancing Technologien die Risiken des Datenaustausches? Würdest du sagen, da kommen neue Risiken hinzu? Auch, oder?

50 [0:42:19.1] **B5:** Na ja, also ich meine, rein prinzipiell und das, wie gesagt, mit dem Vertrauen in die Technologie selber könnte ich natürlich auch was schreiben, dass das nicht Open Source ist, dass niemand nachprüfen kann und dann verkaufe ich das den Leuten als Privacy Enhancing Technologie. Aber letztendlich sorgt es dafür, dass alle Daten an mich geschickt werden. Also ich meine, das kann natürlich alles passieren und auch wenn es Open Source ist, also wir hatten jetzt ja gerade letztens erst so einen Hack, der quasi in Open Source passiert ist, da muss man auch aufpassen. Ich glaube, das wäre so ein großes zusätzliches Risiko. Man könnte, ja, man könnte natürlich zu leichtfertig werden. Klar. Also es wäre natürlich gut, wenn man die PTS auch versteht und was sie tatsächlich dann machen. Weil sonst könnte mir das natürlich jemand auch vielleicht ein bisschen überschwänglich verkaufen. Und dann gebe ich einfach meine Geschäftsgeheimnisse raus. Aber am Ende lässt sich das trotzdem reverse engineeren, weil es halt einfach keine zu starke Regierung ist und jemand mit genug Rechenpower kriegt das auch wieder raus oder wenn er genug Daten von allen einsammelt beispielsweise. Also das muss man schon immer noch betrachten.

51 [0:43:38.9] **I:** Und als letzte Frage was würdest du sagen ist die größte Herausforderung für die Akzeptanz von PETs auf dem Markt?

52 [0:43:52.6] **B5:** Das ist eine gute Frage. Also ich meine, da können die PETs nichts dran machen oder nichts dafür eigentlich. Aber je besser der Use Case ist, wegen dem ich Daten teilen möchte, desto mehr bin ich natürlich auch bereit oder desto mehr sehe ich ein, dass ich Platz brauche. Beziehungsweise, desto mehr habe ich Bedarf dafür. Vielleicht ist das noch die größte Herausforderung, dass es einfach noch nicht so viel, also vergleichsweise nicht so viele große Use Cases gibt, bei denen Firmen PETs brauchen und viel Mehrwert daraus generieren. Das ist wahrscheinlich das größte Problem. Also quasi die Nachfrage danach, wenn man so will.

53 [0:44:45.6] **I:** Aber vielleicht wissen Sie auch noch gar nicht, dass es Anwendungsfälle geben würde.

54 [0:44:50.8] **B5:** Ja, genau das ist die nächste Sache. Und dann haben wir, wenn wir darüber nachdenken und dann kommt jemand vielleicht von einem Startup und sagt Schaut mal, wir haben hier diese Technologie, die würde euch ermöglichen, dieses und jenes zu machen. Da sind wir natürlich jetzt auch schon sehr weit so im digitalen Zeitalter angekommen, dass solche Anfragen häufig vorkommen. Und dann liegt es in der Natur von Algorithmen, dass man sie nicht so einfach versteht, gerade wenn man nicht vom Fach ist. Also ich kann das auch durchaus sehr gut nachvollziehen, wenn man jetzt, um wieder das Beispiel zu nutzen, wenn das Kerngebiet eigentlich ist, dass man gute theoretische baut als Maschinenbauer und alle vier Wochen oder so, trifft man irgendjemanden, der sagt Ja, lass da mal ganz viele Algorithmen nutzen und dann können wir das irgendwie alles verbessern und dann hat man irgendwas mal ausprobiert, das hat halt nicht funktioniert. Und dann denkt man sich Na gut, also ich weiß jetzt nicht, ob ich irgendwas mit meinen Daten verkaufen will. Ich verkaufe ja eigentlich theoretische dann also ich kann. Ich glaube, dass da viele Leute jetzt schon müde sind, gerade halt in den Branchen, die nicht so digital affin sind. Ganz grundsätzlich, weil man damit eigentlich nicht so viel zu tun hat.

55 [0:46:06.0] **I:** Ja okay. Und wenn man dann auch noch so viele Anfragen bekommt, muss man wahrscheinlich auch noch rausfiltern, welche wirklich was bringen und welche nicht.

56 [0:46:15.0] **B5:** Ja genau. Und das ist halt sehr schwierig, die herauszufiltern. Also auch wenn man Programmierer ist, es ist schwierig herauszufiltern, wer mir versucht zu erzählen, dass sie was Gutes, einen guten Algorithmus gebaut haben und wer tatsächlich einen guten Algorithmus gebaut hat. Also das einzuschätzen ist ja super schwierig je nach Spezialisierung. Also ich glaube, man sieht es auch so ein bisschen an dem ganzen Blockchain Hype, der dann auch sehr stark wieder abgeflacht ist, weil da dann auch festgestellt wurde, dass das am Ende alles sehr fancy klingt und sehr schön ist, aber einem gar nicht so viel bringt, wie man vielleicht dachte.

57 [0:46:49.7] **I:** Ja, ja gut. Sehr vielen Dank für das Interview und sorry, dass es jetzt schon echt länger gedauert hat.

58 [0:46:57.7] **B5:** Alles Okay. Ich bin davon ausgegangen.

59 [0:46:59.5] **I:** Ich glaube, das war mein letztes Interview, was ich bisher geführt habe. Also vielen Dank nochmal.

60 [0:47:10.2] **B5:** Gerne. Sehr gut.

61 [0:47:11.8]

62 [0:47:16.0]

1 Interview B6:

2

| | |
|---|---|
| Interview-Nr. | B6 |
| Date of the interview | august 23, 2024 |
| Duration of the interview | 27:03 min |
| Interviewer | Felix Starnecker (I) |
| Interviewee | B6 (Germany) |
| Role | Head of IT |
| Sector | IT sector |
| Specialities | Interview in german |

3 [0:03:44] **I:** Meine erste Frage wäre, ob du Unternehmen kennst, die Daten mit anderen Unternehmen teilen und. Welche Partner das dann wären also mit wem, die Ihre Daten teilen und wofür die Ihre Daten teilen. Oder ob du da niemanden kennst?

4 [0:04:33] **B6:** Ach so ne? Ich häng grad gedanklich bei uns. Weil wir ja durchaus Lohn und Finanzbuchhaltung zum Beispiel für unsere Kunden übernehmen. Das heißt, hier ist natürlich auch ein Datenaustausch und von Lohnbuchhaltungsdaten et cetera notwendig. Das geht ja schon mit ein bisschen Kreativität in die Richtung, ja. Ja, also kenne ich machen wir selber, wenn wir so möchte.

5 [0:05:04] **I:** Dann habt ihr das dann quasi in eurer Software, dass ihr da irgendwie den Datenaustausch vereinfacht, ist das Teil.

6 [0:05:13] **B6:** Ja, genau so ist es bei uns. Wir haben quasi die unsere Kundensysteme laufen entweder bei uns, wo wir dann letztendlich Hosting, Partner oder cloudservice Anbieter sind. Ja, und die Kunden greifen dann über eine gesicherte Internetverbindung und letztendlich SSL Verbindung quasi auf Windows System in unserer Welt zu oder die andere Möglichkeit ist, dass die Kunden die Systeme selbst hosten oder bereitstellen. Ihrem eigenen Rechenzentrum oder bei ihrem Hosting Partner und wir dann über eine entweder auch über eine SSL Verbindung quasi. Über HTTPS oder aber über eine Side to site VPN Verbindung drauf zugreifen.

7 [0:06:08] **I:** OK und gibt es da dann irgendwelche Verträge oder Vereinbarungen mit denen? Die getroffen werden davor, oder?

8 [0:06:19] **B6:** Ziemlich sicher es sind Namen, es sind immer Non Disclosure Agreements. Na ja, die Geheimbündungsvereinbarungen. Wo auch meines Wissens nach hohe Vertragsstraßen draufstehen, wenn das nicht eingehalten wird.

9 [0:06:35] **I:** Und würdest du sagen, dass das eher jetzt, wenn Unternehmen entscheiden, ob sie da ihre Daten allgemein teilen? Ob das eher ein subjektiver Entscheidungsprozess ist oder ein sehr objektiver, also ob.

10 [0:06:48] **B6:** Das ist, das ist maximal subjektiv mein, ich kann es von mir selbst also grundsätzlich und da brauchen wir nicht mal auf auf strukturierte Daten sind sich abzielen, sondern grundsätzlich was, was die Datenhoheit betrifft, das immer eine subjektive Entscheidung, die ganzen Hyperscaler Cloud Anbieter, Amazon, Microsoft, wie sie alle heißen Google ja, das sind ja immer noch beispielsweise verpönt im deutschsprachigen Raum. Jeder möchte schauen, dass die Daten irgendwo im EU Rahmen Festland liegen, um so ein Bild der Hoheit über die Daten zu behalten. Am besten ist aber. Finde ich ein bisschen schwierig, aber am besten ist natürlich für die Unternehmen die Daten liegen irgendwo, bei denen im Rechenzentrum oder in einem, also auf ihren eigenen Server.

11 [0:07:39] **I:** Also glaubst du jetzt wenn so Daten zu irgendwas Produktionsdaten oder irgendwelche Daten geteilt werden, dass da also, dass da einfach eine Person irgendein Manager drüber entscheidet, ob das gemacht wird oder nicht, aber sich so mit Risikoanalysen et cetera eher weniger befasst wird ja OK. Und welche Daten, welche gefahren würdest du beim Teilen von Daten am meisten sehen?

12 [0:08:10] **B6:** Am meisten ist tatsächlich das in meinen Augen, das menschliche Versagen beziehungsweise das technische Versagen, sprich an das Daten an die Öffentlichkeit kommen entweder, weil die Technik nicht also weil die Verschlüsselung nicht so funktioniert, wie sie funktionieren soll beziehungsweise und das haben wir ganz oft zum Beispiel datenschutzverstöße ja, weil weil irgendein Mitarbeiter jetzt wieder nicht drauf aufgepasst hat, wem er jetzt was teilt. Ja, und dann? Die Kundendaten von Kunde A bei Kunde B? Also ich gehe davon aus, dass das vertraulich behandelt wird und nachher unternehmensnamen und sowas rausgeschwärzt werden weitgehend ok.

13 [0:10:06] **I:** Ja, ja OKO. Un welche Risiken würdest du sagen, sind am größten, wenn Daten augetauscht werden?

14 [0:10:20] **B6:** Letztendlich stehen ja auch Datenschutzverstöße auch hohe Strafen. Von der

Gesetzesgebung her haben entgegen ja. Das ist sowas. Was auf jeden Fall, wo das Risiko relativ hoch ist einfach, dass das Eintritt. Und was haben wir noch? Und alles, was, was nicht personenbezogene Daten betrifft mein Industriespionage und so weiter und das ist aber auch so ein Grund, warum sich viele deutsche Unternehmen oder deutschsprachige Unternehmen schwer tun, ihre Daten irgendwie in einer Amazon oder Microsoft. Zu geben quasi ja, diese aufgrund dieses wie heißt in Amerika das Patriots Gesetz? Irgendwie sowas, wo halt sich die amerikanische Regierung abgesichert hat, dass die im Krisenfall letztendlich, und das ist deren Abwägung ja dann auch Zugriff auf sämtliche Daten, die nur Microsoft oder Amazon Cloud oder in der Google Cloud gespeichert sind.

15 [0:11:39] **I:** OK also dann ist eher gar nicht die Gefahr von Konkurrenz unternehmen, sondern mehr die Gefahr von also anderen Regierungen, dass die die Daten benutzen.

16 [0:11:50] **B6:** Genau und in China ist ja beispielsweise mehr als offensichtlich, dass die chinesische Regierung und das ist keine Verschwörungstheorie, sondern kann man auch nachlesen, dass die aktive Industriespionage machen. Und Moment. Bei den Amis ist es ich mein, das ist der Patriot Act hier genau Telefon und internetüberwachung als Kontrollinstanz also, das fällt da schon ein ja, das heißt im im Terrorfall und das wie man der Eintritt.

17 [0:12:44] **I:** Mhm, OK, interessant und. Würdest du sagen, dass äußere Einflüsse auch oft Unternehmen dazu bringen, ihre Daten zu teilen? So also quasi, dass irgendwas von außen, die dazu veranlasst, mehr Daten zu teilen, in Deutschland oder allgemein.

18 [0:13:06] **B6:** Ja also. Es fängt ja an mit mit dem ganzen, nenn es mal am Gesellschaftsgesetzen, also GMBH Gesetz, Aktiengesetz und so weiter soll ja alles noch transparenter werden. Dass hier quasi auch Zahlen veröffentlicht werden, müssen in Zukunft, was in der Vergangenheit vielleicht veröffentlicht wurden, was dann durchaus auch ins ins. Geschäft oder eine Auswirkung aufs Geschäft hat ja, wenn wenn mein Wettbewerb meine Zahlen kennt, können die gegebenenfalls auch hier nach Maßnahmen einleiten, um sich selbst besser darzustellen. Ja, also die Gefahr ist schon da, ja, je transparenter das alles wird.

19 [0:13:46] **I:** Ja. Mhm, OK und dann noch die letzte Frage zu dem Fragenblock was müsste sich ändern oder gegeben sein? Das Unternehmen noch mehr Daten teilen würden, weil es ja also in Deutschland werden Parteien Unternehmen bisher relativ wenig Daten oder sind da relativ vorsichtig. Du, da gibt es irgendwelche Faktoren wo? Unternehmen vielleicht mehr Daten teilen würde, wenn das eintreten würde? Oder braucht sie einfach n bisschen länger, bis bis man da offener

gegenüber steht?

20 [0:14:24] **B6:** Ich glaube, es muss ganz klar der Wettbewerbsvorteil. Herauskristallisieren oder herausgearbeitet werden, dass die Unternehmen bereit sind, mehr Daten zu teilen. Ja, und in Deutschland sind wir einfach so wir leben ja vom Mittelstand, letztendlich sind alles kleine bis Mitte bis mittlere Unternehmen, die für sich agieren. Ja, wo viel Wettbewerbsgedanke auch dabei ist die deutschen Automobilhersteller zum Beispiel die haben sich ja gerade in Hinblick auf Elektromobilität so ein bisschen zusammengetan. Zu der Allianz beziehungsweise Joint Venture draus gegründet, wo dann durchaus, da muss halt jeder seinen Beitrag leisten und seine Daten bringen ja, dass man da gemeinsam arbeiten kann, aber das ist glaube ich in Deutschland also für mein Gefühl. Noch zu wenig ausgeprägt, ja, also da muss schon noch was passieren. Warum machen es die Autos die Automobilhersteller? Weil es sich natürlich ein Wettbewerbsvorteil erhoffen beziehungsweise auch einen einen gesunden Gegenpol darstellen wollen gegenüber Tesla und wie wie die anderen alle heißen ja im Bereich Internet.

21 [0:15:20] **I:** Mhm, Mhm. Dann, ich kann dir jetzt würde der Teil kommen, wo ich dich noch zu Secure multiparty computation paar Sachen fragen würde. Meine erste Frage wäre, für welche Anwendung Anwendungsfälle fändest du Privacy enhancing Technologien wie Secure Multiparty Computation sinnvoll beziehungsweise findest du es überhaupt sinnvoll? Glaubst du, da gibt es irgendwelche potenziellen Anwendungsfälle? Oder glaubst du, es ist einfach noch nicht?

22 [0:16:24] **B6:** Ich kann mir durchaus vorstellen, dass wir wieder zurück zu dem Beispiel kommen der Automobilhersteller, dass es da von Vorteil sein kann, weil die letztendlich noch mehr Daten, was eigentlich natürlich in deren RND Umgebung einfließt und niemand offenlegen möchte, ja, wo aggregiert quasi über 3 oder 4 Hersteller hinweg dann durchaus zu einem gemeinsamen Ziel führen kann, dass da einsetzbar ist? Also ohne wirklich weitere konkrete Ideen zu haben, weil ich bin von der Autoindustrie so weit weg. Da könnte ich es mir vorstellen.

23 [0:17:16] **I**: Aber jetzt für so vielleicht Kunden, die bei euch oder die du hattest bisher Unternehmen glaubst du, die würden da irgendwie drauf, hätten irgendwie nen Sinn damit ja?

24 [0:17:45] **B6:** Ne vielleicht Logistik ja, wenn ich weißt du so am Kostenoptimierung in der Logistik mit ein Schenker und keine Ahnung, wie die ganze Spedition heißen Duvenbeck oder sowas, wenn die sich hier Richtung Kostenoptimierung zusammentun. Ja, und ihre Realkosten quasi reinfüttern die wollen aber natürlich nicht, dass der andere sieht. Und die dürfen natürlich ihre ihre Kosten auch nicht preisgeben. Zwecks Kartell, Gesetzen und sowas ja, aber dass man da irgendwie so ein

Durchschnitts also ganz, ganz simpel gesagt so Durchschnittskosten über mehrere Unternehmen hinweg bildet, dass jeder für sich selber so ein bisschen Maßstab hat, aber optimieren kann oder nicht?
Also ich glaube, da könnte auch.

25 [0:18:51] **I:** Mhm also, so Benchmarking quasi Mhm ja.

26 [0:18:57] **B6:** Genau genau grundsätzlich für Benchmarking Design, Ich glaub ganz gut getroffen. Wieviel.it Budget haben ist verfügbar wieviel Prozent vom Umsatz ist das und so weiter? Und wenn wir da so eine gemeinsame Datenbank hat, ja, man muss alles reinfüttern, sich selber dann aber nur die Durchschnittswerte über seine Branche Rauszieht, letztendlich das, was Gartner für teuer Geld. Die ganze Zeit macht kann ich mir durchaus vorstellen, dass da Mehrwert da ist.

27 [0:19:35] **I:** Mhm und Du hast ja vorhin gesagt, dass diese Entscheidungsprozesse oft eher subjektiv ablaufen glaubst du dann mit so Technologien, die hören sich ja jetzt, wenn man das hört, relativ sicher an sind doch irgendwie erstmal schwer zu verstehen glaubst du das einfach nur das Gefühl dass so ne Technologie dahinter steht? Datenteilungsverhalten also. Sich positiv drauf auswirkt, dass noch mehr Daten geteilt werden, weil man sichereres Gefühl hat, mit solchen Technologien.

28 [0:20:11] **B6:** Glaub da spielt die Gesetzgebung auch eine extrem große Rolle, ja und um das Verständnis und Gesetze sind halt so so schwammig teilweise formuliert ja, dass jeder eher vorsichtig ist, was was Datenteilen betrifft, weil es darf natürlich nicht alles rausgegeben werden. Und ich glaub, da spielt das mehr eine Rolle als die Technologie, die drunter liegt, oder die drüber liegt, wie auch immer ja. Die europäische Welt muss einfach offener werden. Hinsichtlich Datenschutz, Datensicherheit, ja, dass unsere am Unternehmen gegenüber, amerikanischen oder chinesischen Unternehmen auch wettbewerbsfähig bleiben.

29 [0:21:01] **I**: Siehst du neue Risiken, die durch die Nutzung von Secure Multiparty Computation entstehen könnten?

30 [0:21:12] **B6:** Risiken für einzelne Unternehmen oder Risiken für den ganzen Markt oder?

31 [0:21:23] **I**: Eher für einzelne Unternehmen das irgendwie vielleicht.
Daten auch manipuliert werden könnten zum Beispiel.

32 [0:21:40] **B6:** Also. Nein, das Daten, also Datenlecks, sind das eine haben überall, wo ich eine neue

Technologie hab, ist die Gefahr eben da, dass dass weil die Technologie noch nicht ausgereift ist, dass irgendwo ein Sicherheitsleck ist, wo die Daten dann irgendwie abgezogen werden können, das ist das eine für. Unternehmen seh seh ich eher den Mehrwert von dem ganzen er gerade Richtung Benchmarking gedacht. Für den für den Markt an sich sehe ich aber tatsächlich das Risiko, dass wir noch mehr in so eine Preisdrückerei dann auch abdriften könnten. Bei der Maximaloptimierung von allem Optimierung aus dem Unternehmenskontext natürlich immer. Hier Kosten einsparen ja und wenn dann also aus dem letzten Dienstleister quasi noch das letzte bisschen rausgedrückt wird? Eine Gefahr von Marktansicht, da in meinen Augen.

33 [0:22:53] **I**: Ach so OK, ja interessant ja OK.

34 [0:22:59] **B6:** Und vielleicht noch mal zu der Frage von vorher fällt mir nämlich gerade ein, wo das Ganze noch ne Rolle spielen könnte, ist zum Beispiel Bauwesen, wo ganz viele Unternehmen zusammenarbeiten müssen ja, wo wir das eine übergreifende an ich war 4 Jahre lang nächstes Jahr lang in der Baubranche. Habe da mit Secured Data Rooms gearbeitet, wo letztendlich Pläne drin gelandet sind, excelkalkulation und so weiter alles, was relevant ist für so n Bauprojekt für so n Riesenbauprojekt wird da allerdings nur quasi die Die die harten Zahlen einträgt, weißt du in Form von der Datenbank, wie es jetzt hier angedacht ist? Ja, haben alle beteiligten Unternehmen einfach viel besser die Möglichkeit, das zu extrahieren und so kann quasi innerhalb Security kann man trotzdem zusammenarbeiten. Ich muss nicht aus dem Architektenplan Mühelvoll mühsam von Hand irgendwelche Maße rausziehen, sondern das steht alles in so einer Datenbank drin. Ja, und der andere also, dann braucht der Zimmermann braucht beispielsweise nur beispielsweise nur die Maße vom Dach, der kann dann nur auf die Maße zugreifen. Der Maurer kann auf die Maße vom Haus zugreifen also ganz.

35 [0:24:11] **I:** OK, aber sind es dann so Daten, die überhaupt gesichert werden müssen? Es so sensible Daten dann oder wenn die einfach so geteilt werden würden, wird es dann nicht genauso gehen?

36 [0:24:23] **B6:** Ich glaub, das kommt n bisschen drauf an mein, wenn man allein wenn wir drüber reden, dass natürlich so aus Datensicherheitstheorie per Perspektive soll ja jeder nur auf die Daten zugreifen können, die für ihn auch relevant sind, ja nicht alles offen für jeden. Im Bau ist es anders ein bisschen ja, das sind halt die Architektenpläne und jeder Handwerker hat darauf Zugriff. Das ist die Frage will ich das wirklich? Ja, beziehungsweise wenn dann 2 beispielsweise Mehrfamilienhaus für jede und jeder, der so eine Wohnung besitzt, jeder Bauherr quasi von einzelner Wohnung arbeitet, in unterschiedlichen Elektriker zusammen. Auch dafür wäre das dann denkbar, dass man das viel modularer gestalten kann.

37 [0:25:07] **I:** Mhm ja, OK. Mhm, dann passt meine nächste Frage eher nicht ganz gut, bei welchen Sektoren unternehmen würdest du am ehesten ehesten sehen, dass die in Secure Multiparty Computation investieren würden, also dann wahrscheinlich Bauunternehmen vielleicht oder die Baubranchen allgemein?

38 [0:25:23] **B6:** Genau. Bau und Automotive.

39 [0:25:31] **I:** Weil der viel Wettbewerb hat und Mhm OK. Ja gut, dann. Hätte ich nur noch die Frage, was du als größte Herausforderung für die Akzeptanz von Secure Multiparty Computation siehst, aber das hatten wir eigentlich eh schon etwas.

40 [0:25:43] **B6:** Ja, genau also grundsätzlich geht es in meinen Augen nicht um die Technologie, sondern wirklich um die Bereitschaft, Daten teilen zu wollen, ja letztendlich damit einhergehend auch die Gesetzeslage.Ja, darf ich es?Oder da?Das ist ein erste Herausforderung.

41 [0:26:02] **I**: Glaubst du, das ändert sich in Zukunft also, dass in der Hinsicht lockerer wird?

42 [0:26:10] **B6:** Es es muss sich ändern, also in meinen Augen muss ich es ändern.
Vor kurzem auch ein Podcast veröffentlicht da kannst du auch mal reinhören, da sag ich genau das gleiche, dass wir wettbewerbsfähig bleiben als Europa als Deutschland. Müssen wir unser Mindset ändern, müssen wir offener sein für Technologien, auch für das Thema Datenhoheit müssen wir offener sein, um auch beispielsweise sowas wie KI wirklich gezielt einsetzen zu können?

43 [0:26:51] **I**: OK Mhm OK nice, ja mach ich. Dann erstmal vielen Dank für das Interview! Wie heißt denn der Podcast?

44 [0:27:03] **B6:** Kein ding. [Podcastname]

1 Interview B7:

2

| | |
|---|---|
| Interview-Nr. | 7 |
| Date of the interview | september 4, 2024 |
| Duration of the interview | 30:11 min |
| Interviewer | Felix Starnecker (I) |
| Interviewee | B7 (Germany) |
| Role | Managing Director in IT security investment company |
| Sector | Venture Capital with focus on cybersecurity |
| Specialities | Interview in german |

3 [0:00:05] **I:** Ok also der erste Teil, der über (..) ist welche Faktoren allgemein Unternehmen dabei beeinflussen, ob sie ihre Daten teilen oder nicht. Und der zweite Teil besteht dann daraus, wie sich diese Faktoren oder das allgemeine Data Sharing Verhalten von Unternehmen durch Privacy enhancing Technologien verändert. Und speziell eben Secure Multiparty Computation. Meine erste Frage wäre: Mit wem tauschen Unternehmen ihre Daten aus? Was kannst du aus deiner Erfahrung hierzu sagen? Sind es Wettbewerber, Lieferanten, Kunden und wofür?

4 [0:00:50] **B7:** Im Wesentlichen logischerweise sind es eigentlich schon Kunden und Lieferanten. Also typischerweise wird eben so ein Datenverarbeitungsvertrag geschlossen für die meisten SARS Companies. Als Beispiel (..) Wir sind gerade dabei, Investment zu machen in der Company, die ebenso Llm basierte Bots Knowledge Bots zur Verfügung stellt. Da müssen natürlich auch Daten übertragen werden, also zum Beispiel fütterst du eben die Maschine mit Daten. Du wirst alle Dokumentationen zu einer Maschine zum Gerät einfach mal in so einen Raum und wenn der Servicetechniker kein Standardproblem hat vor Ort, sondern ein spezielles Problem, dann kann er quasi wie bei chatGPT fragen, was er tun soll. Und dann soll eben die Maschine die richtige Antwort geben und auf die richtige Seite verweisen. Dabei müssen auch die Daten geshared werden. Und typischerweise agieren da die Sept Contractor als Auftragsdatenverarbeiter und damit bleiben viele von den GPR Compliance Themen sozusagen nicht beim Verarbeiter hängen, sondern bei der Firma selbst. Datenaustausch bei Wettbewerbern ist natürlich immer tricky. Das Beispiel, das ich zuvor gemeint habe mit den Banken, die denken natürlich schon darüber nach: Wie kann man das tun? Da wäre das so ein potenzieller Use Case wie gesagt, Ich habe noch nicht funktionieren sehen, eben nicht aus technischen Gründen, sondern aus kommerziellen oder Risikoabwägung.

5 [0:02:57] **I:** Und bei dem Bankenbeispiel meintest du ja, dass dann doch Vertrauen wichtiger ist und

deswegen hat es nicht funktioniert oder?

6 [0:03:05] **B7:** Wir haben uns einen Use case gemacht mit drei Banken oder sowas um zu zeigen, dass es funktioniert (..) da haben wir einfach halt Fake Daten angegeben und dann war im Prinzip die Idee, ob wir das erweitern auf - ich sag jetzt einfach mal Hausnummer 10. Dann gab es Diskussionen über die richitgen Daten, gute Daten. Das war über ein Jahr lang haben die immer am Vertrag noch diskutiert, also letztendlich kam dann das Konstrukt nie zustande. Und die klare Aussage von dem Team war eigentlich ist es kein technisches Problem. Es gab es entsprechende Untersuchungen natürlich. Ob das alles vernünftig ist und technisch gut implementiert ist. Aber letztendlich hat, dann hat die Unterschrift gefehlt, das kann man spekulieren. Und ich die wahrscheinliche Spekulation ist irgendjemand hatte halt eben dann doch Angst, die Daten zu teilen.

7 [0:03:53] **I:** Na ja, OK ja und würdest du allgemein sagen also ich hab in meiner Literaturrecherche Vertrauen und Kontrolle als wichtige Faktoren herausgearbeitet. Was würdest du sagen, ist wichtiger oder würdest du sagen, beides ist wichtig?

8 [0:04:16] **B7: I**ch hatte gerade vorhin einen, also vor dir ein Telefonat mit mit KPMG, die für unseren vor die Geldwäscheprüfung machen. Es sind so viele Themen, die du dokumentieren musst. Also wurde, wo du kontrollierst, also wo wo wo beauftragte Datenschutzbeauftragte Geldbeauftragte einfach kontrollieren müssten dokumentieren müssen. Glaube ich bekommst du das Thema Kontrolle eigentlich nicht rum.

9 [0:04:43] **I:** Würdest du sagen der Entscheidungsprozess nee, kein Objektiv?

10 [0:04:49] **B7:** Kein objektiver Nein, jedes Bankenbeispiel, dann wäre es ja objektiv, wenn man nachweisen kann, das ist keine Ahnung du du findest 2 sicherheitsaudits und die Stellen fest, dass da keine offensichtlichen Trap Doors sind. Du machst immer das das Bug Bounty Hacking sowieso aber. Das Beispiel zeigt ganz klar das subjektiv da immer noch der Punkt ist also das das sehe ich bei bei vielen Cases das technologisch also wir brauchen gar nicht so weit gehen, dass du MPC oder oder PETs nimmst, wo wirklich dann ja wertvolle Daten und keine contention Daten geshared werden. Gehen wir ne Stufe drunter an anderen Daten auch da seh ich schon, dass relativ oft die Frage ist, ob Cloud Computing eingeführt werden soll beim ganzen deutschen Mittelstand. Es gibt die die großen Cloud Anbieter, die die versichern natürlich auch ja wir, wir speichern die Daten nicht in den USA. Wir nennen den Nutzen unsere Cloud Center in Europa. Da ist Redundanz, da gibt es die Security Audits und trotzdem sagen viele ich will aber nicht alles zu Microsoft geben oder ich hab Angst, dass Google dann doch irgendwie mit dem USA Act auf die Daten zugreift, lieber nicht ist glaub ich auch ein Existenzbegründung für manche lokale Deutsche Cloud, die nie mithalten kann, kommerziell mit einem Hyper scaler, aber trotzdem ihr Geschäft machen.

11 [0:06:02] **I:** Okay, und was würdest du? Was sind die größten Gefahren oder Risiken, die du siehst, wenn man Daten teilt?

12 [0:06:21] **B7:** Also wenn es eben diese hochsensiblen Daten sind, dass die geleakt werden. Mit hohe kommerzielle Risiken, also eine kommerziell vom vom Use Case von Business Case für die Firma und wenn ein Leck nachgewiesen werden kann, hast du auch noch Compliance, also regulatorische Auflagen, also gerade Banken zum Beispiel sind da sehr, sehr reguliert. N anderes Beispiel ist ganz banale Analogie Schlüssel Keymanagement für incription Technologien. Du weißt vielleicht schon mal HSM gehört High Security Module da gibt es n paar Anbieter Thales ist einer der ganz großen, das ist im Prinzip ein harten, harten Rechner also ein Blade sozusagen, wo du spezielle. Hat ja auch implementiert hast die einfach sichere Räume schaffen und Schlüssel zu speichern. Also wirklich im im im Background, sozusagen bei den Issuern, die nutzen HSMS. Viele Banken nutzen das und du könntest relativ einfach zum Beispiel mit Confidential Commuting, so ein HSM in der Cloud also sprich rein in Software realisieren. In Klammern Photonics die Amerikaner die allerersten in Conference Computing Bereich. Die machen hauptsächlich, damit ihr Geschäft eigentlich mit dem keymanagement Lösungen. Aber die, die Bankenregulierung schreibt ganz oft noch vor, dass da steht, noch irgendwo drin. In der Richtlinie Hardware und ganz banal, ja und dann sagen natürlich viele Risk Officer. Wenn ich mich dann im Buchstaben halte, der der Vorschriften, dann kann ich sowas nicht nutzen. Ich muss bei der Hardware bleiben, auch wenn es langsamer ist. Teurer ist Punkt, Punkt, Punkt.

13 [0:08:09] **I:** Obwohl es anders da nicht besser wär oder genauso gut, muss man trotzdem gemacht werden.

14 [0:08:13] **B7:** Ja, "you never get fired for buying IBM". Man ist oft zu zurückhaltend wegen regulierungen bei Datenaustausch Entscheidungen. Das ist schlecht für Innovation.

15 [0:08:23] **I:** Okay, und was würdest du sagen, sind Voraussetzung für den erfolgreichen Datenaustausch? Oder gibt es da überhaupt welche?

16 [0:08:39] **B7:** Ich glaube, das ist das ist im Prinzip ich glaube, das ist in in manchen Segmenten. Gewisse Zeit eingesetzt wird und der Proof da ist, dass da nichts passiert ist. Einfach das Vertrauen zu erwerben, dass da, dass es ich würde glauben, dass das schrittweise mehr Raum greift, also bei kleinen, speziellen Use Cases sozusagen anfängt, die weniger riskant sind, da keine negativen Erfahrungen gibt. Wenn dann irgendwann mal sich die ein paar größere aus dem Quark trauen, das Einsetzen auch darüber reporten, dass es dann entsprechend weiter in den Markt reinwachsen kann.

17 [0:09:14] **I:** Und was würdest du als externe oder interne Einflüsse sehen, wo Unternehmen vielleicht mehr oder weniger Daten teilen beziehungsweise daran gehindert werden?

18 [0:09:29] **B7:** Also Angst vor der Regulatorik würde ich sagen Angst vor dem vor dem Business Case Angst vor Reputationsverlust? Neue Technologie. Muss ich der erste sein?

19 [0:09:41] **I:** Ja, ja und so interne Einflüsse gibt es auch. Also würdest du sagen, dass die Unternehmenskultur der Management ist da wichtig ja?

20 [0:09:49] **B7:** Ja, ich glaube das ist aus meiner Sicht auch ganz stark. Wenn du ein Top Management hast, das sehr progressiv ist, wo vielleicht nicht der der Kaufmann sozusagen das IOS, sondern der Techie der CEOS, die des in speziellen Bereichen einfach mal austesten, wo jetzt gerade Investment machen den Bereich und die auch mit einem Chief Sales Officer sprechen da kein reiner Sales Guy, sondern Tech Tech gig auch ist mit dazu und der hat einfach gesagt wir probieren es jetzt einfach aus Punkt und der schiebt halt dann die Bedenkenträger weg also ich glaube, sowas ist.

21 [0:10:11] **I:** Also wenn man sich da besser auskennt, dann ist man da auch offener gegenüber quasi.

22 [0:10:37] **B7:** Ja, das ist die einfach. Die Risikobereitschaft Innovationstrieb und welchen kommerziellen Benefit hab ich dadurch? Und dann werden einfach die einfach sagen, die ja mutigeren Firmen wahrscheinlich da früher Anfang das andere. Ich glaub auch das Segment also im Bankenbereich glaub ich ist super schwierig, weil die halt hochreguliert sind und wenn die ein Problem haben irgendwo, das wird sofort schweineteuer durch die Regulierer durch Strafen, aber halt eben auch die Reputationsverlust. Es kommt immer raus, es wird publiziert und ich glaub, deswegen haben die da wahrscheinlich sehr, sehr große Bedenken.

23 [0:11:16] **I:** Also in dem Sektor ist noch mal stärker als in anderen Sektoren, quasi ja OKOK Gesundheitssektor wäre das dann auch?

24 [0:11:20] **B7:** Ja. Ja, also in Deutschland ganz schwierig erstaunlicherweise, ich hab vor kurzem was heißt vor einem Jahr ein Startup gesehen, wo es auch drum dieses Thema wie geht man mit Daten, um Patientendatenakte und Austausch und erstaunlicherweise das hat mich total gewundert. Die Firma hatte gar keine PET Technologie angewandt. Ich hätte erwartet, dass die irgendwie MPC Competential Homomorphic, weiß der Geier, was gerade geht sozusagen anwenden. Die kam aber aus der Perspektive. Die hatten einen Datenschutzbeauftragten eines Bundeslandes involviert und es war extrem prozessorientiert also. Wie also prozessvorschrift? Glaube da ist auch mal da ist in diesen regulierten Märkten wie Gesundheit das eine ist die ganze technische Lösung ist aber glaube ich. Die andere Prozessvorschrift ist mindestens genauso wichtig. Da eben, dass es geprüft wird, dass es dokumentiert ist, also der der ganze Prozesskram wie es abläuft ist, steht da ziemlich stark im Vordergrund. Ich war entsetzt, muss ich sagen ich habe nämlich gelernt, dass universitären Umfeld zu Forschungszwecken, das heute bereits alles möglich ist, sprich da die Uniserver sind wir vielleicht nicht so gut. Erwartet und und unter Beobachtung wie jetzt bei einem bei einem Corporate

des Daten zu Datenverlust einfach halt wirklich kommerziellen Schaden nimmt. Da war es OK, aber wenn wir es dann eben versucht auszurollen, dann ist es extrem schwierig also ich hätte ich hätte die Bedenken, dass halt Grad Healthcare als speziell Deutschland super stark reguliert ist. Gematik und. Was auch immer die für einen Blick haben auf die Technologie, dann funktioniert oder eben nicht funktioniert unabhängig von würd ich einfach mal sagen technischer vernünftiger Meinung das ist einfach noch mal noch mal politisch sozusagen regulatorisch beeinflusst. Ja, aber das ist schon interessant, dass genau die Sektoren, wo eigentlich das meiste Potential vielleicht da wird, so im Gesundheitssektor oder im Finanzsektor, dass da dann, wenn wir doch nicht geht, weil es so stark reguliert ist. Gerne Anekdote ich hab angefangen 1995 operativ also nach meinem Studium nach meiner Disc zu arbeiten. Damals bei Infineon im Security IC Bereich 9596 wurde bereits über die Patientendaten, also über die Patientenakte mit Chipkartenzugang diskutiert. Wir haben 2024 und vielleicht kommt es jetzt mal?

25 [00:13:54] **I:** Ja, ich glaub gibt es das nicht seit diesem Jahr jetzt dieses E Rezept oder seit letztem Jahr.

26 [0:13:59] **B7:** Ja, die Versicherer müssen jetzt einen sicheren Zugang halt eben gewährleisten.

27 [0:14:14] **I:** Ja, das ist aber ich glaub, das ist auch nur in Deutschland so, oder? In den anderen europäischen Staaten, die sind da teilweise glaub ich weiter.

28 [0:14:22] **B7:** Also ganz, ganz weit ist Israel zum Beispiel aber bei der Corona Pandemie auch gesehen. Die hatten als erste die Daten. Bei anderen Ländern ich glaub, das ist alles so speziell Regulatorisch in dem Land und aus dem Grund hab ich nämlich bei Healthcare zurückgezogen, weil ich gemerkt hab reines Technologiedenken ist da völlig der Holzweg, da brauchst du einfach die regulatorische Einsicht. Und wer dominiert was in welchem Markt? Wenn du als Startup da rein möchtest, brauchst du dann wirklich die richtigen Leute, die in den Netzwerken sind, die die Märkte verstehen, die Regeln verstehen. Wenn du da rein technisch denkst hast du verloren aus meiner Sicht.

29 [0:14:57] **I:** Obwohl schon viel Potential wahrscheinlich da wäre.

30 [0:15:00] **B7:** Theoretisches Potenzial, ja.

31 [0:15:09] **I:** SO bezüglich Secure multiparty computation ist es so, man braucht keine zentrale Partei, die da irgendwie die Daten dann normalerweise sieht. Es wird alles verschlüsselt und das Besondere ist halt auch auch, dass während Berechnungen immer noch verschlüsselt ist. Genau und? Ja, man muss eben keiner zentralen Partei vertrauen, die da die Daten verarbeitet und berechnet, sondern es geht alles dezentral genau und das wird eben, ich weiß nicht, ob du das gehört hast mit

Secret Shares wird das gemacht, dass man so Daten in kleinere Teile aufteilt. Genau das, das ja genau da ist, hier das Beispiel mit Person AB und C wollen Ihren Durchschnittslohn herausfinden und wollen aber nicht ihre ihren Lohn quasi preisgeben und dann werden die Daten quasi. Zum Beispiel Person A teilt die Daten auf in also die $100 in 50 3020 Personen b in -80 100 180 und dann wird eben ein Secret share. An jede Partei geschickt. Und das dann wieder nach unten. Summiert und dadurch durch diese Secret Shares ist auf jeden Fall die Essenz. Bekommt niemand den totalen Zugriff auf die ursprüngliche Datei, sondern nur auf das Endergebnis, und das ist das der Sinn dahinter. Und ja, meine erste Frage wäre jetzt für welche Anwendungsfälle würdest du am meisten? Privacy, enhancing Technologien oder auch spezielle Smpc sinnvoll finden. Oder wo glaubst du funktioniert, wo würde es nicht funktionieren?

32 [0:17:28] **B7:** Also just weil wir daüber diskutiert haben, wenn wenn halt eben mit Big Data AI Analysen auf auf Daten von verschiedenen Quellen gemacht werden sollen und letztendlich keine Daten des anderen sehen soll. Das ist so n so n Punkt, ich hab vorhin das auch bei der Bank erwähnt im Prinzip. Täglich aktuelle Zinskonditionen aus aus einer Mischung von vielen verschiedenen Zinsdaten bei der Bank sich holt solche Dinge also wo, wo verschiedene Parteien Daten einlegen und halt eben ein Ergebnis daraus haben. Da ist eigentlich ein Interview mit [Unternehmen] wahrscheinlich ziemlich gut, ich mag die ja, jetzt können Sie hierzu ich bin bin ja, ich kam zu spät zum Investment damals die.

33 [0:18:16] **I:** Hättet ihr überlegt in die zu investieren?

34 [0:18:21] **B7:** Hab mal mit den engen Gesprächen, in der in der allerersten Runde da war es aber noch sehr, sehr, sehr früh und die zweite Runde wollen wir dann wieder auch bisschen weiterbauen in unseren Verständnis, da war es dann zu spät. Hatten sie bereits genügend Investoren. Und die, die hatten damals die Waren, die ersten, die drauf kamen. Damals haben alle versucht, dieses Company Computing einfach zu machen für Anwender, wo du heute eigentlich immer einen echt gut informierten IF i so brauchst oder halt einfach alte Spezialisten brauchst, die das anwenden können. Sie haben gemerkt, dass es aber zu komplex ist, sind weggegangen und haben wirklich so ein Use Case, den Securator darum zum Beispiel als Use Case damals entwickelt und auf den Markt gebracht und letztendlich was ich das Gute fand, ist, dass du gar nicht die Technologie nach vorne gestellt haben, sondern einfach den Use Case und halt eben im zweiten Schritt dann den Interessierten nachgewiesen haben, dass ihre Technologie einfach hart ist. Und die Anforderungen erfüllt, aber haben nicht die Technologie verkauft, das fand ich eigentlich ziemlich gut.

35 [0:19:18] **I:** Also erst erst die Nachfrage quasi beweisen und dann sagen, dass es dafür eine Technologie gibt.

36 [0:19:27] **B7:** Ja, denn die Technologien MPC oder alle PETs würde ich mal sagen Hypothese sind extrem, also wer kann die anwenden? Wie gesagt, unsere Seconet spezialisiert die die kommen aus dem Krypton Bereich, das sind sind Nerds in der Richtung und die haben Schwierigkeiten, das gut anzuwenden. Das heißt, wenn du das auf den Markt bringen möchtest, als Technologie verkaufst, dann hast du einen ganz, ganz engen Markt, weil nur ganz wenige Unternehmen überhaupt in der Lage sein werden, das umzusetzen, und es wird ewig dauern, was natürlich auch ihre eigenen Zweifel haben prüfen wollen. Wie sicher sind eigentlich wirklich und halt Schritt für Schritt, das Umsetzen? Deswegen fand ich es bei [Unternehmen] extrem gut, dass die sich 2 Use Case sage ich einfach am Anfang überlegt haben, die im Prinzip den Kunden gegeben haben und dann halt eben die Technologieargumentation nur genutzt haben, zu zeigen, dass es eben halt vielleicht kostengünstiger, einfacher zu installieren ist also, desto mehr Parteien leichter sozusagen an den Tisch kriegst und nicht halt eben on premise. Keiner muss vertrauen, dass der, der rechnet, sozusagen bei Partei A steht oder beim Hyperskiller B und das letztendlich dann die der Vertrauensbeweis durch durch die Mathematik schlechte Technologie gegeben wird. Das fand ich eigentlich ziemlich gut. Ich glaube, wenn ein Start anfängt, mit MPC zu argumentieren, dann ist der Adressatenkreis weltweiter das verstehe ich verdammt klein.

37 [0:20:54] **I:** Ja okay, also würdest du auch sagen, dass die Hauptbedingungen, wieso sich Privacy enhancing Technologien vielleicht nicht durchsetzen oder das Haupthindernisses einfach die fehlenden Use Cases vielleicht oder die fehlende Praktikabilität ist.

38 [0:21:09] **B7:** Ja, ich beobachte seit 2019 es waren die ersten Dinge, die wir uns angeschaut haben. Wir 2024 und die meisten Firmen also die meisten, die ich kenne Krebssen irgendwo bei einer 1000000 umsatzfirma oder sowas also eigentlich eher so POC Ebene noch die sind weg was jetzt aktuell keine Daten die müssten weiter sein und ich hab gemerkt, dass es also der Versuch, die Technologie sozusagen zu verkaufen wird scheitern, weil es einfach a sehr komplex ist, also wenn sie du jetzt in stellfall du bist vertriebst, weil du verkaufst, dann irgendjemanden, du musst anfangen so ein einfaches Beispiel mit Bob, also ABC und und Klammer hast das vertriebst man selber nicht genau verstanden, wie es funktioniert. Was Hochtechnisch ist, das zu verkaufen, wird scheitern aus meiner Sicht.

39 [0:21:50] **I:** Und der ist ja auch noch in kurzer Zeit Verkauf, also erklären das ja.

40 [0:21:56] **B7:** Und das in kurzer Zeit ja, weil nach nach 10 Minuten schalten die Leute ab und sagen ja, verstehe ich nicht, tut mir leid das war s also ich glaube, das geht nur über gute Use Cases also. Einfach zu Realisierende Use Cases und da fand ich [Unternehmen] ein ganz gutes Beispiel. Ich glaub es braucht unbedingt diese ganz einfachen Use Cases, die du recht einfach in den Markt kriegst, solange es nicht ist, wird es funktionieren. Und die einfachen Use Cases sollten nicht

unternehmenskritisch sein. Es war wichtig, aber nicht unternehmenskritisch aus meiner Sicht, weil dann doch die Leute die gewisse Scheu haben, halt ihre höchstsensiblen Daten dreinzustecken, ohne gewisse Profiles und wenn du den Markt genügend Erfahrung hast, also wenn man genügend Referenzen aufbauen kann, ich glaub, dann wird es ganz erweitern können.

41 [0:23:17] **I:** Mhm, OK ja, interessant und zudem Beispiel mit den Banken würdest du sagen das Vertrauen trotzdem, dass man durch SMPC vielleicht mehr Kontrolle noch mal über seine Daten hat, weil sie verschlüsselt sind, ist Vertrauen immer noch wichtig. Ja OK, also ganz ohne Vertrauen geht es nie.

42 [0:23:44] **B7:** Die Regulatorik, so Regulatorik, zwingt dich auch dazu, dass du entsprechende Nachweise triffst also bei bei manchen Segmenten halt eben. Warum das? Warum du das nehmen kannst und darfst oder warum du denkst, das ist Risikomanagbar ist da ist auch Vertrauen natürlich die Grundvoraussetzung, aber ne gewisse Kontrolle oder Nachweisfunktionalität dazu.

43 [0:23:56] **I:** Und welche Risiken siehst du beim Datenaustausch mit Privacy enahncing Technologien? Kommen da neue Risiken hinzu, die es davor vielleicht nicht gab?

44 [0:24:18] **B7:** Wie vorhin schon gesagt einfach dieses dieses so riesige kommerzielle Risiko. Homomorphic ist ein rein mathematisches Verfahren, gibt aber ich sag mal ne Handvoll Firmen weltweit das machen, das heißt wie? Nachgewiesen, dass es wirklich sicher ist Analogie, wenn Du die Kryptoverfahren Standard Kryptoalgorithmen, anschaust RSA oder ähnliches, da ist, da ist jeden Tag weltweit sind die Hacker dran, wieder ein paar Stellen mehr zu brechen. Das heißt, du hast eigentlich permanent den Beweis was ist machbar, was ist nicht machbar? Also du weißt, wo das Sicherheitsniveau liegt sozusagen ich bin immer ganz kritisch und alle Security Leute sind ganz kritisch bei neuen Verfahren und ganz besonders bei Black Box verfahren, die nicht öffentlich gelegt werden. Weil du eben die die Anbietende Firma wird ja immer sagen, das ist sicher und wir haben 2 Leute das rechnen lassen, das ist sicher, aber wenn du deine, wenn du deine Kronjuwelen nach irgendwas legst und vertraust, dann bist du wirklich wissen dass das also wie hoch das Sicherheitsniveau wird eigentlich wirklich versucht, die ganze Welt das zu attackieren und zu hacken und gibt es da einen Nachweis, dass es eben, dass es eben sicher ist und homomafic, die versuchen sogar einen mathematischen Beweis als Nachweis zu bringen, dass es sicher ist?

45 [0:25:18] **I:** Ja ja, also es fehlt, es fehlt immer noch das Vertrauen einfach in die Technologie, weil sie noch so unbekannt ist, ja, Mhm.

46 [0:25:44] **B7:** Ja also für die normale Firma frag mal irgendeinen Mittelständler in Deutschland, ob er MPC kennt. Der wird dich ganz groß anschauen und sagen vielleicht vielleicht hab ich schon mal RSA gehört oder? Oder des für n Verschlüsselungsalgorithmen, ob du dich jemanden frage, mit der

wir sagen ja, die sind seit 20 Jahren im Markt und wenn immer ged und da kannst du vertrauen, wenn der 5 Experten aus seinem Netzwerk fragt Was ist ein SMPC? SMPC. Dann werden die Mehrheit. Wird das nicht wissen, schätz ich mal.

47 [0:26:15] **I:** Und bei welchen Unternehmen siehst du das größte Potential? Also vom Sektor Region Größe.

48 [0:26:39] **B7:** Also ich glaub schon dass wie es wieso oft die innovationsoffenen Regionen da stärker sind. USA, Israel. Vielleicht sogar auch die, die baltischen Länder in Europa, Zentraleuropa ihr ihr hinten dran, würd ich sagen. Von den Regionen her von den Segmenten ich würd eben wie gesagt nicht auf die, also wo das wo wahrscheinlich das größte Potenzial liegt im Financial Services Bereich im Healthcare Bereich. Ich glaub, dass die einfach durch Regulatorik noch ein bisschen ängstlicher sein müssen.

49 [0:27:15] **I:** Ja. Und Größe? Aber egal weil oder wie siehst du das? Wie der wie viele finanzielle Ressourcen man da braucht, um da zu investieren das. Würdest du sagen?

50 [0:27:30] **B7:** Also wenn es eine Gepacket der Lösung ist, die man relativ einfach subscriben kann, sozusagen, was es Ziel ja sein sollte, dann wäre die finanzielle Größe nicht so wahnsinnig entscheidend. Wenn wir es einfach mal so die, die in Europa es gibt paar riesengroße Unternehmen da schauen, wahrscheinlich viele hin, aber die, die haben viel zu verlieren. Deswegen werden die das nur in Randbereichen wahrscheinlich versuchen erstmal zu testen. Die Kernfrage ist, haben die Expertise im Haus, die das einigermaßen bewerten kann. Oder was ich mir noch vorstellen könnte, ist, dass es vielleicht stark auch eher über Berater kommt, weil eben dieser Mittelstandsbereich, wenn die keine Fachleute haben, dass sie dann einfach zum Beispiel ein IT Consultant oder ein Security Consultant holen und sagen wie kann ich das implementieren? Also Beispiel so Themen wie Geldwäsche, Richtlinien oder oder Compliance gibt es sehr viele IT oder einfach Consultants, die das für die kleinen Firmen anbieten, weil die natürlich selber gar nicht die Manpower haben. Alle Anforderungen, die es europaweit oder weltweit gibt, jeden Tag zu durchforsten, das hat sich geändert in den Regularien, das umzusetzen. KPMG oder ähnliche Firmen können Vertrauen erzeugen, das könnte sehr hilfreich sein für die Einführung solcher Technologien.

51 [0:29:00] **I:** Dann bei dir sieht er besser aus können. Ja, das.

52 [0:29:12] **B7:** In den Bereichen im Namen haben. Ich würde mir denken das würde ich als Startup versuchen, so einen Kanal zu nutzen, um da in Markt zu kommen, weil klar, wenn wir beide ne Firma haben und gehen jetzt zum Kunden und sagen vertrauen sie mir, das ist alles sicher. Die werden dir nicht vertrauen, zumindest nicht bei bei ihren wichtigen Daten. Kommt aber halt eben der der IT Consultant oder Provider. Inzwischen haben und hilft Ihnen dabei. Dann ist

wahrscheinlich die Wahrscheinlichkeit höher.

53 [0:30:11] **I:** Gut, das war es jetzt eigentlich also vielen Dank fürs Interview sorry dass das bisschen länger gedauert hat.

1 Interview B8:

2

| | |
|---|---|
| Interview-Nr. | 8 |
| Date of the interview | September 13, 2024 |
| Duration of the interview | 26 min |
| Interviewer | Felix Starnecker (I) |
| Interviewee | B8 (Germany) |
| Role | Head of IT Security \| CISO |
| Sector | IT sector |
| Specialities | Interview in german |

3 [0:03:37] **B8:** Genau also ich bin ich bin [Name]., Ich bin CISO bei der [Unternehmen]. [Unternehmen] macht eine ganze Menge individualsoftware Entwicklungsprojekte im deutschsprachigen Raum mit großen Kunden als Ziel. So bin ich einmal für die Informationssicherheit verantwortlich der Systeme, die [Unternehmen] im Entwicklungsprozess und so grundsätzlich benutzt. Und ich bin gleichzeitig Head of Security, das heißt Verantwortung im Geschäftsbereich für einen Security Team Wir machen eine eigene Security Beratungsprojekte in unterschiedlicher Form von testen über sichere Software, Entwicklung, Training und Organisationen von IT ab Abteilung und genau das ist sozusagen meine Spielwiese.

4 [0:04:18] **I:** Alles klar, OK und aus deiner Erfahrung mit wem würdest du dir denn jetzt sagen tauschen Unternehmen Daten aus? Also erstmal mit welchen Parteien und wofür?

5 [0:04:31] **B8:** Mhm also, wenn ich jetzt mal in unser Kundenumfeld mich umschaue, dann sind das meistens Lieferanten Beziehungen, die dahinter stecken. Das heißt irgendwelche Zulieferungen von entweder teilen oder von Prozessabschnitten und dann werden auch Systeme miteinander integriert. Das hängt ein bisschen von der Branche ab. Also nur wenn du im E Commerce Bereich irgendwie im Handel unterwegs bist, dann geht es irgendwie um Lieferketten von Artikeln und Lieferbeständen und solchen Themen in der Automobilbranche sind es dann eher Teile und Spezifikationen von Produkten und in der in der Logistik. Hast du dann sozusagen die Vernetzung von verschiedenen Logistikprovidern, sei es jetzt Bahn, sei es Flugzeug oder Transport, auf der auf der Straße, wo eine ganze Menge Unternehmen miteinander zusammenarbeiten müssen, um ja auch sozusagen die Pakete von AB zu kriegen.

6 [0:05:33] **I:** Mhm und würdest du sagen, dass bei diesen Beziehungen vertrauen ne wichtige Rolle spielt, also dass sich Unternehmen quasi kennen, vielleicht wissen, dass es andere deutsche Unternehmen sind, oder?

7 [0:05:40] **B8:** Ja. Ja, also Vertrauen ist ja sozusagen genau mein Thema mit. Ich glaube, an der Stelle wird das tatsächlich hauptsächlich aktuell durch Vertrauensbeziehungen in der Zusammenarbeit zwischen Partnerunternehmen oder Lieferkettenunternehmen etabliert. Die es gibt durch die Informationssicherheit natürlich Instrumente wie Prüfkataloge wie ne ISO-Zertifizierung oder ähnliches, mit dem man versucht das Vertrauensniveau in den Partnern im Umgang mit Daten zu erhöhen. Aber am Ende des Tages müssen sie immer noch darauf vertrauen, dass die Unternehmen das im Griff haben und mit ihren Daten ordentlich umgehen. Deswegen würde ich sagen in so einer Zusammenarbeitsmodell ist typischerweise das Vertrauen relativ hoch, weil es technisch gesehen wenig Mittel gibt, sicherzustellen, dass es tatsächlich auch gewährleistet ist, dieses Vertrauen.

8 [0:06:23] **I:** Mhm. Ja ja, aber würdest du was würdest du dann sagen? Durch was entsteht Vertrauen sind das irgendwie ist die Herkunft von dem Unternehmen oder Beziehungen zwischen den? Für den Topmanagern oder wie entsteht das Vertrauen?

9 [0:06:50] **B8:** Das ist ne sehr gute Frage. Ich, ich glaube, es hat was mit der Beziehungsebene zu tun natürlich ne, dass man irgendwie sind 2 Dinge einmal sozusagen die Vertrauen auf auf Managementebene, wenn es irgendwie um partnerschaftliche Zusammenarbeit geht. Und der zweite Teil des Vertrauens wird, glaube ich, durch so Prüfprozesse zumindest aufgebaut, also in der Zusammenarbeit durch Verträge und durch solche Fragenkataloge also wir als Dienstleister kriegen das ja auch immer wieder und unternehmen, die mit uns zusammenarbeiten, die mit uns von uns Software gebaut haben wollen, die müssen ja auch Vertrauen aufbauen. Das passiert über mehrere Stationen. Erstmal ist es eine inhaltliche Vertrauenstellung, d.h. traue ich, demjenigen das zu, dass er das fachlich inhaltlich kann, was ich brauche. Dann werden Verträge geschlossen und dann im Rahmen dieses Prozesses werden. Dann müssen Fragenkataloge beantwortet werden. Und muss man sich irgendwie die Augen schauen und gucken, können wir ausreichend gut begründen, dass wir das im Griff haben? Und die andere Seite beurteilt dann ja das klingt irgendwie sinnvoll, also wenn sie es hoffentlich im Griff haben, so auf der Ebene passiert das aktuell.

10 [0:07:56] **I:** Also auch Vertrauen durch Kontrolle, dann quasi so ja und ja, bei euch müssen sich wahrscheinlich dann die Unternehmen auch, also die müssen euch ja auch vertrauen, dass weil die müssen euch ja auch ziemlich viele Daten wahrscheinlich weitergeben, wie sie ihr Produkt haben wollen und so also da wird wahrscheinlich oft nur vertrauen.

11 [0:07:59] **B8:** Ja ja. Ja, es gibt genau es gibt unterschiedliche Formen von Daten, das muss man glaub ich an der Stelle noch mal separieren. Das eine sind natürlich sowas wie fachliche Konzepte businessmodelle ne womit verdienen wir Geld und wie dann hast du die, sag ich mal operativen Daten, Kundendaten artikeldaten, die sozusagen produktionsrelevant sind da versuchen wir als Dienstleister die Hände vorn zu lassen. Das heißt, möglichst dafür sicherzustellen, dass wir diese Daten nie in die Hand bekommen, sondern nur Systeme bauen, basierend auf Testdaten, weil wir gar nicht die Verantwortung für die Sicherstellung von großen Kundendatenbanken oder Ähnliches übernehmen wollen. Weil wir das ja also ist halt sozusagen einfach nen Pattern, dass wir uns da dadurch absichern, dass wir die Daten nicht in Hand nehmen, klappt aber nicht immer also ne, wenn wir jetzt ne Bank nehmen und wir wir haben für ne für ne Banken online Banking gebaut und die uns Zugriff auf ihren Produktionssystem geben, dann sagen wir nein wollen wir gar nicht haben.

12 [0:09:01] **I:** Okay. Und würdest du sagen, dass es ein subjektiver oder objektiver Entscheidungsprozess ist, wenn sich Unternehmen dafür entscheiden, Daten zu teilen? Also quasi entscheidet, dass eine Person einen Topmanager so oder ist das n, lange risikoanalysen et cetera?

13 [0:09:33] **B8:** Hängt ein bisschen von der Unternehmensgröße an ab? Oder beides? Also es gibt sozusagen bei größeren Konzernen, es ist eher eine längere Risikoanalyse und wie gesagt, bei uns, wir versuchen das eh zu vermeiden, aber wenn wir nicht drum herum kommen sollten, dann gibt es tatsächlich längere Risikoanalysen und so Kontrollmechanismen, wo wir dann einfach geprüft werden. Es gibt auch Unternehmen gerade in der Finanzindustrie ist es so die schicken dann auch mal jemanden vorbei, der uns auditiert. Und in den anderen Fällen ist es teilweise auch echt sehr hemdsärmelig. Das Unternehmen sozusagen unbedacht auch Daten mit uns teilen, die wir eigentlich nicht haben wollen, also gerade wenn Bildungsprozess.

14 [0:10:09] **I:** Mhm OK ja ja. Und welche Risiken siehst du beim bei gemeinsamer Datennutzung beziehungsweise Datenaustausch? Oder was sind die Hauptrisiken oder?

15 [0:10:29] **B8:** Ja, die Vertraulichkeit sicherzustellen. Das ist auch der Grund, warum wir das Versuchen zu vermeiden, was weil wir noch keine gute andere Alternative haben, ne das Konzept was du vorgeschlagen hast, wenn man das technisch sicherer machen könnte. Daten miteinander dezentral zu teilen, ohne sie gleich komplett offenzulegen, dann würde das sicherlich helfen. Das Problem ist du hast ja in diesen Lieferketten zum Beispiel du hast ja keine Kontrolle über die Unternehmen und eines, der dieses Data Loss Prevention Thema ist sozusagen Kernproblem in dem

Moment wo ich die Datenteile muss der andere eigentlich genauso viele Sicherheitsmaßnahmen selber sozusagen operationalisieren und auch wirksam haben und darüber hab ich ja gar keine Kontrolle mehr und das ist n Problem.

16 [0:11:15] **I:** Ja, und gibt es irgendwelche Voraussetzungen, die besonders wichtig sind fürs Teilen von Daten? Also vielleicht das überhaupt möglich ist?

17 [0:11:33] **B8:** Ja, also weiß nicht, in welche Richtung du da guckst, ne das eine sind natürlich vertragliche Absicherungen. Das ist so der Sicherheitscode, den die meisten Unternehmen versuchen, anzulegen, dass man sich dann mit Strafen belegt, wenn man nicht damit ordentlich umgeht oder wenn irgendwas passiert und. Das zweite ist Kontrollmaßnahmen etablieren, ist aber auch, glaube ich dem geschuldet, dass es noch keine gute technische, also es ist nicht etabliertes technische Maßnahmen dafür zu implementieren. Also wenn du so an Confidential Computing zum Beispiel denkst, dann ist das einfach noch nicht etabliert, ja. Und dann, dass die Schnittstellen miteinander sozusagen ausgehandelt werden und auch entsprechend abgesichert werden. Wie zum Beispiel bei Logistikern. Da sind technische Schnittstellen implementieren eine Voraussetzung für den Datenaustausch und die müssen dann auch abgesichert werden. Das ist bei uns jetzt ein bisschen anders, weil wir, wenn wir testdaten irgendwie zur Verfügung gestellt bekommen, dann kriegen wir meistens Zugriff auf die Systeme der Kunden. Wir versuchen, sozusagen dabei zu bleiben, dass die Daten im Zuhause des Unternehmens bleiben und wir in deren Anwendungslandschaft Zugriff bekommen, ohne dass wir die Daten irgendwie raus tragen können, wenn du Systeme unternehmensübergreifend vernetzt aber auch weil Systeme miteinander reden müssen, dann musst du da natürlich auch noch mal absicherungsmaßnahmen treffen.

18 [0:13:15] **I:** Mhm, Mhm, OK und die letzte Frage eigentlich zu dem Themenkomplex kennst du irgendwelche oder würdest du sagen, es gibt externe oder interne Einflüsse, wo Unternehmen vielleicht mehr dann dazu gebracht werden, Daten zu teilen oder auch weniger?

19 [0:13:38] **B8:** Ja, sind hängt ein bisschen von der Kritikalität der Daten ab. Und welche Relevanz sie haben ne, wenn man zum Beispiel jetzt an so KI Modelle denkt und es darum geht, sie trainieren zu lassen, wenn man dann auf einen gemeinsamen Wissenspool zum Beispiel aufsetzen würde und um gemeinsam einen Trainingssatz zu zu trainieren, dann tun sich Unternehmen da durchaus schwer, immer diese diese Wissensbasen und die Daten Töpfen wir auch auszurücken, weil das ja ein Kern Asset ist, was für ihr Geschäft relevant ist.

20 [0:14:26] **I:** Mhm, OK gut. Und dann zum zweiten Teil. Für welche Anwendungsfälle würdest du

Privacy enhancing Technologien am ehesten als sinnvoll erachten, also speziell Secure Multi Party computation oder kannst du dir vorstellen, dass die überhaupt sinnvoll ist oder praktisch anwendbar?

21 [0:14:50] **B8:** Ja, ich kann mir das gut vorstellen, ich mach mal ein Beispiel: ich sitze in Hamburg, da gibt es viele öffentliche Unternehmen und einfach jetzt mal aus dem Security Bereich Rausgegriffen. Die haben öffentliche Einrichtungen und kleine mittelständische Unternehmen haben die Herausforderungen, sogenannte SOCs aufzubauen, also Security Operations Center, wo es darum geht, eine kontinuierliche Überwachung der Netzwerke für Unternehmen durchzuführen und darauf festzustellen, ob es Anomalien gibt. Jedes mittlere bis größere Unternehmen muss sowas versuchen aufzubauen, das ist wahnsinnig teuer und know how intensiv, du brauchst Security Spezialisten dafür. Also gibt es zum Beispiel Initiativen und Ideen, dass das nicht ein Unternehmen alleine macht, sondern dass sich Unternehmen zusammenschließen und es einen Verbund gibt, der sozusagen diese Dienstleistung erbringt. Das Problem ist, dass diese Daten aber hochsensibel sind und kein Unternehmen wird, die einfach jemandem zur Verfügung stellen, das ist so ähnlich wie mit Trainingsdaten. Die werden nicht einfach aus der Hand geben, so dass jemand anders auf mein Daten lernt, nur dass es eben, wenn Security Beispiele ist und ein Hemdschuh ist. Dass man sozusagen diese Art von Services und Betriebsmodell gar nicht betreiben kann, weil es da noch keine gute Lösung für gibt, wie man gemeinsam diese Daten nutzen kann, ohne sie aus der Hand zu geben. Und das zweite, das ist dann eher der Bereich Thread Intelligence. Es gibt der typische Angriffsmuster. Die wenn du jetzt so eine Attacke dir zum Beispiel Anguckst, die hat ein bestimmtes Verhalten und darüber gibt es Daten. Und die will man eigentlich auch miteinander teilen. Aber auch dafür gibt es keine richtig gute Plattform. Und Mechanismen, weil das eben auch vertrauliche Daten sind so und idealerweise kann man solche Daten miteinander teilen, ohne sie aus der Hand zu geben, um aus dem Datenbestand, dass jedes Unternehmen hat zu lernen, ne um Son Beispiel zu machen. Ein Unternehmen stellt fest ich hab ein bestimmtes Angriffsmuster bei mir, bei mir ist was passiert und weil die anderen das auch prüfen, ob das passiert ist. So wie tausche ich das jetzt miteinander aus? Also es gibt bestimmt noch ganz viele weitere Beispiele. Aber Moment, ich hab das jetzt mal rausgekramt, weil es für sensible Daten steht, wo es sich lohnen würde, sie miteinander zu teilen, ohne sie tatsächlich aus der Hand zu geben und zu teilen.

22 [0:18:09] **I:** Ja ja, Mhm ja, ja hört sich interessant an. Und gibt's irgendwelche Bedingungen, die erfüllt sein müssten, damit sich Datenaustausch, dass der erfolgreich ist oder ist einfach, wenn der Use Case passt, dann?

23 [0:18:32] **B8:** Ja gut also der Use Case muss natürlich passen, aber am Ende sozusagen eine systemische Lösung, so Privacy by Design, die einfach sicherstellt, dass ich die Daten nicht

verlieren kann, oder das, was immer wie bei einem, bei dem bei dem Partner, mit dem Ich teile, passiert ich die Vertraulichkeit gewahrt ist. Einfach systembedingt und zwar dem Partner gegenüber und möglichen Angreifern, die vielleicht die Daten auch einfach irgendwie raustragen, dass wir natürlich nice.

24 [0:19:07] **I:** Mhm und würdest du also glaubst du, dass Secure Multiparty computation oder allgemein Privacy enahncing Technologien in der Zukunft den Datenaustausch verändern werden oder dass es viel mehr damit kommt?

25 [0:19:20] **B8:** Ich glaube schon, was ich ich glaube, es wäre ein wichtiges Werkzeug, ne wenn man so auch in die Regulierung guckt, man versucht ja der Regulierer bis hin zur EU. Auch das Thema Open Data voranzutreiben und die Unternehmen dazu zu animieren, möglichst viele ihrer Daten zu teilen, so dass man darauf gemeinsam lernen kann und bessere Services und Schnittstellen etablieren kann. Aber das mache ich natürlich nur. Wenn also, wenn niemand verschenkt, sozusagen seine Daten und sein Kern Asset weil davon am Ende mit dem Unternehmen und die Prozesse werden immer datengetriebener, wenn ich das irgendwie nach außen stelle. Kann ich mir natürlich die Frage stellen ist das ne gute Idee für mein Geschäftsmodell? Also ich bin zum Beispiel bei der, wenn wir in der Energiebranche gucken, bei einem großen deutschen Stromhändler unterwegs, die sich genau gerade mit dieser Frage beschäftigen wenn du Solaranlagen hast und Geräte, die die Daten sozusagen miteinander austauschen, wo kommt gerade Strom, wo ist der Verbrauch und wie wird das ausbalanciert? Ist ja auch die Frage sind das Daten, die man vielleicht mit anderen auch teilen möchte, was ja noch ein paar mehr Anbieter und Lieferanten in diesen in diesem Netzwerk gibt und das Macht bestimmt total viel Sinn, aber es gefühlt traut sich da auch keiner im Moment das zu tun, weil das sind einfach wichtige Daten, die zu den Kernprozessen gehören, die gibt man hier aber aus der Hand.

26 [0:20:37] **I:** Ja, das hab ich auch gelesen, dass das so ne im Energiesektor so n dezentraler Sektor geworden ist und mit man muss die Netzstabilität Stabilität gesichert werden und das wär vielleicht möglich mit sowas. Und würdest du Vertrauen immer noch als wichtig einschätzen auch mit Privacy enhancing Technologien. Oder glaubst du durch solche Kontrollmechanismen ist es weniger erforderlich.

27 [0:21:05] **B8:** Ja also. Aus meiner Sicht, das ist sozusagen meine Kernbotschaft meiner täglichen Arbeit ist diese Vertrauensfrage zu beantworten extrem wichtig und das wird auch nicht weggehen, aber es wird sicherlich helfen. Als vertrauensbildende Maßnahme aus technisch sozusagen technischen Instrumenten zu implementieren, die das Vertrauen gewährleisten. Also ich muss quasi

nicht raten, sondern ich kann mich systemisch darauf verlassen, dass dieses Vertrauen gewährleistet ist. Und das wäre natürlich das würde helfen, also da, offener zu werden.

28 [0:21:41] **I:** Ja ja und siehst du irgendwelche speziellen Sektoren, Unternehmensgrößen, Regionen wo Privacy enhansing Technologien am ehesten oder wo Unternehmer am ehesten bereit sind, in solche Technologien zu investieren.

29 [0:22:02] **B8:** Das kann ich tatsächlich nicht gut beantworten. Ich könnt mir sozusagen vorstellen, dass es bestimmt sozusagen. Ich habe idee, warum kann ich das nicht gut beantworten, weil ich den Eindruck hab, dass ich da einige Unternehmen noch gar nicht an dem Level befinden, sich damit zu beschäftigen und dich deswegen noch nicht so viele Berührungspunkte in die Richtung hatte. Ich glaube, das Potenzial gibt es in allen Branchen sind eher Unternehmen, größerer Ordnung oder? Ich würde es nicht mal in einer Unternehmensgröße festmachen, wo dann einfach die sehr viele sehr viel mit Daten zu tun haben. Da würde sich das lohnen. Es ist ein Wissensproblem, also selbst bei Confidential Computing tun sich viele noch schwer, sich damit zu beschäftigen.

30 [0:22:49] **I:** Ja, und siehst du in welche neuen Risiken, die entstehen könnten mit solchen Technologien?

31 [0:23:04] **B8:** Ja, also das ist glaube ich dann der Klassiker ne, man muss natürlich am Ende sicherstellen, dass auch das irgendwie gut implementiert und sichergestellt wird, damit das Vertrauen auch gerechtfertigt ist, also sozusagen die die Qualität des der der Algorithmik und auch der technischen Umsetzungen nachher die muss natürlich irgendwie wasserdicht sein. Das Risiko, dass die Technologie doch nicht so funktioniert wie gesagt, ist schon auch nciht zu vernachlässigen.

32 [0:23:27] **I:** Alles klar, Ach so ja, die letzte Frage wäre noch was wäre die größte Herausforderung, die du siehst für die Akzeptanz von solchen Technologien auf dem Markt allgemein.

33 [0:23:45] **B8:** Waren wir wieder beim Thema Vertrauen sind ne es ist n Stück weit ein Nachweis dafür, dass das gerechtfertigt ist, dass das gut funktioniert. Also ne wenn es ausreichend Belege dafür gibt, dass man dieser Technologie gut vertrauen kann und dann da bin ich noch nicht auf Paper Ebene irgendwie unterwegs was die Algorithmik anbetrifft, sondern auch, dass nachher die Umsetzung funktioniert. Von technischen Systemen, die das implementieren. Aber ich glaube, das wäre eine Grundvoraussetzung dafür, dass da nicht nur sozusagen wissen aufbauen notwendig ist, sondern ich muss auch systematisch nachweisen können, dass das Vertrauen gerechtfertigt ist, also um mal so ein Beispiel zu machen ich den Tisch oder so, der hätte das mal eine Zeit lang sich mit

Confidential Computing beschäftigt und halt Services angeboten die. Sehr hohe Vertrauensstellung einfach aus Infrastruktursicht schon bereitstellen können. Und haben versucht das dann auch systematisch nachzuweisen das ne? Durch Prüfung durch Audits, durch Tests dass, dass das tatsächlich funktioniert, das ist dann klar am Wissen gescheitert, weil die meisten noch nicht verstanden haben, was sie da, was sie davon haben.

34 [0:24:56] **I:** OK, aber das ist also nur einsetzbar, quasi, wenn es ganz viele Anwendungsfälle schon gibt, wahrscheinlich wo man drauf verweisen kann und einfach es muss irgendwie etabliert sein, damit die Leute dem irgendwie überhaupt vertrauen können.

35 [0:25:20] **B8:** Ja, also ne, das ist ja eh so n Henne ei Problem. Aber man man sollte von vornherein sicherstellen, dass man das irgendwie, dass es irgendwie prüfnachweise gibt, dafür sowohl organisatorisch als auch technisch als auch algorithmisch so und dann werden die ersten schon anfangen.

36 [0:25:39] **I:** Ja, ja, gut, alles klar vielen Dank für deine Zeit.

37 [0:25:44] **B8:** Ich hoffe, es hilft dir hilft dir ein bisschen, wenn du noch fragen hast meld dich gerne sonst auch im Nachgang und das Transkript schicke ich dir zu.

38 [0:25:52] **I:** Aha, ja, mach ich. Perfekt, danke dir gut, dann mach es gut schönes Wochenende.

39 [0:26:00] **B8:** Danke wünsch ich dir auch tschau Felix.

Interview B9:

| | |
|---|---|
| Interview-Nr. | 9 |
| Date of the interview | august 28, 2024 |
| Duration of the interview | 14:59 min |
| Interviewer | Felix Starnecker (I) |
| Interviewee | B9 (Switzerland) |
| Role | Founder \| CEO |
| Sector | IT sector (Privacy Enhancing Technologies) |
| Specialities | No specialities |

[0:00:21] **I:** Also aus deiner Erfahrung mit wem teilen Unternehmen ihre Daten?

[0:00:49] **B9:** Also das hängt natürlich sehr stark von der Industrie ab. Mhm also weißt du, ob es im Medienbereich ist, ein Gesundheitsbereich im öffentlichen Bereich? Deswegen gibt es da nicht wie heißt das o One Size fits all? Also wir sehen, dass es im Medienbereich zum Beispiel, dass hier, wenn die Werbung schalten wollen, zum Beispiel ist. Adidas mit einem Medienunternehmen wie RTL zusammenarbeitet, im Gesundheitsbereich Krankenhäuser mit Pharmafirmen bezüglich Forderung oder Universitäten und teilweise auch Krankenkassen genau das ist so das Einzige, wo wir auch Competitors oder Wettbewerber zusammenarbeiten sehen, ist im Bereich Betrugsbekämpfung, also als Versicherungsbetrug und so weiter. Ah ja, ok, ganz dann auch vielleicht so Cyber Security Themen, wo Leute dann zusammenarbeiten ja, das haben wir auch schon gemacht mit dem Schweizer mit der Schweizer Militär, mit der Börse und einer Bank und der Nationalbank.

[0:02:01] **I:** Mhm OK und gibt es dann irgendwelche Vereinbarungen, vor allem jetzt wenn man mit Konkurrenten zusammenarbeitet?

[0:02:13] **B9:** Hat natürlich, je nachdem aus juristischer Sicht wenn du Daten verarbeitest, auch wenn sie verschlüsselt sind, ab und zu ich sag jetzt mal werden halt alte Gewohnheiten angesetzt, also aber in der Regel sind die Verträge eigentlich hauptsächlich, um kommerzielle Dinge zu klären.

[0:02:30] **I:** Okay und würdest du sagen, das Vertrauen und Kontrolle wichtige Faktoren sind die Unternehmen dazu bringen, Daten zu teilen oder auch nicht zu teilen?

8 [0:03:07] **B9:** Die unemotionale Rechtkomponente zum Thema Vertrauen das Vertrauen ist gut, Kontrolle ist besser. Das man natürlich eine Lösung haben möchte, wo du kryptographisch sicherstellst, dass auch wirklich nur das gemacht werden kann, was auch abgemacht ist.

9 [0:03:46] **B9:** Zu deiner Frage, on ein subjektiver oder objektiver Entscheidungsprozess, also, da geht es immer Business Case getrieben, okay und Gefahren und Risiken, die du in Datenteilen siehst ja, ich sehe eigentlich das größte Risiko, es nicht zu machen Mhm. Weil also auch ein bisschen ich sag jetzt mal ganz gute Thematik mit AI die wenigsten Firmen haben die entsprechenden Datensätze itern zur Verfügung. Das heißt, sie müssen auf externe Datensätze zugreifen können und da sind natürlich datenteil datenkooperation extremst wichtig. Und ich glaube, es ist auch immer wichtig, sich mit neuen Technologien frühzeitig vertraut zu machen, weil wir merken es jetzt schon, wie es schon Business Standard wird bei gewissen Firmen oder auch Industrien und dann Firmen also du weißt ja wie es ist, weißt du gibt es irgendwelche neuen Tools und so, da willst du nicht auf den Knopf und alles funktionieren also vielleicht funktioniert das Produkt aber die Firmen wissen ja auch die Leute haben mit der entsprechenden Expertise.

10 [0:05:26] **I:** Und siehst du interne oder externe Einflüsse für Unternehmen ihre Daten zu teilen oder auch nicht zu teilen?

11 [0:05:39] **B9:** Externe Einflüsse natürlich, wenn es Marktveränderungen gibt, zum Beispiel eben Online Marketing. Dass das Tracking immer schwieriger wird durch Party, Cookies wegen Apple und Google? Und das ist natürlich dann ein externer Faktor für europäische Firmen mit Daten mit eigenen Daten zu kollaborieren. Gesetzliche Themen es gibt ich glaube, ab 2026 oder 27 ein Gesetz für die Pharmaindustrie, um auch Real World evidence Daten für die Zulassung zu nutzen.

12 Müsstest du mal recherchieren, das ist natürlich dann auch noch mal so ein Faktor genau und intern absolut also du hast natürlich wie überall ich sag jetzt mal die Innovatoren, die Slow Adopts genau. Spannend finde ich, dass es meistens die Marktführer sind, die da voraus sind. Vielleicht haben sie einfach dann mehr Zeit, sich um sowas zu kümmern.

13 [0:07:29] **I:** Und was wären Herausforderungen oder Gründe, wieso man keine Daten teilt?

14 [0:07:43] **B9:** Ja, es gibt natürlich ne Ressourcen weißt du jetzt ein neues Projekt zu stellen? Innerhalb einer Firma ja, das der Cache einfach nicht da ist, also weißt du, es macht auch wirklich nicht überall Sinn ja, wir merken auch, dass es Firmen gibt, mit absurden juristischen Ansprüchen also ich hab kürzlich mit dem großen deutschen Konzern gesprochen die meint „Oh, wir würden das

nie machen“. Dann habe ich gesagt, kennt ihr das Unternehmen aus 20 Kilometer von [Region] ja, das ist das Größte in der Region, und die nutzen unsere Plattform. Dann waren sie offener. Also verstehst du, ich nenne das immer so ein bisschen emotionale, also fast schon religiöse Züge weißt du, da kannst du dann auch nicht mit rationalen Argumenten kommen. Es ist komischerweise gerade von Leuten und von Firmen, die nicht so technologisch sind, also weißt du und die eigentlich gar nicht zu verlieren hätten also es sind dann aber hier die mit den großen Daten setzen und die die Technologie getriebenen Firmen, die eigentlich eine viel höheres Risiko hätten, aber die halt die Technologie und so weiter verstehen und dann sagen ja easy macht Sinn?

15 [0:09:11] **I:** Und welche Anwendungsfälle würdest du für PETs, speziell SMPC sehen?

16 [0:09:21] **B9:** Also wir haben uns auch damals, als ich die Firma gegründet habe, ganz bewusst für Confidential Computing entschieden. Aus folgenden Gründen, Grund 1 ist es Enterprise Production ready also, das heißt, es kann eingesetzt werden. Grund 2 es wird vor den größten Tech Firmen favorisiert eingesetzt, also zum Beispiel jetzt gerade kürzlich, weiß nicht, ob du dieses Announcement von Apple gesehen hast. Mit Open AI mit diesen Modellen mit Secure Computing, das ist genau grad wieder zur Verfügung, oder es macht jetzt Apple auf jedem neuen Handy. Dann Microsoft als Cloud führender Cloud Anbieter hat er entsprechend in groß investiert und auch Google die komplette neue Werbeinfrastruktur mit 2 Milliarden Umsatz.

17 [0:10:00] **B9:** Und ja, wir freuen uns da technologisch auch spürend mitführend zu sein. Zu deiner ersten Frage, also für welche use cases? Ich glaube also in einer perfekten Welt werden natürlich Daten Slash Informationen immer maximal sicher verarbeitet? Ist natürlich in der Praxis noch nicht gängig, aber ich glaube um einfach sämtliche Risiken von Data Breaches, Missus of Data so klein wie möglich zu halten. Das ist so ich, wir glauben auch irgendwann, dass das so wie heißt das? In Poker nennt man das Table Stacks Steaks also Industriestandard wird genau.

18 [0:10:51] **I:** Und wie könnte sich das Daten Austausch Verhalten durch PETs ändern?

19 [0:10:59] **B9:** Sagen wir mal, es ändert sich zum einen, dass Datenkollaborationen, die vorher unmöglich waren möglich werden. Ja, und da, wo sie schon stattfinden potenziell effizienter, weil du nicht mehr irgendwie weißt irgendwelche Mittelsmänner, oder wie heißt das Treuhänder und sowas brauchst also es wird viel dynamischer. Und dass du natürlich Daten nutzen kannst, die vorher zu sensitiv waren, das ist schon extrem groß und durch Kollaboration auch Tools schaffen kannst.

20 [0:12:01] **B9:** Bezüglich den Faktoren Trust and Control hatten wir ja oben schon ein bisschen

gesprochen, dass du natürlich vertrauen quasi technologisch garantieren kannst, also zumindest vertragliches vertrauen. Es gibt natürlich immer noch das emotionale Vertrauen oh ja, keine Ahnung das Internet ist böse. Aber wir haben ja auch nie Zugang zu den Daten, das können wir technologisch garantieren, also sie müssen uns auch nicht vertrauen, und sie müssen der Technologie per se auch nicht vertrauen, weil sie das Kryptographisch oder die können das verifizieren also. Das heißt also wenn wir präsentieren und sagen wir hey? Gucke ich gerne den Code an, mach die Tests und das ist natürlich extrem stark, ne?

21 [0:13:11] **I:** Und welche Firmen/Sektoren glaubst du investieren am ehesten in solche Technologien?

22 [0:13:19] **B9:** Ja also ich glaube, es sind natürlich eben diese Daten betrieben sind ne wo du natürlich solche Sachen überhaupt brauchst. Dann Größere das ist wir arbeiten nur mit Großfirmen zusammen, deswegen kann ich das jetzt nicht für uns beurteilen, da habt ihr, glaube ich auch Health Sektor oder genau hast du gesehen Nummer 1 im Medienbereich, gefolgt von Gesundheitsbereich. Regionen ist spannend also auf der einen Seite die größte Anbieter, davon kommen aus Amerika, aber wir finden Europa eben ein spannender Markt, weil wir regulatorisch einfach einen anspruchsvolleren Markt haben als Amerika, wo man natürlich mit PETs mehr erreichen kann.

23 [0:13:56] **I:** Und gibt es neue Risiken, die durch die Nutzung von PETs entstehen? Die es vielleicht davor noch nicht gab?

24 [0:14:05] **B9:** Ja, ich sag jetzt mal es ist wichtig ein gewisses Grundverständnis aufzubauen und eben die Frage es gibt ja verschiedene Technologien und die sind auch für verschiedene Probleme die richtigen und das muss man dann auch entsprechend kombinieren. Bei einer falschen Nutzung entstehen natürlich Risiken. Und zu den größten Herausforderungen einfach das Problem mit den First Movern. Es wird immer einfacher, sobald eine gewisse Anzahl umsetzt.

25 [0:14:53] **I:** Ja perfekt, vielen Dank. Dann haben wir sogar doch in einer Viertelstunde geschafft.

26 [0:14:59] **B9:** Ja, kein Problem. Meld dich gerne, wenn noch Fragen aufkommen.

1 Interview B10:

2

| | |
|---|---|
| Interview-Nr. | 10 |
| Date of the interview | september 14, 2024 |
| Duration of the interview | 41:14 min |
| Interviewer | Felix Starnecker (I) |
| Interviewee | B10 (Netherlands) |
| Role | Founder of cryptotech-company; Employee in research facility for data sharing |
| Sector | IT-Security (Privacy Enhancing Technologies) |
| Specialities | No specialities |

3 [0:03:51] **I:** Yeah. So my first question would be like from your experience with whom do companies share their data like is it like with competitors, suppliers, customers and for what purpose?

4 [0:04:07] **B10:** Yeah, it's a very broad question. But when I look at professional technologies, I'm always make the distinction between sharing within the company and sharing between companies because I know that also companies share a lot of information between different parts of their company, so different subsidiary, different regions, different brands within each multinational. And I think that's also a perspective you could easily incorporate. But typically, companies we see a lot of sharing in supply chain. So as the so you have a supplier that supplies you something and then you have a buyer that you, you sell it to. So that that's typically in supply chain, but there can also be. Organisations around it that, like regulators that you need to share data with. There's also a sharing between competitors. Sometimes is more like. Doing sort of statistics together. Where was tracking KPIs? Or I think it's mostly on the. North core business stuff, so bank sharing data on. I don't know security topics, for instance.

5 [0:05:26] **I:** So you mean like no sensitive data? More like general?

6 [0:05:29] **B10:** No, I think it is quite sensitive data because you're sharing, my bank is now we're under, we're being hacked currently by this and this and this virus. They share those kind of signals in the Netherlands, there's a nicer use case about it where they actually use secure multiparty computation for that.

7 [0:05:46] **I:** For cybersecurity issues, you mean?

8 [0:05:51] **B10:** Yeah. So it's cybersecurity risks that they share and it's called secure net by SC. What says it NCSA. I'll type it in the chat so you can.

9 [0:06:05] **I:** So they share when they get attacked by somebody and then they share to other companies.

10 [0:06:09] **B10:** Yeah. Or sort of questions like have you in the last year were you ever the victim of ransomware? And then did you also pay the ransom and how much was it? That kind of data it's sensitive, but it's more sensitive in sort of a reputation. We typically say this is more reputational risk than competitive risk or privacy risk. But that's actually also a reason for companies to share. And I think in healthcare, for instance, companies share a lot of data around effectiveness of their healthcare solutions or hospitals share a lot of KPIs with insurers, but also amongst each other to create benchmarks for instance.

11 [0:06:46] **I:** Yeah, OK. Mm hmm. And would you consider a trust or control more important for companies to share the data or would you say they are, is trust important for companies?

12 [0:07:03] **B10:** Yeah, depends how you define trust. Would be my sort of answer that we cannot do anything with but the.

13 [0:07:14] **I:** Yeah, like, I don't know. I've read much that maybe like European companies are more likely to share the data with other European companies and not like more reluctant to share it with maybe Chinese companies.

14 [0:07:29] **B10:** Yeah, I think because I worked a lot in healthcare and people in healthcare are quite open to sharing data. It's not so unusual in healthcare that you share data for research or for improving. It's more sort of sector with also sort of a societal benefit in mind. But also banks would really like to share data about fraud, for instance. But sometimes they're not allowed to. But they should want to share that data with other banks. But then banks that they try, I think there's some trust like if I look at the Netherlands, there's a couple of big banks and they're quite happy to share. But there's maybe some external players or new incumbents or maybe FISA is becoming too powerful. And then so it's more like they sort of sometimes form this. I know, yeah, it could be sort of more of like a coalition thing. I want to share with people that I think are a bit similar to me and I don't want to share it with people that are very different. So just like people in everyday life, I think companies sort of behave a bit the same sometimes. Same regions, common goal helps all.

15 [0:08:39] **I:** So you would say it's not trust, it's more like companies who are more same.

16 [0:08:45] **B10:** Yeah. Or are you already used to maybe there's a sort of a degree of cooperation in a

sector where it's maybe a degree of competition, some sector are like cut throughout competitive towards each other and some sectors are more cooperative. Also there's regulation, so some sectors are really not allowed to like in telecom or in banks. All these regulatory there are some regulated sectors. It is also a regulatory risk to try to share data because are you not sort of creating unequal competitive playing field.

17 [0:09:20] **I:**Mm hmm. And would you say if like companies think about sharing data, is it a subjective or objective process when they decide like is it the top management?

18 [0:09:30] **B10:** I would say it's subjective and I think I learned it through privacy, non technology because if you start then they say oh, we cannot do this. We want to do this kind of data sharing within, I don't know within the supply chain or in healthcare between hospitals and insurers. But privacy regulations make it impossible. And then we come with privacy technologies and with lawyers and we say look, these lawyers and this technology and, you know, says, hey, you can do it, it's not. And then you go layer deeper and it well. But there's actually some personal distrust or just people. Yeah, just all the regular stuff at cooperation between companies. It needs to be right on all the different aspects and not just if we've solved some privacy issues, then all of a sudden the data will start flowing freely. You need to have it all on all those topics. So we need to have the right contracts in place. They need to be trusted, but also sort of trust at the board level and personal trust. These people should like each other and if they don't like each other, it's much harder to get something going. So yeah, I think I learned it's way more complex than just one thimble trust the thing.

19 [0:10:56] **I:** Are there always contracts involved when sharing data? Like is it always with?

20 [0:11:00] **B10:** I think there if you go into privacy nonsing technologies, I think only in synthetic data generation sometimes it's done without a contract. I know some cases like the Dutch. I'll write it down Ika and I'll synthetic data set. IKIKNL is the Dutch cancer registration, so they have all the patient data of people with cancers and different. It's a super very good data set. They created the synthetic data set and you can just download it or maybe you need to click three buttons or something, but then you can just get this synthesised data set. They do it to further Cancer Research, but also to on their own side. There's also some internal incentive. I think they have quite a they were sort of overflow with requests by researchers saying, can I have access to this data. So they built sort of a two stage approach where at first you get the synthesised data set. Then you can play with it and see if it's really something that fits your research. If yes then you have a very long and tedious formalised process to get the actual data. But it means that there are people have more time to work on the they can help to help researchers with a good case better because they're not distracted by two hundreds cases that are actually that would have nothing to gain by getting this

data.

21 [0:12:25] **I:** Yeah, yeah. And what risks do you see when companies share data or what are their biggest risks or are there risks?

22 [0:12:45] **B10:** Yeah, I think if it's personal data, then there's a privacy risks. That's a risk that you would reveal. Some personal information and that people could edit and could get fined for that, for instance, in by the privacy authorities. The bigger risk if it's sensitive personal information, so that's medical records, but also like police records. And there's a lot of categories in the GDPR when data becomes more sensitive. Data about your religion, data about your race, data about if you're a member of a union yet or no, those kind of. Yeah, that's extra. Have you ever been in gaol?

23 [0:13:27] **I:** Then you have regulatory. Yeah.

24 [0:13:31] **B10:** That's sensitive data. You don't want people to learn that your genome. You can think of so many aspects, and then there's more competitive risk. So if my competitors would learn this information, they would be more wiser and they would and I'll read it, come with different prices or gain some market share or. Those kind of things. And then there's reputational risk. I think those are the in my mind, the three biggest risks.

25 [0:14:02] **I:** And the biggest risk just depends probably on the sector where you are in.

26 [0:14:07] **B10:** Yeah. So in, I think in healthcare, it's just sort of a data leak. I think they're scared of the financial consequences, but I think maybe even more of the reputational risk just sort of hospitals. I don't know screwing up with your patient data just is yeah, it makes them look bad as AI think. I think that you get more into the will drives will drive the CEO of hospital. But they typically don't want their hospital to be in those kind of negative press and it's for CEO quite important.

27 [0:14:46] **I:** And do you know any like external or internal factors which influence companies when they share data like? Or like external probably maybe lost or something. This direction and internal would be the culture of the company.

28 [0:15:06] **B10:** Oh yeah, yeah. Yeah, I think externally it's mainly laws, but then of the laws also get put into sort of companies specific policies. So you have the GDPR which just says something in general about how you should do data processing and then a company would make its own policy based on the GDPR. That's typically quite sort of. These are the 20 steps you need to do, and these are the 50 people you need to talk to before you can send any data outside of the company so that the whole process becomes quite complex. And I think sometimes companies also make these processes even a bit too strict or too. And they really want to avoid those risks. So they put a lot of

process in place which just also makes it expensive and slow and frustrating for people. So yeah, so that comes from the law. But then on the other side I think there's also some risk averseness in culture that can be hard for data sharing because it's quite easy to say, well, there might be some risks here, so let's just not share anything. I've seen a lot of those meetings where you typically there's already four people in.

29 [0:16:32] **I:** But if you're not here, you probably have done different risks, like maybe supply chain risks or other risks?

30 [0:16:41] **B10:** Yeah, but that's usually already. And maybe that's also one of the problems I think, for privacy enhancing technologies, it's not that we're now not sharing data. If there's a lot of cases where actually data is already being shared and there are a lot of cases where people think it's a really bad idea to share data and then privacy technology or somewhere in the middle and they can sort of open up a few of the cases that were previously impossible but now possible. Or you can say some of the data sharing that was a bit risky, they can maybe derisk that. And that's also what I saw in healthcare that there's quite some data sharing going on where people just do it very sort of the old fashioned way. They just e-mail each other Excel files because they've already been doing that for 10 years. So every month they send in an Excel file and there's some contract behind it. That was 10 years ago, quite a good contract and now you could start saying well this is beginning to feel a bit risky. But as long as no one has no one sort of raises that alarm bell in the company, then they can keep doing it for two more years. I think that's sort of it. It's sort of quite a narrow sport where perhaps technologies all of a sudden make stuff possible that wasn't possible before.

31 [0:18:06] **I:** Yeah. This would be like the next question for which use cases. You think privacy enhancing technologies like smpc are most useful, or it could be practical.

32 [0:18:20] **B10:** Umm. I think that it depends how you want to define a case 'cause if it's sort of super sector specific.

33 [0:18:28] **I:** Yeah, because I often hear that like many people would say, the technology is nice or sounds nice, but then they don't see any practical use case where you could really make use of it.

34 [0:18:45] **B10:** Yeah, I think. I suppose that what is that you can do with this technology and I think one of the things is to single out individuals without. Who have certain characteristics without learning anything about all the other individuals? So like if you want like fraud detection stuff, say OK, I only want to know the 10 people who have all the fraud signals and not all. Not the 10,000 people who are. And I think there's a nice example in the Netherlands for that I always use. It's we have gun laws there. So people are allowed to have a gun. If you have a certain licence or you're a hobby shooter or a hunter or something that you can have a gun. Like very few people have guns

and also some people have. What's the word for that? A psychosis? So there's a second like a depression, clinical depression. Sequos I think probably in German. If you have both, you have a problem. In that case, they will come to your house and try to take your gun away. Which means like but 99% of people with a psychosis don't have a gun, so it's really very unnecessary that their data is being shared with police, because if police learns that you have a psychosis, they might use that in in criminal investigations in these neighbourhoods. I know there was a murder. Maybe it's the guy with the psychosis who did it, that's sort of the risk that you didn't run. And then with MPC, you could sort of single out. And that you only get three hits a year of people who have both and not learn all the other patients with the psychosis. As the police, I think that's a nice sort of ethical case. But. But in reality, in Dutch law, it's already allowed for this data to be shared with police. So it's not that we're doing this with MPC now, we're just doing this as we speak. This data is being shared with police now. But I think it's for this sort of singling out that's quite interesting also for fraud case, if I want the the only concern a few people and not too many and then you can just make it more privacy by design. And also in other cases, we're sort of getting statistics or getting like shared risks. I think that's also a good thing, like shared risk scores over portfolios or we have so many bank accounts, what are the top 100 bank accounts that are of each other? Other organisations also feel that they're quite the dodgy, so we should investigate only those and not the others. That's one sort of aspects. Then the second aspect is doing statistics. Sonoverall overarching statistics over different organisations, which is.

35 [0:21:58] **I:** For research, for example, or for health.

36 [0:22:00] **B10:** Yeah, for research. For policy. That's my company was more in that field. So we did. We had healthcare data from hospitals, healthcare data from insurers and healthcare data from municipalities and we could sort of do the overarching statistics. So we could say people who you got a certain intervention from the municipality, they actually led to lower health care costs at the insurer. So this is really good policy. So please, let's keep doing that policy or you can measure. I think in the Netherlands now there's quite a debate about how would you call that, let's say waiting lines for hospital. So waiting times for mental health, for instance, there's quite a long wait time. But then people start applying to a lot of different clinics, so no one knows how much. So if every person applies for four clinics, we think that the actual waiting list is four times too big. So that's kind of that kind of sort of statistic. It's we call it, secondary use of data in healthcare.

37 [0:22:39] **I:** OK, it's interesting because my last interview partner said that he doesn't see MPC at all for healthcare or not so much because it's more like the data is too sensitive and they're too hard barriers to. I don't know the people are working there, don't trust this technology so much, I don't know, he said.

38 [0:23:37] **B10:** Yeah, I would say otherwise 'cause I've been in the projects where we actually share this kind of data.

39 [0:23:38] **I:** But I also thought it's very useful for this because yeah, you can do so much with it.

40 [0:23:53] **B10:** Yeah, I could send you some slides, et cetera about what's going on in. But it's all in Dutch. I don't think that's 'cause that the Healthcare is. It's very sort of German Healthcare is in German and Dutch Healthcare is in Dutch, so it's. But in the Netherlands there are a couple of life examples where multi party computation is being used in healthcare.

41 [0:24:16] **I:** Hmm. Yeah, yeah, I think Netherlands is also like a bit further with healthcare and data. I'm not sure if it's there.

42 [0:24:24] **B10:** Yeah. I think I have the feeling that in Germany people think the best privacy is to just share everything. Had to just do not share, like the best if you want privacy, you pay with the bargirls and then you have the ultimate privacy. And in the Netherlands we are more. I think we see more than in Germany we see technology as a solution rather than a threats or something. So typically the Netherlands is a bit quicker to accept like mobile payments, etcetera. And then sometimes we learn the hard way that there is indeed a privacy risk.

43 [0:24:39] **I:** Yeah, I would say it was like this.

44 [0:24:59] **B10:** But I'm always confused. If you go to Berlin that you need to bring cash 'cause. I don't even have a wallet. I don't even have it anymore and I never have.

45 [0:25:10] **I:** Yeah, but in Germany you need it. I needed it today and I didn't have it, but yeah. My next question would be like do you think that the data sharing behaviour of companies will change in the future with privacy enhancing technologies like? Is there more shared in the future or different data types?

46 [0:25:34] **B10:** Yeah, I think.mI think the general trend is that more data will be shared in the future. Maybe not even so much because companies want to, but also because it will be mandated by the EU. So there's a lot of laws, like the data Act, but also very sector specific acts like the European health Data Space Act or the, I don't know, in insurance it's a FEDER FIDA. That's an act. That's an act that just forces insurers through or financial institutions throughout Europe. To open up and to start sharing data, and if it means I can go to my bank and I said, look, I have this app and this app needs my bank account data. You're now mandated to give the data to me so I can put it into this app so that will open up. I think a lot of data sharing much more than privacy technologies will.

47 [0:26:28] **I:** OK, but then you have to share data normally not with like smpc or something like this or?

48 [0:26:40] **B10:** Yeah, it doesn't really specify how you share it. Typically it's about sharing data to the owner of the data. But also for feeder, I think it's more like I give the mandate. If I want to have my bank data in an app for I know and that sort of would analyse my spending behaviour and then give 10 saving tips for those kind of apps. Then I would mandate the app to on my behalf, go to the bank to request that data, but then sort of the law sort of specifies and how you can get that mandate and the banks are obliged to comply to the mandate, but then it will be just raw data. It will not be NPC Ed or anything 'cause then you would learn less. So some of these technologies, some of these laws specify how you do it and typically they just say share it just in the blank, just in the plane. So I don't think it will help. What I do think is that GDPR has some clauses in it where it says that you should take. Yeah, you should. You should always strive for maximum amount of privacy given sort of the technological capabilities of that are current, so and once MPC sort of becomes more and more current at some point GDPR will always say oh it's true. Why are you not using privacy enhancing technologies? Look because all the other hospitals are using them and now there's like this turning point in GDPR. But it's sort of it keeps raising the bar.

49 [0:28:00] **I:** OK. And then privacy?

50 [0:28:18] **B10:** So I think that will be sort of on the long run if it gets fully incorporated.
And also where it sort of gives a direct gain like. The tech companies are using privacy technologies basically to sort of protect their monopoly a bit, and so advertisers will not learn that it's a concrete person, but they will just learn it's someone with an interest in MPC and then they start serving you ads, but they will never learn who you really are. So basically it defends their Kingdom a bit and there you see a lot of. I think all the big tech players are now using multi party computing homomorphic encryption. It's getting more and more into their. I can send you some links. There's a nice blog. And I just sent it to somebody else. Let me see if I can find it. Yes, this is it. I found this quite.

51 [0:29:11] **I:** Yeah. I saw that Microsoft and Google also like they offer. Like privacy enhancing technologies.

52 [0:29:21] **B10:** Yeah, I think and in iOS 18, there's a sort of a library for homomorphic encryption. So app developers can start using. Building blocks for, but if you go to this website, I think they've written like 7 they've analysed I think 7 big tech players and what are they doing in privacy nonsent technologies. It's quite interesting. But it's mostly in sort of the advertisement deck where I got a corner.

53 [0:29:38] **I:** Alright, OK, interesting. Oh yeah, but nice. I will look at it and you've, like, do you think the trust factor changes with secure multiparty computation that it's not so important anymore to have trust or is it still?

54 [0:29:46] **B10:** Yeah, I think it's also the you put some of your trust in technology. So you don't need to trust in people or in processors, I think that's quite important. But it's also one of the things we learned at links I think you see it at all these MPC companies that they start off by trying to sell multi party computing and they at some point also on their websites they start talking more and more about data, collaboration and governance and those kind of terms. Because it's sort of the mix. It's you also need to. And you still trust is also in that there are good contracts and that for instance with MPC. That you know that you still have this control you mentioned control organisations still want to have control. They want to limit how others can use their data. So we allow you to do compute an average over our data, but we'll not allow you to, for instance, a nice example for us was in healthcare. If you do certain statistics and then the population age, you want to know what are all the men older than 65 who are millionaires who live in a very small house. Then your population becomes very small, and you're talking about like 5 people. And then if you compute an average of five people, you already learn quite a lot about those five people. And so we had some rules saying, well, if somewhere in the computation the population dropped between 20 below 20, then it should be cut off and you cannot, receive the answer. Those kind of settings, I think.

55 [0:31:18] **I:** Yeah. So you need other control mechanisms as well to, yeah.

56 [0:31:46] **B10:** Yep because MPC is not gonna solve this for you. You can just make. It's quite easy to hack an MPC system. You just enter a data set, then you run all the computations, and you get some answers. And then you make a new data set that has one different field and you put it in and then you see what's the change? OK then I can see what's the effect of this. So I've learned. Actually, your data about this one field. At those guys, so it's actually so you should prevent that you can sort of keep messing with the system in such a way that you could sort of say that we do population this first 100 patients and then just from patient two to patient 101 and then from patient three to 103 and then you just stick it up and then you can basically compile the whole data set. So those kind of sort those could be sort of attacks you could say that those are attacks on the system and you want to prevent them? But also you want to have a contract that says if you're data scientist is doing this kind of behaviour. Yeah, I expect you to fire him.  Alright, so if we're cooperating, we won't. I want to have this contract in place. I want to have sort of in, in the technology some way that I can sort of have like a list of settings saying this is the allowed algorithm or these are the boundary conditions for the algorithm and I need to trust in the technology that's sort of the whole stack and not just the technology.

57 [0:33:06] **I:** Yeah, that would be actually my next question, if you see like other risks who emerge from privacy enhancing technologies, like for example.

58 [0:33:34] **B10:** Yeah, I know. Let me first answer. Maybe your examples are the same or different. That would be interesting. I think 1 risk is if you try to get everything sort of. If you use MPC it's basically black box. So you can sort of get into sort of algorithmic decision making without understanding why the algorithm says something. So that's I think that's a risk. There's also a risk in just the complexity of it. So it's hard to explain. It's hard to explain to lawyers. It's hard to explain to people, to citizens, it's hard to explain to management, it's hard to explain to people.

59 [0:34:07] **I:** You mean that they use it then wrong or just?

60 [0:34:10] **B10:** Yeah. So there's a risk of people using it wrong or understanding it wrong, making the wrong assumptions. I think that I would consider that a risk and I think also in well, the field that we were in, we for instance, we also did some on premise it stuff and it's just it's also an IT risk. It's companies really don't like. You messing in their IT systems with all these new technology that they hardly understand.

61 [0:34:42] **I:** Yeah, I thought also like about maybe data manipulation risk that maybe some if you share with a competitor your data and 11 puts maybe really good data inside and the other one maybe wrong data or bad quality of data, I don't know is this possible?

62 [0:34:52] **B10:** That's good. Yes, I would consider it another attack that you should make. Make a certain.
Let me think it could be if you use. Like sort of benchmarking for hospitals, but it could also be between companies for how green is your company in comparison to competitors or how big is the gender gap in your company compared to competitors? Yeah, it's quite easily to sort of just put in garbage data because you want to show that you're very good or you want to mess up. Yeah. So I think there's need a risk for manipulation. That's also I just gave you like sort of this stack. I think there's one more thing in the stack and it's also all the trails. So you should always also be able to. The sort of replay, a certain calculation, so there'll be like one year ago we ran this computation, the answer was 15. How did we then come to 15 if you would enter your if there's a dispute. Can we still sort of trace it back, saying, OK, but then? Oh, then you must have entered the wrong data because with the data set that you now say you put in it doesn't match with the hash that we stored and it will also give a different answer.

63 [0:36:15] **I:** Uh. And you cannot trace it back. Is it not possible or?

64 [0:36:27] **B10:** No, in our platform we had all the trails where you could sort of timestamp and also

so at least you could if people were willing they could jointly recreate a certain transaction, a certain computation. So you can always sort of prove that it's really happened in that way and that nobody manipulated the inputs for instance.

65 [0:36:34] **I:** OK, mm hmm. What couldn't conditions are needed for companies to invest in Western privacy enhancing technologies like which type of company is maybe more willing to invest? Like from the size, sector or region.

66 [0:37:08] **B10:** Yeah, I would say now big US and China tech firms are investing in Prussian technology. Also in countries like Singapore, et cetera. There's quite some things going on. Also, Singapore government for instance, is investing in nets from my experience. We then, as I've seen it in healthcare, so it's mostly the insurers who are interested because they really have something to win if healthcare becomes more efficient and effective, it's basically good for their business. But I think it's still quite complicated for small companies. So I think, and I think maybe even more importantly within the companies, it's typically the people who are investing in it are more are still sort of more the innovation functions rather than sort of the operational part of the company, so also the healthcare part is typically the innovation departments of the health insurers invested in these kind of projects. And then the next step was OK, now it works. Now let's set it. Let's get it into it embedded into the business side of the organisation A real business. And then there's also a lot of hurdles so. I would say it's not really.
It's definitely not a mature market or something.

67 [0:38:45] **I:** Yeah. Yeah. Mm hmm. And the last question would be, umm, what do you see as the biggest challenges for the acceptance of pets on the market?

68 [0:39:00] **B10:** I think the biggest challenge for pets is the fact that they're sort of collaboration technology. And so it means that it's you. You don't sell your pet solution to one customer, you need to sell it to at least two at the same time.

69 [0:39:14] **I:** Mm hmm. So you need also to convince 2 at the same time.

70 [0:39:16] **B10:** Yeah. Which is in terms of business development, that's always one of the two that has less of an appetite or says well, yes, super interesting, but maybe in 6 months 'cause now I'm full or busy or so that timing is hard. But also the fact that it's sort of, it has so many angles to it. So it has a tech angle, but it also has a collaboration or organisational angle it has. A legal angle. Yeah. So that means that you're actually not having two companies at the table, but you have for each company 5 different departments at the table. So it's ten, yeah. So you can just draw a matrix of this and you can say how many, how many lines are there over which people will not understand each other. And there are quite many. I think that's the major hurdle for pets.

71 [0:40:03] **I:** Mm hmm. And is it then necessary that maybe like the top management has to introduce it? Because if it's only doing like one department then it's not possible?

72 [0:40:23] **B10:** Yeah, I would say that before you start any project where you try to implement the MPC, you really should get top management support because they can manage sort of this cross department cooperation. And then because you need at least two or three, like sort of a coalition of the willing sometimes as people call it so you also need those people at the other department or sorry at the other company. So you need to have this buy in on both sides.

73 [0:40:49] **I:** Mm hmm. OK. Yeah. Thank you very much for the interview and your time. It's got a bit longer, sorry for that.

74 [0:41:00] **B10:** Yeah, no problem. Let's not. And when will you?

75 [0:41:07] **I:** Submitted.

76 [0:41:07] **B10:** Yeah.

77 [0:41:09] **I:** I think like hopefully at the beginning of October depends how many, how long it takes now with the next I need still to do like 10 interviews. Yeah. Depends how fast they get the people. But yeah, probably beginning of October. So if you're interested I can send you the if you want.

78 [0:41:14] **B10:** OK. Yeah, that'll be nice. Yeah.

1 Interview B11

2

| | |
|---|---|
| Interview-Nr. | 11 |
| Date of the interview | september 11, 2024 |
| Duration of the interview | 27:06 min |
| Interviewer | Felix Starnecker (I) |
| Interviewee | B11 (Belgium) |
| Role | Research Engineer at cryptotech-company |
| Sector | IT-Security (Privacy Enhancing Technologies) |
| Specialities | No specialities |

3 [0:01:41] **I:** So I would start like with the first part is in general about data sharing between companies. And the second part is about the influence of secure multiparty computation. And I think I will put the focus on the second part because I think your experiences mainly in secure multiparty computation, I would say. Yeah, OK. Do you want to explain like for short what's your working like? What are you doing?

4 [0:02:54] **B11:** Uh, so I'm a research engineer, which means that like I do some studies trying to find out like what? Like state, state of the art, research is applicable to the product of my company. So I'm looking for some scientific results, but I'm also responsible for implementing those results in part.

5 [0:03:15] **I:** Mm-hmm. OK. And yeah, to the first question. From your experience with whom to companies share data like is it? Is it competitors or suppliers or customers and for what purpose?

6 [0:03:43] **B11:** I think it's a bit of everything. Uh, so from the experiment I got my company, I think, uh, I'm not sure about competitors. I think uh, it's always, I mean currently for example we have use case where we have a big company and their network of suppliers. So they want to exchange information between them. Uh, the same also happens sometimes that the companies they wanna get some information from their customers. But there is also an additional thing that there might be companies that sort of have like different types of data sets and those data sets they overlap on like certain interesting data and essentially the companies they want to merge forces to extract some value from for both companies. So you can think of, for example, it's a super popular use case. I think you can think of companies that have some advertisement campaigns. Think of Nike, for example, and you can think of companies that provide actually placed for advertisements like newspapers, and I don't know social networks. Uh, like websites? So they want to collaborate

essentially to kind of pass information about, like some users that I don't clicked on some button but at the same time they won't buy Nike shoes something like this.

7 [0:05:18] **I:** Uh, yeah. OK, OK. Mm-hmm. And like in my research I identified a trust and control as like important factors for companies when they share data. Would you say that both is important or what would you say is more important like?

8 [0:05:46] **B11:** So I mean both are important, but of course trust is kinda like vague thing. It's really hard to quantify in any sense. And this is why actually people want to use, uh, secure multiparty computation or other privacy enhancing technologies because now they put trust into like fear of breaching the law, breaching the uh but yeah.

9 [0:06:20] **I:** Yeah. I asked so you mean with secure multiparty computation or something else? They don't have to trust anymore so much or.

10 [0:06:31] **B11:** They should trust mathematics, which is like much stronger, and the assumption than trusting like some lawyers.

11 [0:06:35] **I:** Yeah. And would you say the decision process of companies when they share data is more like subjective or objective? Like do they have? Would you say it's more like dependent on the top management, they just say yeah, let's do it all. Let's not do it. Or is it a long process with risk analysis before they are sharing data?

12 [0:07:05] **B11:** Four, it's a hard question because I think it's both, uh, so it's subjective in the sense that people understand the need for this ologist so they understand. I mean the problem is that you have the current procedures. Involve a lot of like legal action and a lot of time, so people spend a lot of money on that or on the other hand, there are these mathematical technologies like SMPC, working encryption like zero knowledge proofs and stuff that might be easier, right?

13 [0:07:26] **I:** Uh-huh. Yeah.

14 [0:07:43] **B11:** So there's no legal action. You can essentially in theory switch as fast as possible. Uh, but the thing is that, I mean, there's this objective side. And then once we as a company having like some pets technology, we start implementing something, then we stumble upon some subjective stuff, how user friendly our application is how actually engineers at the company where we want to deploy stuff. They perceive a technology where using whether they understand it or not, and am, I think like the problem currently is that people don't really trust the idea how it works.

15 [0:08:19] **I:** Umm. Because they don't understand it or.

16 [0:08:34] **B11:** Yeah, because they don't understand. It's not really like some common thing. Uh, and people are like sometimes confused really confused about like what they have to do and this is why we, yeah, we sometimes lose prospects clients on really subjective grounds just because people kind of do you have time to even yeah.

17 [0:08:49] **I:** OK. Umm, OK yeah. And what risks of data sharing do you see when you don't use privacy enhancing technologies like?

18 [0:09:09] **B11:** One factor it's always the problem that yet data can be leaked. Sometimes it's easy to identify the leak and the source of the leak, and then, like we have legal action on that. But it's not always possible and there's always this risk, which is kind of non negligible.

19 [0:09:34] **I:** OK. Yeah. And like when companies share the data, do you think there are some external or internal influences them which encourage to share data or like are also barriers not to share their data?
Like I don't know. For example, the culture of a company or the management, I don't know.

20 [0:10:02] **B11:** Uh. Sometimes we come across that there are some cultural things that the companies. But I think it's usually defined by loss by for example, there is a good example of your imagine like a big informational Company Apparating both in Germany and United States. So when Germany, there's GDPR. Uh, and yes, there's like, if you're not in California, there is essentially nothing. And the thing is that 2 entities of the same company, they can't really share data. Uh, so they have to kind of comply in a sense, they have to find like some third party maybe within that company that can do that due to G and stuff.

21 [0:10:56] **I:** Mm-hmm. OK. Then to the second part to privacy enhancing technology and especially secure multiparty computation for which use cases do you think privacy enhancing technologies, especially like SMS PC or like homophobic encryption or like useful especially?

22 [0:11:27] **B11:** Uh, it's a brown thing, but so essentially it's useful in every use case where we you have sort of. If we talk about the SMPC, if you have, uh, several beak entities with kind of significant computational resources and those images, wanna compute something on mutual data. So some common function? Uh, it might be auctions, for example. It's like canonical use case. Uh, it might be a data set intersection.

23 [0:12:09] **I:** Umm yeah.

24 [0:12:10] **B11:** Uh, it might be data retrieval in any form. It might be a training machine, learning algorithms and stuff. I think in practice the most popular thing, which also most employed one, is

this set intersection thing. Uh, because essentially every it private uh messenger, such as the signal WhatsApp, I messenger? Probably telegram. We don't know much about them. They have this functionality. Uh, just across like contacts on your phone with contacts already present in the database of that particular messenger. They have to technology. A homomorphic encryption is different in a sense that it usually assumed that there is one big entity that performs all the computation. There is weak computationally weak client that outsources the computation to that big entity and uh, I don't think so that yeah, like some very good use case was found for homomorphic encryption. The companies, they're struggling, I think.

25 [0:14:04] **I:** And do you think secure multiparty computation will change the data behavior, data sharing behavior of companies in the future? Like that they share data or share data differently. I don't know.

26 [0:14:24] **B11:** I hope so. I think that UM, so once the companies will understand that. They can share certain sensitive data without significant leakage. I think it will just speed up collaborations between different companies. And also opens up certain data sets that like some companies just have free to send somewhere now.

27 [0:14:56] **I:** Umm OK, but so do you think they share more data in total or just more sensitive data?

28 [0:15:05] **B11:** I think in total there will share more data.

29 [0:15:08] **I:** Hmm. OK. And how do you think I have a like? Yeah, privacy enhancing technologies influence on the trust and control factors like would you say that trust gets a less important?

30 [0:15:28] **B11:** Yeah, trust will be less than it's also a little bit tricky because it's a PC. It's not a silver bullet, and there's always the problem that results of SMPC collaboration might leak little bit more than one company wanted to leak to they're counteragent uh, I counterpart.

31 [0:15:39] **I:** Hmm yeah. You know, why? Do you mean like I didn't understand like.

32 [0:16:01] **B11:** So what could be a good use case? So let's say. Ah, there's that intersection thing, right?
So the companies they want to intersect their datasets. And so one company has like IDs of users, IDs of customers. Plus I don't know where salaries another company has the ID, the same IDs of customers, but like maybe different number of customers and I don't know like what maybe like their health data. I don't know whether there are SQL something. And uh. It might have both parties. They don't know what exactly is present in the data sets of their counterparts, right? And it might happen that essentially they completely overlap on ID of clients and think then it's like both parties

or one of the parties if collaboration is set up in this way, one of the parties will end up with two datasets like completely. And this of course will be bad, right? Because the entire date is just moved to another part and this is why companies are also a little bit reluctant to do with MPC in this way because uh yeah, you might lick much more. This is super nice use case, but even like if you lick just like some intersection but not a big one, but then like one of these collaborating parties collaborates with somebody else, also gets a little bit more data. So they essentially dealing immunized users like extract more information than needed, yeah.

33 [0:17:56] **I:** Hmm. And do you see like are there new risks like developing through secure multiparty computation or like this OK and what do you think?

34 [0:18:11] **B11:** Yeah, serious. Yeah, this is very exactly, yeah.

35 [0:18:15] **I:** Like I've already heard something about like data manipulation and stuff. If, like one of the players is one of the actors, is not like puts in the algorithm like some data which are like wrong or not the data quality what is expected? I don't know. Would you say this is a risk?

36 [0:18:39] **B11:** Yeah, it's pretty hard to avoid such behavior actually.

37 [0:18:47] **I:** OK, OK. Umm. And what conditions are needed for companies to invest in privacy enhancing technologies like for do you need anything like prerequests? And yeah, which type of companies? More willing to invest in pets?

38 [0:19:11] **B11:** I think big companies. Uh, we companies, as I mentioned, like big companies with big network of suppliers, they're quite interested and actually kind of eliminating these legal framework of handling like of sharing data with their suppliers and they would be happy to have like some unified system in place where they could do it easily. A sector like I think food production. So if you look at big companies, big consortiums that, yeah, food production, also household items and stuff.

39 [0:19:49] **I:** Food production. OK, OK.

40 [0:19:55] **B11:** Uh, like you can think of?

41 [0:19:57] **I:** By crazy if I thought more about like maybe automobile sector or health sector, because there's more sensitive data.

42 [0:20:05] **B11:** Uh, actually, I didn't come across those companies.

43 [0:20:06] **I:** Why would you say food production then?

44 [0:20:15] **B11:** Uh, because they have this thing that they have a lot of I mean, you can think of companies like Mars, for example. So Mars has, like a lot of different brands and a lot of different suppliers in every country. So there might be a small company in Germany I don't producing just one part of sneakers. I want some sugar for example or some stuff and for Mars it's important actually to retrieve as much information as possible from that company about specific qualities of that sugar. And maybe also find better supplier. Uh producing sugar and stuff, so they want us to sort of kind of get like, all the information about, like all possible suppliers and quality of your products. And then for them, it's quite a critical thing. Uh, there's like use case where I'm involved. It's [company] and uh, they, for example, they wanna contact us, test different sense for perfumes. Uh, like you can think like they're like multiple use cases and in this I'll come about production. I haven't heard that anybody wanted from those companies to incorporate this NPC. Uh, which I think is weird because they also have different suppliers and they have quite a lot of sensitive data.

45 [0:21:56] **I:** And the app.

46 [0:21:59] **B11:** The closest I think. Through that, maybe companies like Bosch. They invested already a little bit into SMPC. They opened a research group recently and I guess, yeah, they want to do something with it.

47 [0:22:20] **I:** OK. But I've also heard that big companies sometimes have their own like data sharing platform, haven't they?

48 [0:22:32] **B11:** It's surprising, but not really. It's the usually use what is called data clean rooms. These are other companies providing these because like legally those companies, they can't really set up this data sharing mechanisms on their premises, yeah.

49 [0:22:42] **I:** Yeah. So then only one last question, what do you see as the biggest challenges for the acceptance of privacy enhancing technologies on the market or reasons against investing?

50 [0:23:07] **B11:** I think technical issues and user friendliness. So the most efficient SMPC protocols, they assume three different parties involved. It's really, I mean, which means that yeah, for the company that wants to sell as MPC, they have to find the group of free customers. Uh, that wanna deploy this data? Uh, it's also possible to do with two parties in somewhat efficient way, but again the same problem have to participate time.

51 [0:23:49] **I:** And so you need to convince not only one company you need to convince many.

52 [0:23:57] **B11:** Yeah. But there are like specific use cases where it's not needed as I mentioned like this set intersection thing where you have messenger or some service that wants to intersect data of

their users with their datasets. And then provide the results to users, which is important. And then yeah, it's a good use case in this case. But also user friendliness. People don't really understand how it works. People are really reluctant to deploy this stuff.

53 [0:24:40] **I:** Because they don't know. They don't understand it.

54 [0:24:43] **B11:**Yeah, because they don't understand.

55 [0:24:44] **I:** And OK do you think is there anything else interesting for my topic? What I have maybe should have asked but not did ask.

56 [0:24:57] **B11:** What is the focus of your master thesis? Is it like economical aspects of that or I don't see?

57 [0:25:00] **I:** And it's more economical aspects here, not the technical only like which factors? Influence companies when they want to share data and yeah, like how privacy, enhancing technologies.
Umm change these factors? Maybe Yep. So, but it's probably that there's less trust needed and less risk involved. Maybe, but also different risks emerging, yeah.

58 [0:25:34] **B11:** I think from what I see is that the companies got a little push after GDPR was introduced, so they couldn't understand that. Yeah, I mean, we can't really handle data in the same way as we did before. Then the second push was when browse their started to remove third party cookies. And that was also quite important for the advertisement market in web. So I think it's mostly like the main motivation to avoid, uh, more time with some legal teams. Gotiations and staff sign certain documents, etcetera, etcetera. I guess it takes a lot of time. Maybe not that much money, but time is really except here.

59 [0:26:37] **I:** OK. Yeah. Interesting. Then thank you very much for the interview. Thanks, because it's not so easy to get interview partner for like because it's 20 minutes is also time and yeah, but very nice of you.

60 [0:26:59] **B11:** My pleasure. Yeah, I wish you good luck with that.

61 [0:27:02] **I:** Yeah. Thank you. You too. Goodbye.

62 [0:27:06] **B11:** Thank you. Bye bye.

1 Interview B12:

2

| | |
|---|---|
| Interview-Nr. | 12 |
| Date of the interview | September 10, 2024 |
| Duration of the interview | 24:08 min |
| Interviewer | Felix Starnecker (I) |
| Interviewee | B12 (Spain) |
| Role | Chief strategy officer |
| Sector | Data and cloud |
| Specialities | No specialities |

3 [0:01:09] **I:** Then let's start with the questions. Umm, like from your experience with whom the company share data and for what purpose.

4 [0:01:25] **B12:** Yeah. Well, I think in general company share data with whomever they have to as part of their to say, well, they're corporate process is they're value chain, right. So if you are a manufacturing company, you definitely have to share the data with your suppliers. Quite possibly with someone helping you to understand your clients or market your clients. And really, the purpose that I would say at the moment is mostly about operational efficiency. So to really get the processes going and there are contracts and agreements that regulate this change of data. But I think each of them are quite differently. So the question and I guess that's part of the big question for the European Union and call people like you and you know what are some of the enablers, these article actually is what are some of the enablers for private sector trust in data sharing beyond you know sharing data that you have to if you want to build things or market things.

5 [0:02:27] **I:** Mm-hmm. And would you say trust is an important factor when company share data?

6 [0:02:43] **B12:** 100% you mentioned here control or trust. I don't. I'm not sure what you mean by control.

7 [0:02:50] **I:** Umm, actually it's more about the trust factor. Control is just like control can be like this technology. I'm researching secure multiparty computation. With this technology you can have more control over data, but also like contracts as a kind of control mechanism.

8 [0:03:08] **B12:** Yeah, I guess that's the reason why you need it. To be able to exchange data is we need to have some level of trust. You know which typically will be done by lawyer agreements or

this kind of stuff. So, you know, if we're able and that's kayaks is sort of mean value proposition is to try to have a more automated flow or more generic framework for what flow trusts or for what trust or control means you know and I'm sure you're familiar with things like that. The idea, say data space connector and all these things which are also technical means.

9 [0:03:47] **I:** Yeah. OK. And like maybe in combination with trust like what you say that it's a subjective or objective process when companies decide to share data.

10 [0:04:02] **B12:** Well, it's subjective in the sense of a cultural. I think companies have a culture and some companies will be more willing to not just sharing data, but in general trying new things and some of it wouldn't be as much. But I think there's definitely an objective process as well in the sense that if you are able to and this is part of what actually goes on in, in sort of being in the world at the moment, it's trying to also come up with an understand sort of an economical or financial framework or model for understanding how benefits from big sharing. And I'm gonna share with you something. I'm looking it up so this is happening this week. Yeah, I would share it on the chat, too. So this is a conference one day conference at the police dolphin, which is being one of the biggest universities in Paris when it comes to economics. And if you scroll down you will see that there're four elements and the first is the economics of sharing data and the 2nd is the business model of orchestrators data sharing orchestrators. So these are sort of theoretical frameworks or theoretical framework really to try to understand. Hey, can I make money? Can I benefit from sharing data so that will be more objective, but then culturally some companies will be more interested than others, so there is definitely a cultural factor which is completely subjective.

11 [0:05:36] **I:** Yeah, but also like, would you say like if top managers know each other or something that they're more willing to share the data than if they don't know.

12 [0:05:47] **B12:** Yeah, totally.

13 [0:05:48] **I:** Yeah. And what risk do you see in data sharing?

14 [0:05:52] **B12:** Well, there's many, right? I think probably one of the largest ones and I'm sure people have told you this is reputational risk or some regulatory, you know, whether somebody has a breach and it comes back to you. But I think there are also competitive risks. People who might say, well, my data, I think in general there tends to be an idea that your data is worth more than there is I think data can be. I'm actually writing a chapter in a book and I found a very interesting article by the seal of the Spanish gigantic bank called Banco Santander. And she actually, I'm sure you've heard of data being described as the new oil. Let me find it for you. Umm. Some people actually describe it. I think it's Tim O'Reilly who is the founder of O'Reilly. Uh, editorial. So it's this editorial company which is heavy on tech. So he describes it as the data is the new sand because basically says well, oil

has some value. I mean obviously for petrol you have to transform it, but sand is useless. But when you transform sand, it becomes silicon and then you can actually make a lot of things. But what he was trying to say is that you need to transform it that data by itself is useless and an opportunity of team who's the CEO of this Spanish band said it's the new oxygen. Because what she was trying to say with that is that's also poisonous that if you have bad data, you can actually make bad decisions.

15 [0:07:33] **I:** Yeah, I have to.

16 [0:07:33] **B12:** I send you so there's a lot of risks. And do I think they're overestimate or underestimate it? And I think UM, they are non estimated. In other words, I don't think they're underestimated. I think there's not a clear framework for how to measure risks, and that means that people either are being underestimated them or overestimating them. Ohh man, I'm not sure.

17 [0:08:02] **I:** So they just can't estimate it, right?

18 [0:08:05] **B12:** It's not easy. I mean, you could do it, but I just think in general companies are not, you know, it's more academic work that could be done about estimating the risk of sharing data and you're asking what are the prerequisites for company successful data changed and companies well. The first thing I would say is you need to have a data problem to solve. That's the key and a motivation. Why are you sharing data? Sorry, I'm sending you a lot of links, but based because I think there is a lot of discussion to be held about this and obviously, you know, we cannot spend 8 hours, but I do have thought a lot about this questions.
Let me find something that I also found.

19 [0:08:50] **I:** Nice. Yeah. Thank you very much for the links. I will look into everything.

20 [0:08:59] **B12:** Let me find them. This is something I rolled. Maybe two years and a half ago.

21 [0:09:08] **I:** But if like there is a motivation to share data or reason, then there's also maybe things like interoperability, isn't it?

22 [0:09:20] **B12:** So I think to me, interoperability is the way to do it. It's the tool, but it's not the reason to do it. I'm trying to find you something I wrote. It's actually, it's kind of stupid. I mean it doesn't require here. It's in Spanish, but I mean I need translator which will do the job for you.

23 [0:09:37] **I:** Yeah, I can also speak up in Spanish.

24 [0:09:37] **B12:** So we wrote this. There you go. So it's a development model for use cases in data spaces and if you scroll down to the last page, it's a very simple road really it's a path and that has like eight different steps. And the first one is finding your business problem and this is something I

wrote when I was leading the data office for the Spanish Deputy Prime Minister. And so really this is very simple, but basically what it tries to say is you shouldn't start by understanding why do you need to share data. Because if you don't have the incentives, then data sharing will not happen. So to me, that's the first prerequisite. And then of course, you have to think about how we're going to model the data, what that entails and step #3 in that. I don't know if you have it in front of you. It's a requisite specification that is about governing how to participate in data sharing. How do you build the trust? What does trust mean for each organization? That's actually the most complex one.

25 [0:10:42] **I:** But what you said with the incentives to share data that they are so important, what is if companies, maybe they don't know that there is some reason to share data because they just have too less knowledge about it or is that even possible?

26 [0:10:57] **B12:** You're right. I have three, no four types of incentives, which are very, very simple and very high level. The first one is very simple to build a data marketplace, so everybody understands this because there's data marketplaces already in place and also we understand what marketplace is right. So that's almost like a breach in, let me actually. Sorry, I'm going to follow you with information, but I think that's useful when let me get it.

27 [0:11:26] **I:** Oh, that's good.

28 [0:11:28] **B12:** This is my job, right? This is what I do for my guy.

29 [0:11:31] **I:** Right now?

30 [0:11:32] **B12:** Yeah. So I think about this all the time. So that's why when I saw your questions, I was like of course I can answer these. This is exactly what I did.
So can we download this?

31 [0:11:39] **I:** Nice. Yeah, of course.

32 [0:11:42] **B12:** This is a bit longer than you need, but I'm gonna point you to the right place. Things just like 678 pages you owe me. Like half a page in this. Ohh sorry this is the older version. Now I wanna send you the key version. Good. Just a second. I'm going onto my so yeah. And if this is in English so OK, so I'm gonna put it on the screen. I'm gonna take the text instead of writing that. I'm just gonna take that photo, but an image and then I will send you the actual document, but here you say I've given four different reasons, but really they're more like glasses of reasons.

33 [0:12:26] **I:** Perfect. Nice.

34 [0:12:30] **B12:** Yeah, like sharing marketplace. So it's basically, hey, I have detailed assets, I have

data, I have algorithms. Umm, you know why do it again if somebody else has it and he can sell it, you know?

35 [0:12:44] **I:** Yeah, I understand.

36 [0:12:45] **B12:** And there's a big corporation in Spain, a big association called in the CIA, which is an industry association for AI. And the whole point of it was exactly that one company spend millions and millions of euros to develop a data platform and then they thought, well, you know, somebody else has to do this needs these, too. So why don't we make money out of it? And I'm sure you've heard of AWS.
Amazon Web Services.

37 [0:13:09] **I:** Yeah, of course.

38 [0:13:10] **B12:** So AWS was not originally created to sell. Cloud was for themselves, and then they realized they could actually make a lot of money. As a matter of fact, they make more money for many AWS than from Amazon. The second reason is cost sharing. It's basically saying, hey, I need to do something that is extremely complex. For example, the German automotive industry needs to compute the CO2 for the government and it's extremely expensive to compute CO2 by yourself. And so they do share the cost. The third one probably is the most open one joined innovation you can really do a lot of things when you put data together because data by itself is useless. But when you join it and then the 4th one, it's more about making sure that no one is like there's no Amazon in the middle that takes like or booking, which takes like so much percentage of, you know, Netflix or whatever. Well, maybe not so much Netflix, Airbnb or Uber, you know, so those are to me the main reasons. I mean to download this and I will send it to you in a little bit so.

39 [0:14:11] **I:**I see. And what for external or internal influences? What would you see there?

40 [0:14:37] **B12:** Yeah, and governmental. I think governmental is a big, big deal. I think governmental influences are huge, and so if you are able to, that's a very significant external influence.

41 [0:14:46] **I:** Uh, like which regulations they make?

42 [0:14:56] **B12:** Yeah. And also because you know the data act, for example forces companies to share data. So then there's no question you have to that's it. Umm, internal. That's an interesting one. I don't think any company. Well, actually I don't know if you've ever heard of this concept of data mesh, which is this organizational structure for sharing data in a company. It's actually very, very popular, but it's very difficult to actually execute in practice. Umm, data mesh is really to me the

main reason for an internal company. To you, you're not doing external sharing, you're doing internal sharing, but still sharing. Umm, initially just really like that for you.
If you have a lot of things.

43 [0:15:49] **I:** Yeah, thank you I understand. OK, then let's go on with the questions, because otherwise we can't come through. And like for privacy and enhancing technology, shall I? Do you know I'm secure multiparty computation?

44 [0:16:06] **B12:** I've heard of it. I've seen it in working, but I am not an expert like I wouldn't be able to program it so.

45 [0:16:09] **I:** Yeah, okay, but it's enough when you know how it works. And my first question would be like do we see use cases for secure multiparty computation or do you think it's only a nice technology but not practical relevant?

46 [0:16:32] **B12:** I don't know. I mean, I'm sure there are, but you know, because we had [company], we work more on the general framework. We don't work on a specific use cases and they are done by the companies and data. Umm, I think I'm sure there are, but I'm not aware of any that I know you know.

47 [0:16:50] **I:** OK. And would you say that, umm, that this is a technology?
Who will change the data sharing behavior of companies? Do you like in the future?

48 [0:17:00] **B12:** No. Umm, no. And I would say I'll tell you why, because I think it's a technology solution, but umm, it's almost like saying hey. We have a lot of technical tools. How come there's still cybercrime? How come there's still crypto ransom? Well, because usually problems come from the weakest link. So technically speaking, we have infrastructure that is completely almost perfect today. However, there's still a lot of data breaches, and that comes from, you know, fishing, usually from fishing from people in the company clicking on a link that they shouldn't click. So what I'm trying to say with that is that secure multiparty computation is probably a fantastic way, but does that eliminate every single human risk? The answer is no.
So I still think you probably need lawyers and contracts and legal security. So would I think SMPC will be fantastic? Yes. Would I think we'll eliminate all the risks? No. And also it's gonna get more and more technical right? So and that moment like only like really techies and geeks can actually understand it. So, umm, you know, it's almost like blockchain, right? So blockchain works. I have technically fantastic, but where is really blocking being used in practice for like Bitcoin? But I mean like really making a significant change, no one.

49 [0:18:40] **I:** Because two less people understand that you would say

50 [0:18:46] **B12:** Your next question is the conditions for data change with pets.

51 [0:18:52] **I:** The other antitrust, would you say the trust factor is also still needed?

52 [0:18:52] **B12:** Well. Yeah. And trust is defined and different levels, right? Not just technological. Technological is one, but I think culturally is another one.

53 [0:19:04] **I:** Uh, yeah. So trust is always needed. Doesn't matter which technology is used, it's always needed.

54 [0:19:12] **B12:** Yeah. And I like the same for example like think about self driving cars, right? And they might actually be safer than human cars because, well, they drive a lot more. And they don't. But still people are like really scared. And there's like where you're not scared to drive with that. That's your driver. Like, who might be tired? You know, so that my point is, I think humans get used to a certain way of behavior and then changing that can be difficult.

55 [0:19:34] **I:** Yeah. And to the conditions for companies who would rather invest in uh, privacy enhancing technologies, would you say there are differences or it's just use case?

56 [0:19:47] **B12:** Umm. I'm not sure, to be honest. I think, yeah, I think they really need to have a business use case and a cultural ethos or cultural fabric that they are ready to try it? I think if you're ready to data share, umm, privacy and technology will help you a lot. But if you're not ready to share data culturally or business driven, then privacy enhancing technologies will not do anything like if you're not ready to drive in a car, it doesn't matter how good the car is, yeah.

57 [0:20:20] **I:** Yeah. OK. And to the next questions, how privacy enhancing technologies change the risks? You already said that there's always still risk.

58 [0:20:33] **I:** Would you say the risks are getting differently than they are now?

59 [0:20:38] **B12:** Ohm, they're more complex to understand, right? Because they're technology driven, so, and I mean I think we are minimizing the probablity for risks, but also like, for example, with cybersecurity, what happens is now there's everybody. Everyone is connected, so the footprint of potential risks has really gone up a lot like 25 years ago, the technology was worse, but the amount of damage you could do was minimal because the world was analog. Now the world is so digital, even though technology is a lot better, the footprint or the surface in which you can do damage is way, way bigger. So you know you're going to, I don't know, 23 and me, and suddenly you find that you're exposing genetic DNA of people you know. So maybe the probability of risks can made smaller, but companies could start share more and more sensitive data and so so the damage is bigger when something happens.

60 [0:21:28] **I:** But to the cybersecurity like one guy I interviewed he said that this would be a good use case for prior secure multiparty computation, because then like companies can share who was attacking them. And then they yeah, I don't know, better prove that for it.

61 [0:21:48] **B12:** Yeah. I mean [company] doesn't work necessarily on the tech solution. It's more of a tech solution for trust, but not so much for the actual data sharing.

62 [0:22:00] **I:** Yeah. And what do you see as the biggest challenges for the acceptance of pets on the market?

63 [0:22:08] **B12:** Well, I really think it's all about the willingness to share data.
So like if you're gonna share data, private enhanced technology should be like no brainer. You have to go there. And why wouldn't you do it, right? This is like saying, oh, I wanna drive through Germany, but I'm not willing to go on a plane. Why? Like, do you wanna walk? But if you don't wanna go to Germany, if you just wanna go to like next door, I mean, if you're in Spain, if you're in Germany, but you're already in Germany. But then the whole point is, I think the biggest challenge for the substance of pets is technology complexity. So, like how much talent you have to how well bundled the services from technology providers are. So like they make it simple? Or do they require a lot of work for you to put on top. Then you need many people who know about this stuff. And then really the most important one is that desire to actually benefit from sharing data? So again, going back to part one of your general questions, like is there a framework where we understand how to share data, how to benefit from sharing data, what the risks are? I think that will really, and it's gonna depend. You can apply it differently to different companies and different scenarios, but it can really help you. Umm, understand what your state?

64 [0:23:23] **I:** Touchable.

65 [0:23:24] **B12:** Yeah. And decide to go for it or not go for it. Or as I think you will go for it in specific circumstances like just you don't take a plane to go for 100 kilometers, you could, but what's the point, you know?

66 [0:23:27] )**I:** So if we have a framework, then maybe companies can estimate the risks and everything better and also the benefits and then?

67 [0:23:43] **B12:** And that's part of the guy I asked. That's what the Guy acts Institute and I don't know if you've ever heard of the lighthouse projects from Gaya with catena eggs with mobility that space.

68 [0:23:46] **I:** I know, yeah.

69 [0:23:53] **B12:** These are projects that really want to share data, and the idea is that they can really

teach the rest of future projects how to do it.

70 [0:24:05] **I:** Yeah, I understand. So thank you very much for the interviews.

71 [0:24:08] **B12:** No worries.

1 Interview B13:

2

| | |
|---|---|
| Interview-Nr. | 13 |
| Date of the interview | August 19, 2024 |
| Duration of the interview | 28:29 min |
| Interviewer | Felix Starnecker (I) |
| Interviewee | B13 (Germany) |
| Role | Founder |
| Sector | IT Sector (Privacy Enhancing Technologies) |
| Specialities | Interview in german |

3 [0:01:55] **I:** Aus ihrer jetzigen Erfahrung: Kennen Sie Unternehmen, die Daten mit anderen Unternehmen teilen? Und wenn ja, welche Beziehung haben die Unternehmen zueinander?

4 [0:02:15] **B13:** 2:15
Das ist sehr unterschiedlich. Ich kenne Unternehmen, die Daten miteinander teilen, das passiert insbesondere da, wo es mir bekannt ist, unternehmensverbänden also innerhalb einerselben Branche, wo Unternehmen ein Interesse daran haben sich untereinander besser zu verstehen, wie sich der Markt gerade bewegt und wie sie sich auch selber, wie sie selber auf dem Markt dastehen. Da passiert das in aller Regel über Intermediäre, in diesem Fall beispielsweise Juristen. Oder Notare, die dann zur Geheimhaltung verpflichtet sind.

5 [0:002:51] **I:** Mhm, Mhm, OK und würden Sie sagen, dass der Entscheidungsprozess bei solchen datenbezogenen Entscheidungen eher subjektiv oder objektiv abläuft? Hat da immer eine Person nur entschieden, über Daten Entscheidungen oder geht das dann über ganz viele Instanzen?

6 [0:03:19] **B13:** Das sind in der Regel verschiedene Ebenen, je nachdem auf welche Daten das natürlich betrifft und insbesondere, wenn es sensible Daten sind, sind in der Regel immer die Geschäftsleitungsebene beteiligt häufig auch eine juristische Ebene und natürlich ist auch meine wirtschaftliche Entscheidung. Der Bereich, in dem wir uns bewegen, da ist es vor allem auch einfach die Frage wie groß ist der wirkliche Nutzen davon, die Daten gemeinsam zu verarbeiten? Das ist in der Regel die große Herausforderung, und das ist das, woran es am ehesten scheitert, ist mein Eindruck, dass man den Unternehmen, das eine kritische Masse von Unternehmen sich beteiligen würden, so dass nachher am Ende für alle ein ausreichend großer Nutzen, da es sich daran zu

beteiligen, da ist es dann häufig einfach eine Business Entscheidung im Sinne von im monetären Sinne lohnt es sich für uns, Daten zu teilen oder auch eine Datenaustauschplattform mit anderen zu initiieren?

7 [0:05:46] **I:** Okay, und über welche Kanäle werden Daten normalerweise geteilt?

8 [0:06:00] **B13:** Das, was ich am häufigsten gesehen habe, wenn es um sensible Daten ging. Wie gesagt über Juristen oder über allgemein über Intermediäre wenn es wirklich sehr sensible Daten sind, dann in der Regel ein Notar, der zur Geheimhaltung verpflichtet ist, der dann alle Daten einsammelt und da und die dann auswertet.
Wenn ich mir jetzt etwas weniger sensible Daten ansehe wir waren beispielsweise ein Gespräch mit einem Unternehmensverband, wo es dann darum ging, Krankenstände über verschiedene Unternehmen zu vergleichen. Da war es einfach der Verband als Intermediär, der diese Daten eingesammelt hat. Das ist natürlich, wenn es für die Verbandsmitglieder in Ordnung ist, völlig ausreichend, aber ist natürlich im rechtlichen Sinne weniger sicher als beispielsweise ein Notar. Aber kurz gefasst in der Regel war der Austauschweg über juristische Entitäten. Oder ja, doch juristische Entitäten und darüber abgesichert, nicht auf eine technische Art und Weise.

9 [0:07:06] **I:** Weniger über Plattformen, der Austausch oder?

10 [0:07:09] **B13:** Ja weniger, aber Plattformen hier haben wir auch Gespräche mit verschiedenen Organisationen geführt. Allerdings sind bei diesen Plattformen.
Es wird sehr viel entwickelt, ich hab aber noch nicht miterlebt, dass darüber wirklich ein aktiver, produktiver Datenaustausch stattfindet.

11 [0:08:31] **I:** Welche Faktoren würdest du sagen beeinflussen den Datenaustausch zwischen Unternehmen?

12 [0:08:39] **B13:** Also äußerer Druck auf jeden Fall da, da auf jeden Fall der ich meine.eu Data Act, der da eine Rolle spielt und der dafür sorgt, dass Unternehmen wirklich Druck haben, auch Daten irgendwie zugänglich zu machen in der einen oder anderen Form. Darüber hinaus Wertschöpfung. Unternehmen teilen Daten nicht aus Jux und Tollerei. Damit Unternehmen bereit sind, Daten zu teilen, muss für sie ganz, ganz, ganz, ganz deutlich sein was ist der Nutzen davon, wenn der nicht klar ist, ist es sehr, sehr schwer, unternehmen davon zu überzeugen. Vertrauen auf jeden Fall auch. Hier auch einfach daher kommend Unternehmen unter oder zumindest die Entscheidungsträger verstehen in der Regel nicht auf einer technologischen Art Ebene. Wie sicher ist etwas wirklich? Sie

sind daran darauf angewiesen, den Aussagen Dritter im eigenen Unternehmen oder auch Externer zu glauben. Und hierbei spielt einfach als psychologischer Faktor vertrauen eine erhebliche Rolle. Weil in der Regel sind die sind die Entscheidungsträger selber nicht in der Lage, die technische Sicherheit einer Lösung dafür zu beurteilen.

13 [0:10:33] **I:** OK, würden Sie sagen, dass da bestimmte Punkte gibt, die das Vertrauen sehr stark oder sich einfach auf das Vertrauen auswirken, wie ich weiß nicht die Region aus den Unternehmen kommt, die die Mitarbeiteranzahl, die Webseite, also das Design der Webseite.

14 [0:10:54] **B13:** Im weitesten Sinne schätze ich einfach ein Track Record eine Website ist super, um einfach auch Leuten etwas in die Hand geben zu können, sodass Sie sich informieren können, dass Sie auf einen selber stoßen können. Aber in meiner Erfahrung nach und meiner Einschätzung nach ist der wichtigste Faktor, wenn es um Vertrauen geht, zu wissen: Andere haben diese Organisation auch schon vertraut. Also da einfach nen Track record. Um sagen zu können andere haben denen schon vertraut und haben damit gute Erfahrungen gemacht, deswegen kann ich denen jetzt auch trauen. Darüber hinaus natürlich auch ne gewisse technische Expertise, die belegbar ist über Publikationen, Titel beispielsweise auch das hilft. Ist aber im Endeffekt eine Unterstützung, um Vertrauen zu erlangen, aber kein Ersatz und an der Stelle auch einfach dann insgesamt einen sag mal professionelles Auftreten.

15 [0:12:03] **I:** OK, alles klar, denn dann würden wir jetzt zum zweiten Teil kommen. Das ist ja ihr Fach Fachgebiet, das muss Ihnen nicht mehr erklären. Für welche Anwendungsfälle würden Sie sagen, ist Secure Multiparty Computation speziell geeignet beziehungsweise welchen Nutzen würde es bringen für Unternehmen?

16 [0:12:41] **B13:** Secure Multiparty Computation ist insbesondere dann interessant, wenn es darum geht, Daten Unternehmensübergreifend zu nutzen, nicht nur in einem Unternehmen selber beziehungsweise besser gesagt Organisationsübergreifend. Ein Unternehmen kann ja auch aus mehreren Organisationen im Sinne von mehreren Abteilungen bestehen und wenn es beispielsweise aus rechtlichen Gründen innerhalb dieses Unternehmens grenzen zwischen verschiedenen Abteilungen gibt, zwischen denen Daten nicht ausgetauscht werden dürfen, dann kann es ja auch hier multiparty Computation eine Lösung sein, diese Grenzen nicht aufzulösen, aber eine sinnvolle Nutzung der Daten zu ermöglichen, ohne dabei gegen diese Grenzen zu verstoßen darüber hinaus insbesondere auch dann, wenn es darum geht, über verschiedene Unternehmen hinweg gemeinsam Daten zu nutzen, insbesondere wenn es um mehr als jetzt 2 Unternehmen geht, wenn 2 Unternehmen geht, wenn es um einfach nur 2 oder 3 unter, wenn es um 2 Unternehmen geht, ist

beispielsweise homomorphic encryption häufig auch eine sehr gute Lösung, insbesondere da hier mittlerweile auch spezialisierte Hardware entwickelt wird. Aber Secure Multiparty Computation insbesondere dann, wenn es 3 oder mehr Parteien gibt, die gemeinsam Daten nutzen möchten.

17 [0:14:07] **I:** OK und was haben Sie festgestellt sind Voraussetzungen, die es braucht, damit Unternehmen Privacy Enhancing Technologien oder speziell Secure Multiple Computation investieren würden?

18 [0:15:06] **B13:** Dass die wichtigste Voraussetzung ist, dass eine Gruppe, die dazu bereit ist oder die Interesse daran hat, bereits besteht. Gewissermaßen das empty Room Problem wenn niemand, wenn man ein altes Unternehmen fragt, willst du das nicht machen, dann wird die Antwort sehr wahrscheinlich nein lauten, weil dieses Unternehmen müsste erst mal selber loslaufen und weitere Akteure hinzuholen. Wenn es allerdings schon eine Gruppe von Akteuren gibt, die bereits ein erklärtes Interesse daran hat, Daten gemeinsam zu nutzen, beispielsweise über einen Branchenverband. Dann geht es wesentlich einfacher. Aber da es wirklich die Voraussetzung, dieses Problem zu lösen. Und es ist natürlich sehr, sehr hilfreich, wenn bereits Daten in einer hohen Qualität vorliegen und in einer bereits digitalisierten Form. Gerade für den Anfang im Einsatz von MPC ist aber auch beispielsweise einfach eine manuelle Eingabe über ein Webinterface möglich, sodass dann die Verarbeitung der Daten noch im Webbrowser passiert. Und dass keinerlei Informationen den Webbrowser verlassen, die irgendetwas mehr verraten würden als das, was sie verraten sollen. Zum Zwecke dieser Berechnung in dem Sinne, dass die Daten jetzt schon irgendwo in einer Datenbank aufbereitet vorliegen, ist eine große Hilfe ohne Frage, aber kein Must have.

19 [0:17:03] **I:** Und das ist jetzt auch nix, was sich nur große Unternehmen leisten könnten, weil es eben viele finanzielle Ressourcen braucht, dass es für jedermann eigentlich zugänglich, also egal ob klein ist oder großes Unternehmen.

20 [0:17:17] **B13:** Das ist das, was wir vorhaben, das ist für jedermann zugänglich ist. Stand heute ist es aber in der Tat immer noch so, dass wenn Unternehmen diese Technik nutzen möchten, werden sie selber in aller Regel hohe technische Expertise benötigen auch von Kryptographen wie mir. Um die Technik selber auch sicher einzusetzen es gibt nun mal dabei viele subtile Fallstricke, die man beachten muss, um die Berechnung nachher wirklich sicher durchzuführen und ohne eine hohe technische Expertise im Unternehmen wird das in der Regel nicht gegeben sein. Damit das Sache wirklich auch funktioniert. Gleichzeitig ist dann natürlich auch so diese hohe technische Expertise, die wird sich nen, die wird sich für ein kleines Unternehmen in aller Regel nicht lohnen, in House zur Verfügung zu haben, die wird sich eher dann lohnen, extern zur Verfügung zu haben, wenn sie

damit es einfach auch skaliert oder halt wenn man selber als Unternehmen so groß ist, dass man genügend Anwendungsfälle hat, dass es sich lohnt, die Inspirative selber Haus an zu haben.

21 [0:18:28] **I:** Heißt das, die technische Expertise wird durchgehend gebraucht? Nicht nur bei der Implementierung zu Beginn, sondern eigentlich die ganze Zeit über, dass immer jemand dabei sein muss, der sich damit sehr gut auskennt.

22 [0:18:44] **B13:** Das nicht unbedingt das kommt drauf auf den Anwendungsfall ab, an, wenn wir ihren Anwendungsfall nehmen an verschiedenen Unternehmensinhalt eines Unternehmensverbandes zu vergleichen und das System auch mal aufgesetzt ist, dann reicht es in der Regel, für das Aufsetzen jemanden dabei zu haben, sollen allerdings wiederholt verschiedene Anwendungsfälle immer wieder evaluiert werden. Dann ist es in der Regel von Ministerien in der Regel notwendig für verschiedene Konfigurationen, wenn Sie denn erstellt werden sollen auch Experten zu haben.

23 [0:19:17] **I:** Mhm ja und sehen Sie neue Risiken? Also normalerweise ist ja SMPC dafür da, um Risiken beim Datasharing zu senken, sehen sie aber auch Risiken, die eventuell hinzukommen könnten für Unternehmen, wenn sie MPC implementieren.
Oder nutzen?

24 [0:19:39] **B13:** Risiken, die hinzukommen. Mhm. Das größte Risiko, dass ich aktuell sehe, ist an der sind an der Stelle. Fehler im Umgang mit der Technik? Und Fehler in der Implementierung wenn Unternehmen SMPC nutzen, um sensible Daten auch Unternehmensübergreifend zu verarbeiten. Er fordert das natürlich, dass sensible Daten verarbeitet werden. Wenn es technische Fehler in der Software gibt, die diese Daten verarbeitet, dann können dadurch sensible Daten öffentlich werden. Die wenn die Daten erst gar nicht verarbeitet worden werden, natürlich auch nicht öffentlich geworden wäre. Genau das heißt, wenn ich die Daten gar nicht verarbeite, sind sie natürlich ideal geschützt. Wenn ich Sie verarbeite und in der Technik, mit der ich Sie verarbeite, ein Fehler enthalten ist dann könnte natürlich öffentlich werden. Die andere Sache ist der sichere Umgang, damit der der richtige Umgang mit der Technik, was Secure Multiparty Computation liefert ist an der Stelle eine sogenannte ich meine, es ist ja doch eine sogenannte Input Privacy das heißt, es wird nur so viel über die Ergebnisse verraten wie nachher das die Ausgabe auch über das Ergebnis verrät, je nachdem welche Berechnung ich aber durchführe, kann das mitunter immer noch sehr, sehr viel sein. Ein banales Beispiel ist dafür, einen Mittelwert zwischen 2 Werten zu bilden. Wenn ich einen der beiden Werte kenne, dann kenne ich auch den anderen, wenn ich den Mittelwert erfahre, also nehmen wir als Beispiel 2 Menschen wollen ihr Durchschnittsgehalt sicher ausrechnen, dann können Sie das sehr, sehr sicher ausrechnen. Miteinander jeder hat am Ende nur sein eigenes Gehalt und den

Mittelwert. Es ist aber an der Stelle dann banal aus dem Mittelwert und dem eigenen Gehalt das Gehalt der anderen Person auszurechnen. Sie sehen also, wenn ich eine falsche Berechnung mit Computation ausführe, dann sind zwar alle technisch und auch mathematischen Sicherheitsgarantien nach wie vor vorhanden, allerdings kann der falsche Verständnis der Technologie an der Stelle dann oder die Missachtung dieser subtilen Gelegenheiten dazu führen, dass am Ende das Ergebnis mehr verrät, als man gewollt hat.

25 [0:23:22] **I:** Und bei welcher Art von Unternehmen würden sie einschätzen ist die Bereitschaft, PETs speziell MPC zu benutzen oder darin zu investieren, am höchsten? Beispielsweise von der Größe vom Unternehmen, vom Sektor wo ist ihnen aufgefallen, dass am ehesten die Technologie zum Einsatz kommen könnte?

26 [0:24:02] **B13:** Meine Einschätzung an der Stelle ist, dass es gar nicht mal um den Sektor unbedingt geht. Die Frage ist eher wie agil und innovativ sind unternehmen also? Desto eingefahrener eine Industrie ist, desto konservativer auch indem sie in ihrer Arbeitsweise, desto größer ist die Hürde. Andererseits wiederum wer selber an der Front der Innovation in seinem Bereich ist und ständig auch weiter Innovation leistet, wird wahrscheinlich eher bereit sein, auch MPC einzusetzen.

27 [0:24:44] **I:** OK also eher die Einstellung vom Unternehmen und nicht andere Faktoren.

28 [0:24:47] **B13:** Ja, ich halte ich halte eher die Kultur und Einstellungsunternehmen für relevant als eine spezielle Branche, Region oder ähnliches.

29 [0:24:56] **I:** OK. Und wo denn sehen Sie die größten Herausforderungen für die Akzeptanz von PETs auf dem Markt.

30 [0:25:12] **B13:** Das Empty Room Problem für mich. Ganz klar man muss wie gesagt, ich halte SMPC dann für interessant, wenn viele Parteien teilnehmen. Die Herausforderung ist, dass diese vielen Parteien zusammenzubringen. Hat man die nicht, ist der Nutzen wesentlich geringer und dementsprechend auch natürlich der Anreiz, das zu tun, hat man das Problem gelöst, wird es für viele weitere Parteien von Interesse sein, Dazuzustoßen alleine auch schon, weil die Kosten für das System oder das System zu betreiben natürlich auch, desto besser skalieren, desto mehr Teilnehmer dabei sind.

31 [0:25:56] **I:** Ok und ändert sich der Einfluss des Vertrauens durch SMPC?

32 [0:26:05] **B13:** Ja, man muss weniger in das Gegenüber Vertrauen, mit dem man die Daten teilt und wesentlich mehr in die Technologie und die Mathematik dahinter.
Das auf jeden Fall also es ist gewissermaßen so. Wir versuchen, Vertrauen in andere Organisationen darin, dass die sich richtig verhalten, das Richtige mit den Daten machen, in Vertrauen in Algorithmik und Mathematik zu ersetzten. Das heißt, man muss danach weniger dem Konkurrenten vertrauen, weil sonst wird man mit dem wahrscheinlich gar nicht zusammenarbeiten. Dafür müsste man aber desto umso mehr in die Technologie vertrauen.

33 [0:27:03] **I:** Dann meine letzte Frage wäre aus dem gestrigen Interview wurden als Gründe gegen Datasharing zwischen konkurrierenden Unternehmen mittels SMPC genannt, dass das Risiko der Datenmanipulation zu hoch ist, also das heißt, wie man auch andere Unternehmen dabei kontrolliert, dass sie nicht irgendwie manipulierte Daten in die Technologie eingeben oder auch schlechtere Qualität von Daten.Was würden Sie dazu sagen?

34 [0:28:29] **B13:** Es kommt auf verschiedene Dinge an, wenn wir stark standardisierte Daten haben, beispielsweise Daten aus Geschäftsabschlüssen, die natürlich einfach aus rechtlichen Gründen schon stark standardisiert sein müssen. Dann sehe ich das Risiko als sehr gering an, dass falsche Daten eingegeben werden. An anderen Stellen kann das Risiko natürlich wesentlich größer sein. Hier gibt allerdings zu bedenken wenn die Berechnung durchgeführt wird, um allen einen Mehrwert zu ermöglichen, dann beraubt sich die Partei, die verfälschte Daten beiträgt, selber dieses Mehrwerts geschädigt, damit zwar auch die anderen Parteien, die auch teilnehmen, ist allerdings auch zu bedenken schädigt. Die haben sich damit auch selber, sie hatte wahrscheinlich genauso gut gar nicht teilnehmen können, was sie von Anfang an frei stand und hätte die Daten dann hätte, wenn sie falschdaten beiträgt und einem nachher ein sinnloses Analyseergebnis erhält, dann hätte sie es auch ganz sein lassen können. Ist dann die Frage, ob die Kosten der Teilnahme nicht den Wert des Schadens bei Konkurrenten übersteigen.

1 Interview B14:

2

| | |
|---|---|
| Interview-Nr. | 14 |
| Date of the interview | August 27, 2024 |
| Duration of the interview | 26:09 min |
| Interviewer | Felix Starnecker (I) |
| Interviewee | B14 (Germany) |
| Role | Chief Product Officer |
| Sector | IT sector |
| Specialities | Interview in german |

3 [0:01:42] **I:** Ja, könnten sie sich ganz kurz vorstellen, ja, was machen welche Expertise sie in dem Bereich haben?

4 [0:01:51] **B14:** Also ich bin jetzt selber bei meinem jetzigen Arbeitgeber seit vielen Jahren für das Produkt selber verantwortlich als Chief Product Officer. Hab hier ein größeres Team von UA, Spezialist, Spezialisten und Produkte, die sich um die Weiterentwicklung unseres lösungssportfolios Portfolios für sichere Datenkollaborationen kümmern. Davor habe ich lange Zeit auch im Bereich Cyber Security und Information Security gearbeitet. Hardwaresicherheitsmodule Gateway basierte Verschlüsselungssysteme in Bezug auf E Mail Endpoint Security Firewalls VPN in unterschiedlichen Rollen.

5 [0:02:36] **I:** OK, Mhm und aus Ihrer Erfahrung: Welche Daten teilen Unternehmen am ehesten und mit welchen Partnern und wofür?

6 [0:02:54] **B14:** [Unternehmen] ist eine Plattform, die als Sarstedt zur Verfügung gestellt wird, und wir haben jetzt hier B to B im Fokus, aber generell findet man ähnlich gelagerte Lösung wie eine Dropbox auch im B Two C Umfeld und der Use Case ist halt wie bei uns. Klassische Situation letztendlich Unternehmen A hat letztendlich sehr viele Verträge, die die dann letztendlich Dienstleister zum Beispiel für das Unternehmen sind und oder meinetwegen auch her Hersteller, die Unterauftragnehmer in diesem Umfeld sind und die tauschen letztendlich mit dem Auftraggeber oder anderen Partnern projektbasierte Unterlagen aus und das funktioniert dann halt traditionell über [Unternehmen] und dann bieten wir dafür werden unterschiedlichste Schnittstellen genutzt. Dort können automatisierte Einlieferungen von Dateien oder Dokumenten stattfinden oder über gewisse

Fronten, die wir zur Verfügung stellen und es gibt keine spezifische Industrie. Aber die Industrien haben sich selber jetzt zum Beispiel, wenn ich den Automobilsektor dort nenne, sich selber eigene Zertifizierungsprogramme auferlegt, wo letztendlich nur zertifizierte Produkte dafür eingesetzt werden dürfen. Zum sicheren Datenaustausch in diesem Vertical, also zum Beispiel OEM, tauscht nur über letztendlich eine zertifizierte Lösung mit TR ONE oder TR two suppliern letztendlich Projektunterlagen aus. Es gibt aber auch andere Beispiele zum Beispiel im Professional Advisory Umfeld, also letztendlich Wirtschaftsprüfer, Unternehmensberatung, die dort auch letztendlich über diese Tools mit ihren Kunden agieren, genauso das gleiche gilt für generell Berufs geheimnisträger. Ob das jetzt Ärzte oder Anwälte sind, es gibt aber auch letztendlich bafin regulierte Unternehmen aus der Kreditwirtschaft oder aus dem Banken und Finanzbereich, die da mit sensitive Dokumente austauschen.

7 [0:07:44] **I:** Ach so OK. Und würden Sie sagen, dass auf ihrer Plattform jegliche Datenarten geteilt werden oder werden eher weniger sensible Daten geteilt oder höher sensible Daten?

8 [0:08:05] **B14:** Mhm auch so sehr sensible Daten besonders aus dem Bereich Juristische Beratung, Wirtschaftsprüfung, Wirtschaftsabschlüsse genauso wie auch Patientendaten zwischen Facharzt und meinetwegen Allgemeinmediziner. Also diese Anwendungsfälle gibt es auf jeden Fall und wir haben auch letztendlich eine Zertifizierung, wo man Daten nach Datenschutzklassen unterteilt und da hat unsere Lösung letztendlich den Zertifizierungsgrad für die höchste Datenschutzklasse.

9 [0:09:21] **I:** Und welche größten Gefahren oder Risiken würden sie beim Teilen von Daten speziell sehen?

10 [0:09:31] **B14:** Mhm also generell ist es ja auch so, man hat immer letztendlich das Bedürfnis, dass man den Zugriffsstark möglichst einschränken kann also dieses nicht und nordprinzip. Dass man auch Rollen letztendlich in diesen Datenräumen definieren kann, was jeder doch zu tun hat der eine darf meinetwegen Dateien lesen, andere darf Dateien lesen und löschen. Der andere darf auch Dateien letztendlich dort hochladen und man möchte letztendlich den Leuten zur Erfüllung ihres Jobs die richtigen Berechtigungen geben und aber auch gleichzeitig das ganze Revisionssicher gestalten. Das heißt, dass auch eine dritte Stelle oder jemand, der dort mit beauftragt ist, zum Beispiel eine Möglichkeit hat, über gewisse Audit Trails nachzuvollziehen, was mit den Daten zu jedem Zeitpunkt geschehen ist oder ob die Leute, die dort mitarbeiten dürfen, auch immer die richtigen Berechtigungen haben.
Und das alles ist letztendlich Bestandteil des Produkts und aus dem Bereich Data Governance oder generell eine einen sicheren Ort zu haben, wo ich Daten verarbeiten, speichern oder weitergeben

kann.

11 [0:11:39] **I:** Wenn jetzt, wenn Sie Kunden von Ihnen anschauen würden Sie sagen, dass der Entscheidungsprozess, ob die Ihre Daten teilen oder auch nicht teilen, ob das eher ein objektiver Prozess ist, wo viele Risikoanalysen et cetera erst durchgeführt werden oder ob das eher ein subjektiver Prozess ist, wo viel mit rein spielt, wie einfach Vertrauen in den Anbieter.

12 [0:12:01] **B14:** Also ja, die meisten sind die Prozesse relativ stringent, also letztendlich der Kunde geht erstmal von der Datenklassifizierung aus und dann werden gewisse Tools bei den Kunden dafür freigegeben. Das heißt Daten, die einen gewissen Klassifizierungsniveau unterliegen, dürfen nur über bestimmte Tools getauscht werden. Das kann letztendlich eine interne Verfügbarmachung sein, genauso wie intern nach extern und nach dieser Datenklassifizierung und der Freigabe von gewissen Tools arbeitet der Kunde und da gibt es auch das kann man nachvollziehen. Meistens sind die Seiten sogar extern erreichbar. Bei größeren DAX Unternehmen gibt es dort Landingpages, wo die Mitarbeiter genauso extern informiert werden, welche Tools dort zum Einsatz kommen dürfen zum Datenaustausch.

13 [0:13:51] **I:** Ja, ja OK. Dann zu privacy enhancing Technologien beziehungsweise spezielle Secure Multiparty Computation. Sie benutzen es nicht bei [Unternehmen], aber überlegen Sie das vielleicht einzuführen?

14 [0:14:13] **B14:** Nein, nein, nein. Nicht in der Überlegung. Aktuell ist es nicht in Überlegung, das Ganze zu tun, weil hier ist es ja so, dass Leute ganz gezielt eingeladen werden, Dateien zu tauschen. Wo letztendlich die so jetzt zum Beispiel dieses Millionärsproblem, was gerne bei SPMC genommen wird, nicht das Problem ist, weil letztendlich der Teilnehmer Circles transparent und man möchte halt projektdaten, die wichtig sind in dieser Teilnehmergruppe, die in virtuellen Datenraum arbeiten, untereinander austauschen und dann überlegt man sich, wer diese Dokumente oder Dateien überhaupt sehen darf. So und wenn man jetzt diese Anwendungsfälle von SPMC nimmt, die kommen aus gewissen Applikationen halt heraus, und wir stellen ja eine eigene Applikation selbst zur Verfügung. So und jetzt wär natürlich aus dem Innovationsprozess letztendlich eine gewisse Applikation dafür zu definieren, die ein gewisses Problem löst. Und diese Applikation stand heute gibt es bei uns nicht, wir würden aber letztendlich könnten wir uns mal vorstellen in Zukunft, weil wir auch eine Technologie anbieten. Die sogenannte Seal Cloud, wo wir Daten in News also wenn man die verarbeitet schützen kann. Vor privilegierten Angriff vor privilegierten Zugriff, also zum Beispiel von unserer Operator Seite, könnte man sehr sensitive Applikationen letztendlich schützen, so und wenn man jetzt so eine Applikation schreiben würde für Spmc, die man gerne halt

letztendlich in der Cloud nutzen würde, hätten wir so eine Basistechnologie dafür, um so eine Applikation zu hosten. Aber alles was halt hätte ich damit zugehört, so ein ganzes Key Management was man noch braucht das man es letztendlich von so einem Independent Software Vendor mitgebracht werden und.
Wir sind sehen jetzt aktuell auch aus Gesprächen oder Marktbeobachtungen jetzt hier bei uns keine SPMC Applikation, die wir in naher Zukunft anbieten würden

15 [0:17:01] **I:** Und sehen Sie Anwendungsfälle für Security multiparty Computation oder Potenzial?

16 [0:17:09] **B14:** Ja, also auf jeden Fall sehe ich das Potenzial also die die Frage ist oder die ich hatte ich im Vorfeld versucht zu beantworten, also wir sehen uns da jetzt nicht als potenziellen Anbieter in naher Zukunft aber Anwendungsfälle sind vollkommen legitim.

17 [0:17:29] **I:** Mhm OK und sehen Sie auch irgendwie? Dinge an MPC, die vielleicht nicht praktikabel sind oder wodurch vielleicht die Technologie nicht akzeptiert wird?

18 [0:17:41] **B14:** Also die Akzeptanz hat ja immer letztendlich wie bei vielen Fällen damit zu tun irgendwie, oder? Wie stark sind die Voraussetzungen, die vorhanden sein müssen, um so ein Angebot erfolgreich zu zu machen? Also letztendlich ist es erstmal von der Anbieterseite natürlich ein klarer Investment Case. Und letztendlich von der Konsumentenseite her ob ich jetzt ein Investmentbank bin oder meinetwegen jemand, der so so ein Treuhänder für gewisse Daten sein möchte, um eine Plattform zur Verfügung stellt, wo verschiedene Problemstellungen wie zum Beispiel vertrauliche Daten aus dem Gesundheitswesen, dass die für Forschungszwecke vertraulich zur Verfügung gestellt werden können und genutzt werden können, anbieten möchte. Der hat natürlich letztendlich immer die Maßgabe in gewissen Particle damit erfolgreich agieren zu können, und da ist halt natürlich die Frage einerseits sind die Teilnehmer dort gewillt, ihre Daten zu tauschen, das ist ja meistens so das Problem, dass man derjenige der Daten besitzt, diese ungerne shared, außer er hat einen eigenen Vorteil davon. Und ich würde sagen, das hängt sehr viel davon ab, wie das Modell letztendlich definiert ist. Also man müsste denjenigen dazu motivieren, seine Daten zu teilen, und das würde dann nur letztendlich tun wollen, wenn er da auch letztendlich raus ein Benefit für sich erzielen kann. Das heißt, wenn ich medizinische Studien durchführe und meine medizinischen Daten aus den Studien zur Verfügung stelle und diese medizinischen Studien sind sehr teuer, da wäre ja für mich jetzt letztendlich wichtig, dass wenn ich das tue, damit ich etwas zurückkomme, was für mich paritätisch vom Wert her ist. Das heißt, ich möchte letztendlich ein Return haben und sobald solange das nicht klar ist, glaube ich, werden viele Marktteilnehmer nicht unbedingt motiviert sein, ihre Daten doch zu teilen. Und auch wenn man das Angebot oder ein

Produkt, was man entwickeln möchte, dadurch besser machen würde, dann ist halt die Frage welche Risiken gehe ich damit ein. Könnte ich vielleicht ein Produkt selber so gut letztendlich definieren? Mit den Daten, die ich alleine habe, ohne dass ich Daten von anderen dazu brauche. Ist mir letztlich mein Vorteil dann halt größer, als wenn ich meine Daten teile. Zwar noch von jemand anders noch mal Daten dazu bekommen, dafür aber ein Risiko einzugehen. Ja, es geht jetzt letztendlich darum, wenn ich jetzt ein Automobilhersteller bin, habe ich sehr viele Daten von meinen Fahrzeugen, die sich weltweit bewegen. So und die versuche ich natürlich zur Weiterentwicklung meiner Full Self Driving Systeme am besten zu nutzen. So ich möchte da erster am Markt sein, der letztendlich mit den Daten vollkommen autonomes Fahrzeug am Markt hat. Wenn ich meine Daten jetzt sharen würde, könnte ich natürlich aus der denke oder aus so einer Burggrabenmentalität denken, ich helfe eher meinem Mitbewerber dadurch, dass ich den Pool an meiner Milliarden Datensätze mit jemand anderem Teile so. Und genauso aus der medizinischen Vorstellung wenn ihr zum Beispiel Apple. Wenn ich einen Fehler mache und zum Beispiel eine Apple Watch hab und meine Daten freiwillig mit Apple Teile, freut sich ja erstmal Apple. Aha, dann wär die Frage, ob Apple motiviert wäre, diese Daten dann anonymisiert in einem Pooling zur Verfügung zu stellen.

19 [0:21:05] **I:** Ja, genau mit SMPC könnten sie es ja dann zur Verfügung stellen und würden es aber nicht wirklich teilen eigentlich.

20 [0:21:19] **B14:** Ja richtig, aber jetzt gedacht, wenn ich den Nutzer meiner Apple Watch bin und die Daten letztendlich Apple zur Verfügung stellen würde.Würde dann Apple motiviert sein, die Daten von Millionen Milliarden Benutzern, die dort auflaufen, noch mal in ein Data Pooling mit anderen Device Hersteller noch mal zu teilen. Ich vielleicht als Einzel Nutzer möchte meine Daten mit Apple teilen, um Apple besser zu machen. Aber letztendlich, um das Gesundheitswesen nach vorne zu bringen, würde Apple dann seine Daten von allen Nutzern nochmal mit anderen Herstellern teilen oder mit Gesundheitsbehörden ein besseres Bild oder die Gesundheit der deutschen Bevölkerung zu bekommen? Ich glaube, da bin ich relativ skeptisch.

21 [0:22:01] **I:** Ja, das würden Sie wahrscheinlich nicht, aber wenn es jetzt ihnen helfen würde, angenommen jetzt zum Beispiel, man könnte Automobilhersteller, könnten ihre Produktionsprozesse miteinander vergleichen oder dass man sich über Cyberattacken austauscht, wenn verschiedene Unternehmen zum Beispiel ihre Daten teilen, wann sie wo angegriffen wurden durch irgendwelche Cyberattacken und des dann quasi miteinander teilen mit SMPC und dann hätten alle Vorteil.

22 [0:22:34] **B14:** Mhm. Wie ich das bis jetzt in meinem Berufsleben erfahren habe und jetzt auch aus verschiedenen Use Cases auch Datentreuhänderschaft auch im Bereich der Auto mobilhersteller,

dass sie letztendlich die Fahrzeugteile Daten teilen, an einen Treuhänder, um letztendlich diese Daten zur Verfügung zu stellen egal jetzt für welche Anwendungsfälle, dass die Daten zum Beispiel Bedarfsfall genutzt werden können, um ein Unfall Gutachten zu erstellen oder andere Dinge. Da war man in der Vergangenheit eigentlich nicht bereit, so das zu tun. Und auch bei den Bereich Cyber Security and Response. Dass man letztendlich für einen Incident dieser vorransichten Daten. Oder bei einem Incident forensische Daten zur Verfügung stellt, um Produkte besser zu machen. Oder um die Anatomie eines Angriffes zu erkennen. Bis jetzt war es halt so das Unternehmen sehr ungerne diese Daten zur Verfügung stellen. Einfach vielleicht auch aus dem Hintergrund heraus, dass man dort irgendwelche, wie soll ich das richtig ausdrücken? Ableitung machen kann, dass das Unternehmen meinetwegen nicht richtige Vorkehrungen getroffen hat.

23 [0:23:31] **I:** Also die Daten quasi gegen das Unternehmen, dann selber zu verwenden.

24 [0:23:40] **B14:** Ja ja richtig also, wenn man jetzt zum Beispiel sieht, dass erfolgreich Angriff erfolgreich war, einfach, weil ein System nicht up to date gehalten worden ist, weil zu viele Legacy Systeme doch im Einsatz sind und kein vernünftiges Patchmanagement für vorhandenes oder dass man auf gewisse Exploits nicht richtig reagiert hat, dann ist das sehr nachteilig. Jetzt könnte man natürlich überlegen, dass man diese also zur Verfügung stellen wie bei SPMC, dass man keinen Rückschluss letztendlich mehr auf Teilinformationen eines einzelnen Teilnehmers bekommen kann. Ich glaub, da muss aber viel Überzeugungsarbeit geleistet werden. Generell begrüße ich ja den Gedanken, ich bin dabei noch relativ skeptisch, ob die Unternehmen da wirklich sehr positiv drauf reagieren würden. Ich glaubte dieser Gedanke, dass das alles informationstechnisch, und von der Privacy her sicher ist, das wird relativ schnell akzeptiert, aber ich glaub da spielen auch einige andere Faktoren eine große Rolle. Generell sollte man ja gegenüber diesen Technologien sehr offen agieren, wenn die letztendlich helfen, entweder der Allgemeinheit zu dienen oder letztendlich auch bei dem Produkt Weiterentwicklung helfen kann. Als ich heiße das erstmal vollkommen willkommen, aber die Frage ist wirklich, ob das jetzt letztendlich Impact erzeugen kann. Spmc vielleicht nach langer Vorlaufzeit mit vielen positiven Use Cases, die doch dargestellt werden können, dann hätte das sicherlich ein viel größerer Erfolgschancen.

25 [0:25:11] **I:** Ja, ja. Positive Use cases sind glaub ich immer von Vorteil. Dann hätte ich eigentlich zudem nur noch die eine Frage und zwar ähm, bei welchen Sektoren oder Unternehmensgrößen oder allgemein Faktoren von Unternehmen würden Sie am ehesten Anwendung von SMPC sehen?

26 [0:25:24] **B14:** Ich glaub das kommt drauf an, wie tief die Wertschöpfung bei den Daten ist. Und ich denke, bei so einem Scoring System zum Beispiel im Bereich der Kreditvergabe, und es sind

besonders größere Unternehmen, die unheimlich viele Daten in kurzer Zeit für ihre Geschäftsprozesse verarbeiten wollen. Und von mehr Daten von Zulieferern gleichzeitig profitieren. Die sind dort viel Investitionsbereiter. Also im Finanzbereich und bestimmt auch in der Medizin. Weil ich das ist ja an sich, auch ein sehr großer Markt mit weiteren Entwicklungschancen. Auch im Bereich Telemedizin was ansteht und dass man die Leute, die heute einfach schon sehr affin sind, gegenüber von neuen Technologien, die tagtäglich auch über Smartwatches und andere Devices Daten sammeln können. Die dann halt unter SPMC Vorgabe verarbeitet werden. Ich glaub, da wird sie sich n Investment relativ positiv bemerkbar machen und deswegen war da ein großer Investitionsbereitschaft definitiv vorhanden.

27 [0:25:58] **I:** Okay, damit sind wir eigentlich mit allen Fragen durch. Vielen Dank!

28 [0:26:09] **B14:** Gerne!

1 Interview B15

2

| | |
|---|---|
| Interview-Nr. | 15 |
| Date of the interview | September 05, 2024 |
| Duration of the interview | 25:06 min |
| Interviewer | Felix Starnecker (I) |
| Interviewee | B15 (Sweden) |
| Role | Founder |
| Sector | IT-Security (Privacy Enhancing Technologies) |
| Specialities | No specialities |

3 [0:00:07] **B15:** Yeah, that would be easier. Because reading is faster than processing voice, right? And I'm a (..) I'm a good reader. Usually I, I said. It's good as well to mix it because then I can see your. Reaction as well, right? If I think that is really completely out of the blue or crap, then I I can see that. OK, maybe I need to rephrase it.

4 [0:00:25] **I:** Yeah, yeah. So now you can see the the screen.

5 [0:00:32] **B15:** Go ahead. Yes, I can see this.

6 [0:00:36] **I:** OK. So here you can see my topic again like corporate data sharing and what factors influence data sharing and the impact of privacy enhancing technologies. And 1st part is about general questions about data sharing.They are the first question would be like from your experience, with whom do companies share data? Is it more like competitor suppliers, customers and for what purpose? Most of the time?

7 [0:01:04.0] **B15:** Yeah. If you rank those three actors that you mentioned there, of course most of them are easier to share with their customer because anyway it is a potential for that monetization, right when they share data, they can actually monetize it as well. So that's why it's easier. The hardest part actually is against their competitors. So but this depends on which industry as well. So you have to analyse it from. The level of competitors competitiveness in the industry, right? And you can analyze as well from the economic side of it whether. They are actually operating in a thin margin, et cetera, et cetera, that that will color as well there.

8 [0:01:44] **I:** OK, which?

9 [0:01:45] **B15:** How open they are, yeah.

10 [0:01:46] **I:** Which industry would you say is the hardest or the the best to share with competitors?

11 [0:01:52] **B15:** For example, in the financial side, it's a little bit harder.
Banks, for example, right? Insurance company payment provider a little bit harder and that's why they have been pushed by regular regulation like PSD 2 in order to open a little bit right the data sharing.

12 [0:02.07] **I:** Mm hmm. Yeah, yeah. OK.

13 [0:02:08] **B15:** Because there's this open data initiative as well, right?

14 [0:02:12] **I:** And would you consider, like, trust an important factor for companies when they share the data? Like, do they have to trust other companies when they share the data?

15 [0:02:24] **B15:** Definitely. I mean, the trust is the is the sometime it's not being talked through in, in, in the surface, but that is the underlying theme, right. Whether I talk, you're not going to screw me, whether when I share the data with you, you're going to protect the data. You're not going to use it beyond the stated purpose, et cetera, et cetera. So there is this factor of trust in there.

16 [0:02:49] **I:** And the effect of trust is also needed if there are like control. Like if companies have control options, how with them with who they can control their data flows, they still need trust like it's always needed. You would say, yeah.

17 [0:03:08] **B15:** Yeah. You always like it. You know it. You know you you have the trust in verify, right? So you still have the agency like Independence as well. In order to check, right? Because if I ask you all the time, like, hey, have you treated my data? Well, you will most likely see it right. But if I have a method to verify it independently without me asking you yeah about that, then it generate a lot of trust. Because I I feel that I have the agency in order to test it myself. Right, so that is. Very important. I I don't really need to rely on it, OK? Trust me. Trust me, but I have a method if I independently, that is really important.

18 [0:03:43] **I:** OK. So you say with control options trust is created. Yeah. OK. And would you say that the decision and process of companies when they think about sharing their data with another company, is it subjective or objective decision making process like is it? One top manager who decides. Or I don't know is it a risk analysis?

19 [0:04:14] **B15:** Usually it's both subjective and objective, because there's human in the loop, right? And then usually there's a governance on each side as well. It's not only that you want to share, but

whether I want to keep your data as well. That's a question, right? Because you impose a risk. If I take your data dynamically, it's under my custody, right? And I might get snare into a legal entanglement in the future. If I don't really like be a good guardian. Your data, because we we we sign a legal contract for example, right? So there's this aspect as well that you know it's not really like one guy decide. Usually it's a. It's a team, a few executive. They're the one who's taking care of it, and also because you know, this is, sometimes it turns into existential risk, right? Because if you look at regulation, there's a hefty penalty, right? So when there's an existential risk to a company, usually you report to the board not only to the CEO, but there's a Direct Line to the board because. If you really screw this up, then yeah, you get half the penalty. I mean, if you read the news yesterday, Citibank changed their. Data. Executive because they get 136 million U.S. dollar penalty, right?

20 [0:05:29] **I:** Uh huh. Oh, we're fine.

21 [0:05:33] **B15:** They they have to revamp this. No, actually, they they didn't really kick him out, but they actually strengthened by sharing that role for the data management. So that that is how it is done. So it's never like a single person is making such a decision. As a consequential potential consequential damage to the organization. So.

22 [0:05:50] **I:** Yeah. So it's a always a risk analysis involved.

23 [0:05:58] **B15:** Yeah. And also I mean I think when you talk to me, I'm mostly my target segment actually is the big corporation. It's not really small, medium enterprise that don't really even have an IT system. Yeah, the CEO might decide it right. If it's only like 10 people organization, most likely the CEO is going to say yes and no, right? But if you're talking to AT&T, then like 200,000 people or Vodafone or ERT. Yeah, for example. Or in Germany, for example, right? So it's that.

24 [0:06:25] **I:** I want but I wonder probably the risk analysis is also a bit subjective, maybe because it's so hard to estimate the risks if you don't know much about it. And also like there are other risks like for your company then you if you share the data then you have some risks, but maybe other risks like cybersecurity risk can be. Made smaller or also like supply chain risks if you share data can make smaller so it's a better risk risk tradeoff like some risk get bigger, some smaller I would say.

25 [0:06:50] **B15:** Yeah. Yeah. No, no, but that's, yeah.

26 [0:07:00] **I:** Subjectivity.

27 [0:07:01] **B15:** But normally the big company they have quite thought through, it's less subjective in the big company because they are like consultant from BCG and all those things to create all this

risk framework in order to identify 25 and also mitigate those risks, right. That's why sometime they. Are a little bit like slow to adop. These kind of things because you have to go through this battery of tests and process as part of the, you know, CIA, the the the Triad and the security, right.

28 [0:07:33] **I:** Uh, huh. OK, OK. And what is the biggest risk you would see in data sharing? What can happen?

29 [0:07:36] **B15:** Yes. The first one actually when you share the data and then it contains some sensitive data, it could be like a personally identified information. There's a penalty, right? For example, there's a penalty to that, so that is one risk. So it's like more regulatory risk. The second risk actually is this right when you share the data, it's still your data you share with someone else. If they are sloppy with their management of your data then. Guess what? You get the penalty as well, even though it's not really you. Actually, but it is your partner. Who's you know, due to the sloppiness, they don't really manage the data in a in a well manner. Then you're on hook as well. The 3rd as well. If you actually screw this up, your brand is tainted. I mean this. I'm I'm talking more on on the segment that I'm trying to address, right? So this is like, yeah, telcos, banks and all this guy, right? They're public company. Most of them right, if these things happen, you know there's data breach when they actually collaborate with others. Then it's tainted the reputation. It's their brand, so it's very hard for me to trust a bank that they're leaking my data and has no good security posture. I don't really trust those bank and also it's really bad for their market capitalization, right, because they're a public company to attract investment, they need to maintain a good. Brand image in the investment market as well, so.

30 [0:09:08] **I:** Representation. OK. Yeah.

31 [0:09:13] **B15:** Those are just some example that I I know that those are the big risk.

32 [0:09:15] **I:** Yeah. And would you say companies overestimate the risks by data sharing or underestimate? Because I've read much that, for example, in Germany companies still share.
Way to less data compared to other countries maybe? I don't know.

33 [0:09:33] **B15:** Yeah. No, no. But that that is true as well. I mean, I I've been talking to Deutsche Telecom, Deutsch Bank and others guys, so I know. They they share very little, actually, and mostly they share almost the information level. So they have a data, not really information sharing. I call it. If you look at the the layer data and all the data information, knowledge and all those things. And so they they share little, but that is due to the fact, yeah, there's a risk, yeah. So that that's why they they don't share it at at will as well and also it contains some privacy data. So they need to take care of it and it contains sensitive data as well, right.For example, I can understand my competitor just by looking into their sensitive data. I can anticipate their growth, for example, by knowing their their

data as well, right?

34 [0:10:23] **I:** Yeah, yeah, yeah. And would you see any pre like? Pre require for successful data exchange.

35 [0:10:34] **B15:** Yeah, it usually is. It's like this. People share data because there's there's a business benefit, right? Either, for example, yeah, you. You you fight fraud. It just costly for them or you create a new product that actually generate revenue, right? So the economic side of it has to be taken care off before we actually even talk to, I mean about the technology, right? Because technology is just a neighbor, for example, using smpc homomorphic encryption or learning and all the stuff, it is just the how. Right? But. The the business benefit has to be there either to reduce risk, reduce cost or to increase revenue through collaboration. You know data sharing.

36 [0:11:16] **I:** Mm hmm. And do you see any external or internal influences that who encourage companies to share the data?

37 [0:11:25] **B15:** Yes, one of the strongest one actually is the top down from regulation. If the law and regulations say you have to share data, then they will share right for certain purpose.

38 [0:11:35] **I:** Yeah, yeah, yeah. OK. And internal stuff like inside of the company, are there some?

39 [0:11:43.0] **B15:** Yes, there there are. You know the data, right? I mean, you know this. I mean, you're paying on this one. The data is costly for the company to keep, right? If you don't really have a path to monetize it, this one and zero, you have to pay the electricity bill, right? So if you look at the the profitability of the company, the less profitable of them, the more eager they are to explore the potential. Yeah. To monetize the data to share the data in order to explore. Potential of data sharing.

40 [0:12:16] **I:** Mm. OK and. Now to the second part. About like the influence of secure multiparty computation on data sharing or in general privacy enhancing technologies. For which use cases would you consider privacy enhancing technologies such as smpc useful? More most useful or practical?

41 [0:12:45] **B15:** Yeah. If you look at SMPC specifically, you're compared to the other teniques, right? Because there are multiple tenants for the privacy and. Security actually strong. If you don't really have a lot of data that you can share because of the computational. So the overhead right when you actually share this data using this smpc, so that is usually the the main difference than the other, right? For example, so if you actually just consume like a small data in order to collaborate over this small data high speed data, then yeah, you can use it. And maybe not really so much in real time because of the computational require and also you need to guarantee the computational as well on

each node, right? In order to process that. So that is usually dictate as well. What use case that? Fit for this smpc. Slightly different. I mean, if you compare to Federated learning for example, or homomorphic encryption, maybe it's quite similar to snpc, even though it's slightly different application. But the simple one, for example, differential privacy. It's a little bit easier and usually you don't really need too much. Strong privacy requirement and then you can actually use a simpler 1, right?

42 [0:14:11] **I:** And would you say privacy enhancing technologies will change the data sharing behavior of companies in the future?

43 [0:14:19] **B15:** I think we. I mean, if you anticipate the future, yes, there will be. A need for the industry and society to collaborate more, right? So they need to prepare this whether they like it or not. Otherwise, they're going to be because the value of data actually is in relation to the other data, right? So this data enrichment thing, right? The value of my Facebook data is nothing, right? I mean but. Let's say my cohort. They actually tend to buy this and this and this in this location. During this time they spend this much that is more important for the advertiser, for example, in order to target me rather than my own personal data. Yeah.

44 [0:14:57] **I:** Yeah. So when if they have to share more data with other companies, then obviously they will need more like things like smpc, you would say like this kind of technologies, yeah, OK. And would you say the trust factor is still important if, like if you through smpc, you have more control over over data sharing, but would you say trust is still? An important thing.

45 [0:15:26] **B15:** Yeah, yeah, especially for a new technology, right? And as you mentioned, this is the level of knowledge in the organization that want to adapt, adopt it, right.

46 [0:15:35] **I:** But then the companies have more to trust, not other companies with who they share data, but more like companies like yours who offer this kind of technology. Maybe.

47 [0:15:42] **B15:** Yes. Yeah. I mean, of course that makes it easier for them to collaborate because they have a party, they have a counter. Party that is going to be liable, right? Satanically, we are on the hook, right? We offer them as a service, right in order to use this different techniques in privacy enhancing technology using smpc, Federated learning etc etc. Makes it easier for the adoption because then they trust us. Then they say OK. Well, you take the the technology risk, we take the business risk, right? So this is a kind of a way to address it because they don't really have the resources. Well, in order to pursue this themselves.

48 [0:16:22] **I:** OK. And what conditions are needed for companies to invest in the privacy enhancing technologies like? Size, sector, region.

49 [0:16:33] **B15:** Yeah, I I think I already mentioned about the profitability as well, right? We are willing to share willing to adopt the technology if they they need the money. That's one thing. So there's an economic reason for it. And the second one, due to the regulatory pressure that they need to do that way, for example, you want to collaborate with your advertising, for example in the attack. Then yeah, there's a regulatory pressure to do it in such a way to preserve the privacy, right? So then they have to use this technique in order to share data, because today they already share the data, but they share it. Then you know authentically, they surrender everything. But you need to consent and you need to protect all the privacy and all this thing so that that is another way to to and the region that region is different as well. I mean some region are more or we have to look into country by country because. This is related to the data protection law, right? So not so much on the region, but more on the country by country basis. Europe actually, with GDPR of course. It's quite advanced in protecting privacy, so as long as you can protect the privacy there. Share but in some country I mean across country, across border, across state jurisdiction. Data sharing is very difficult. For example, the Chinese is never going to let you to take their Chinese data outside into Europe, for example. No way. It's big issue or yeah, it's it's against the law.

50 [0:18:00] **I:** And. (..) And. Size of company. Would you consider this as important for using smpc like?

51 [0:18:09] **B15:** I think the smaller the smaller one is more. Easier for them to embrace, but whether they have a money actually in order to pay that, that's another question, but usually to embrace it, for example. They're more willing and also they are quick to make the decision because you only have like one or two stakeholder that you need to convince compared to if you try to convince Deutsche. You need maybe to talk to around 20 people in order to get their approval to get into.

52 [0:18:41] **I:** Mm hmm. OK. And how do privacy enhancing technologies change the risks of data exchange? Like are there?

53 [0:18:48] **B15:** Definitely it improve it, reduce the risk of dataclassion for sure. And. There. There is some risk related to the technology, right? Of course, there's like, yeah, I channel attack and all this different type of attack like data poisoning and all those things, right? But if you compare it sharing the data. Without the that technology, then of course it's very easy to compare. But the only thing that you can compare actually if they don't really share data, of course, yeah. It's just, yeah, it's it's easy to compare, but if you compare like if they share data. The current technics like. Yeah, of course this snpc and all this guy. They have the advantage.

54 [0:19:39] **I:** What would you say? Maybe data manipulation? That's like you share through smpc data with another company and then the the one company's not honest and puts like manipulated data into

the algorithm which is.

55 [0:19:44] **B15:** Yes. Yeah. No, that's that's poisoning attack as well, right? So you can actually poison one well and then they get injected into the the system, right? Of course, there's a risk, but of course there are ways well to address that kind of challenge. It's more like it's only a software engineering side as well. You can actually protect the integrity of the data as well.
And you can assume as well there's a bad actor, but usually in the in AB2B. Side if it's a corporate onboarding so then it's a known entity. It's not like anybody can just jump in and then share their data to be part of it, no. In the B2B. In the B2C it's made it slightly different. You can actually just, yeah.

56 [0:20:33] **I:** Yeah. And my last question would be what do you see at the as the biggest challenges for the acceptance of privacy enhancing technologies on the market? Like you know.

57 [0:20:45] **B15:** I think it it's the on the economic side because anyway we really need to.
Ring. So it's very we need to be selective on using the use case that really can highlight the the benefit of using this and then of course addressing as well some of the concerns related to the technology because this is quite an awful techniques and people a little. Bit skeptical. You know when when you ask the change in the beginning? So that is how we should, you know, there are a lot of hand holding knowledge sharing. So they understand as well they can test it themselves. They can verify it themselves. Then it's easier. But yeah, at the end, I mean, there's a computational cost as well, right? The calc. That's why I said first one is the economics, right? Because people don't really do technologies for the sake of technology itself, right? They're they're doing it either to reduce risk, reduce costs or increase revenue. Those are the three things. As long as you can find those three things and then you can quantify it and it has certain sides, then they're willing to to adopt.

58 [0:21:38] **I:** OK. OK then. Thank you very much for the interview. And that's where all my questions. I don't if you would say there is something else interesting I should have asked then yeah.

59 [0:21:58] **B15:** No, but I think that that is the that's the crux. Actually, I think I mean, you know as well you have been working in the technology. There's a way that we use it. There's a path to use it, but in in real life, usually it's a combination of few things. You know, maybe you need to combine smpc with encryption or inject a little bit of the noise into the system through a combination of smpc with differential privacy for example. In reality, we are. We are talking about that, that kind of approach. So rather than a single. And all of your single technique that can address all the requirement, because usually the requirement is not really like what we have in the academic world, right? It's very significant using all this, what you call it the the data that we generate ourselves. So it's a little bit slightly different than in in reality for example. I mean, I can give you on the telecom

side, right, on the telecom side, they process billion of trans. In real time. Can can smpc actually do it right? I mean, that's why I said it. It might apply into particular setting, but there are a lot of like engineering requirement. Let's say it is to stop fraud, right? You want to stop the fraud in real time, because otherwise it's too late. Somebody's already losing money, right? So and then they collect, let's say T-Mobile US, right? They process like. 1012 billion transaction per second, right? And your algorithm handle like such a performance requirement. That's a question, right? So we really need to look into this different aspect as well when we actually put it in in real life.

60 [0:23:35] **I:** OK, OK. Yeah, thanks for all, was a very nice interview.

61 [0:23:51] **B15:** Yeah very interesting, I mean, if you have the report later, right? I want to know as well the the outcome and then because you collect statistics right from different interviewers.

62 [0:24:21] **I:** Yeah. Yep, I yes, yes, I can send you the the thesis when I'm I'm finished. But it still needs maybe one month until it's finished. But yeah, I will send you a a copy.

63 [0:24:55] **B15:** Thank you, Felix. Have a nice day, Dan.

64 [0:25:03] **I:** You too

65 [0:25:06] **B15:** Yeah. Keep in touch. Yeah. Take care.

1 Interview B16:

2

| | |
|---|---|
| Interview-Nr. | B16 |
| Date of the interview | september 9, 2024 |
| Duration of the interview | 14:11 min |
| Interviewer | Felix Starnecker (I) |
| Interviewee | B16 (Netherlands) |
| Role | Tech Lead |
| Sector | IT sector (Privacy Enhancing Technologies) |
| Specialities | No specialities |

3 [0:02:27] **I:** OK, let's start. My first question would be like from your experience with whom the company shared the data and for what purpose? Like, is it competitor suppliers, customers?

4 [0:02:42] **B16:** So we are mainly active in the healthcare sector now. We also exploring all the sectors, but that's basically our our main focus. And at least in the Netherlands, companies already share the data. So it's it's like hospitals insurers. All kinds like health organizations, they already do share data. Umm, but often they use like trusted third parties to do that. So they basically trust one independent party. And to do this analysis, umm. And then, of course, they have, like, NDA's and those kind of things. But of course, of course, with the with the rise of multiparty computation, they are now looking into other ways to do this in a more privacy preserving way. But this the the purpose is mainly to improve healthcare. I'm seeing like because in advance the the generation which has a higher age. People are becoming a larger group of people in the Netherlands and then of course a more Healthcare is needed to show they try to really optimize the costs and how to help people in the best way possible. So that's mainly why they do it.

[0:04:12] **I:** OK. And it's no, uh, in general, would you consider trust as an important factor for companies when they share their data like that they trust each other?

5 [0:04:26] **B16:** Definitely. So the main of course they. Of course, the insights will be shared among the parties, but it's definitely in a important factor. You stay insights in the and useful way also by law and then once you also need to have like some grounds to do analysis, you can just do like, OK, I just want to analyze how data and it's done. You would really need some legal grounds to do this kind of things, so also based on those things.

6 [0:05:05] **I:** But like for example for secure multiparty computation, are there legal grounds already? I've heard it's not so sure if like for example like for a personal personalized data, if it's already legal to share with MPC.

7 [0:05:27] **B16:** Yeah, that is a very tricky. I think in the Netherlands it's very tricky. Uh, but of, but it's of course the law is a bit behind, at least in the Netherlands, and I think in the European Union the laws also bit behind. But they kind of like allow it for now. But you need to just for example it's stated in the GDPR and also they did touch variant of the GDPR stated that you need to have legal grounds to do these kind of things.

8 [0:06:14] **I:** Would be good to have clearance. But a general question again to the data sharing process of companies: When they decide if they should share data or not it would you say it's a subjective or objective decision process?

9 [0:06:24] **B16:** I think it's both. I think in many business decisions, it's like a personal connections. Of course, with people you know from other organizations, so that's of course kind of subjective, but there's also, of course, also companies have objectives or they need to comply to law to do these kind of collaborations. Sometimes companies are forced to do it, sometimes they are not. This also makes a difference in the decision process.

10 [0:07:01] **I:** Okay, and do you see risks in data sharing when it's done without privacy enhancing technologies? What are the biggest risks? Would you say?

11 [0:07:15] **B16:** So without discrimination partitions, data sharing happens already when you share the data. Of course, you make basically make a copy of your own data, and when you send it out, it's gone. So it's  out there, it's at the other party and you don't know what happens with it, of course. Even if the other party doesn't have any malicious intent, they can be hacked or there's all kinds of ways that you're at least you're not in control of your own data anymore. So that's definitely one of the big risks. And of course you also need to trust like a third party to do these kind of things. Umm, but that's it. It's the same problems that you have as well, but then at least even an extra party comes in where you also have an additional risks of leaking or your data.

12 [0:07:54] **I:** And do you think there are any like external or internal influences for companies to like share more data or also share less data or share no data?

13 [0:08:06] **B16:** What I explained already, I think the healthcare sector is getting more expensive than the glands and also like a lot of elderly people. Yeah, people are getting older and the group of older people are getting big is getting bigger. So more Healthcare is needed. So that's like a pressing thing to together create better healthcare and there's also from the government side. Stimulating these kind of collaborations as well, yeah.

14 [0:08:44] **I:** And would you say, would you say privacy enhancing technologies like SMPC or something else will change the data sharing behavior of companies in the future or in and in what way that they share more data or more sensitive data?

15 [0:08:57] **B16:** Umm, I think it's hard to say for now, but it's getting easier to do this kind of things. If they would share more. I think there's there are also other defendants in parallel to MPC that allows people to do more data sharing like of course, uh data science is not really coming up of course. So, uh, the companies are getting their data structures and having this process of preparing data. So that's of course an influence that allows companies to have this data ready to be shared. So I think it's a natural thing that more data will be shared, but it's maybe not due to MPC. I think the MPC part is mainly doing it in the privacy preserving way, but I don't think it's like ohh now we have MPC, now we can share more data on it's, it's.

16 [0:09:47] **I:** Yeah. And do you think like MPC changes like the need of trust between companies or does is it still needed?

17 [0:09:56] **B16:** Yes, and still need it, I think. So what we always say is like MPC is not the sale for bullets. If you don't use MPC in the right way, you can still leak data.So for example, you can do everything encrypted, but you say like hey, what is the average salary of this of a person? We use mail lives on this street on this House number. Then it's very easy to just extract informational one person, so it it's encryption is not everything. It's also about the questions you ask, so you definitely need more rounded and also has trust. Indeed, you still need the trust, yeah.

18 [0:10:41] **I:** OK, but what can you do except trust against such misuse of MPC?
Like can contracts help to avoid this?

19 [0:10:49] **B16:** Of course you have legal contracts, so you can also do things legally, but also you can do technical kind of enforcement to make sure that things are within the bounce that you agreed on.

[0:11:09] **I:** OK. And what conditions are needed for companies to invest in privacy enhancing technologies? Like are there companies who are more willing to invest than others? Is its sector specific or size specific or region specific maybe?

20 [0:11:22] **B16:** The first one is that MPC is still kind of new, right? So it's definitely the ones that are more innovative. So then they must have money to innovate and also they are already a bit mature and data. So some companies have a lot of data, but like what? I mean, with mature is like that they have their that's proper data warehouse and lakes and they know what kind of data they have in structured way and they know what value it has. Those companies are more willing to continue because if you just call random Company, we doesn't have data, but they have no clue how to even structure it or or share it or I don't know that kind of things. Yeah.

21 [0:12:06] **I:** OK, but financial resources, for example, would you say this is important?

22 [0:12:11] **B16:** Yeah, I think money is, of course one thing, but humans even more. Uh, that's even sometimes more important in these they're like, yeah, we do have the budget, but we don't have the people to actually do these experiments on innovation, yeah.

23 [0:12:31] **I:** And how would you say can privacy enhancing technologies change the risks of data sharing or like, are there some risk maybe getting like not so important anymore and some getting more important or even new risks of with these technologies?

24 [0:12:42] **B16:** The risk, of course, the risk that you're taking away is that you'll have to share your actual data. So you actually want to gain insights together and not really learn about it in lying data. So that risk is taken away. And of course it's applying to the law right, that the TRP, the GDPR, you have to comply to. So you comply into that, a new risks. Well, as I will say like you can still it's not a silver bullet, so you need more than only MPC to have the secure data collaboration. And of course, that's all new. So people need to learn about it, understand things. So there are definitely also new risks that are coming up.

25 [0:13:14] **I:** Okay and my last question would be like what do you see as the biggest challenges for the acceptance of privacy and hunting technologies on the market?

26 [0:13:21] **B16:** Oh, that's a good one. What we see is, like, uh, we do a lot of talks and these kind of things we trying to make companies aware like hey this there is like this new technology that's there

but it's still new and also very vague for people it's very hard to grasp. It is kind of like magic, right? And it's of course you can explain all the best medical things, but people just don't want to know and really not understand it. And so they think the concept is quite fake. And of course, like for now also like the legal, the legal things are also really at least in the Netherlands, like these personal identification numbers that there is no clarity on how it's how it will work legally. It's also kind of withholding companies to do it because they're like, I'm not sure if we're doing the right and otherwise we can get sued or whatever.

27 [0:14:05] **I:** Okay, this was my last question. Then thank you very much for your time.

28 [0:14:11] **B16:** Yeah. You're welcome.

Interview B17:

| | |
|---|---|
| Interview-Nr. | B17 |
| Date of the interview | september 9, 2024 |
| Duration of the interview | 23:06 min |
| Interviewer | Felix Starnecker (I) |
| Interviewee | B17 (Germany) |
| Role | Data governance research |
| Sector | IT sector |
| Specialities | Interview in german |

[0:04:58] **I:** Ja, also die erste Frage wäre aus deiner Erfahrung mit wem tauschen denn Unternehmen Daten aus, sind das eher Wettbewerber, Lieferanten, Kunden und wofür werden Daten ausgetauscht?

[0:05:15] **B17:** Okay, wir haben bislang immer grundsätzliche Datenteilungsszenarien unterschieden, nämlich zum einen nehmen Unternehmen mit öffentlichen Stellen, also B to G, bietet Business to Government das ist eine Bereich, das hast du zum Beispiel, wenn jetzt gerade im Mobilitätsbereich oder sowas wenn jetzt irgendwelche zum Beispiel Carsharing Anbieter durch bestimmte Auflagen, dass sie ich weiß nicht in welcher Stadt du bist, gibt es so E Roller und so die du mieten kannst?

[0:05:40] **I:** Ja

[0:05:43] **B17:** Die haben zum Beispiel eine Sondernutzungserlaubnis. Ansonsten dürften die hier gar nicht rumfahren und daran gebunden ist zum Beispiel, dass die Unternehmen bestimmte Daten anliefern, das heißt sozusagen einfach Teil von irgendeiner anderen Vereinbarung, die sie dazu zwingen, Daten mit öffentlichen Stellen auszutauschen. Es gibt auch andere zum Beispiel eben Mobility Data Spaces oder sowas die ja zunehmend auch kommen. Dass Daten auch einfach so geteilt werden müssen, wenn das denn so kommt zumindest im Gespräch ist das sozusagen, dass dieses Modell bietet. Dann haben wir das B to B Szenario, also Unternehmen mit Unternehmen und da haben wir natürlich erstmal die Regeln des freien Marktes, also wenn sich beide einig sind, dann machen sie das wenn nicht, dann lassen sie es so, das heißt wir untereinander einigen uns und eben die großen Fallstricke aus dem Wegräumen oder so abfedern, dass die Interessen gewahrt bleiben also konkret eben natürlich einmal der Datenschutz, dass keine personenbezogenen Daten geteilt

werden, dass es Probleme gibt oder eben und das ist auch der ein extrem gewichtiger Punkt natürlich die Geschäftsgeheimnisse, ja, dass die gewahrt bleiben, weil kein Unternehmen hat jetzt irgendwie Lust, mit einem Wettbewerber zu teilen, irgendwas irgendwelche Daten zu teilen, aus denen sich irgendwelche Informationen über die wirtschaftliche Leistungsfähigkeit ableiten lassen würden. So ganz entscheidender Punkt B 2 B und da ist im Grunde da haben die ja völlig freie Gestaltungsmöglichkeiten eigentlich also solange die sich einig werden, tun sie es nicht, dann lassen sie es eben. Und das dritte Szenario vielleicht abschließend noch um das komplett zu machen ist B to s also Business to Society im weitesten Sinne. Also das umfasst dann zum einen eben irgendwelche Open Data repositorien, also sozusagen die Daten werden von Unternehmen der Allgemeinheit zur Verfügung gestellt, damit wer auch immer daraus einen Mehrwert ziehen kann, das können eben engagierte Bürgerinnen und Bürger sein, zum Beispiel in Citizen Science pro. Oder zum Beispiel auch Forschungsinstitutionen. Und da gibt es auch dann noch mal eine ganze Spannbreite verschiedener Nutzungsmodelle von komplett Open Data bis hin zu: OK das und das Forschungsinstitut bekommt für ein bestimmtes konkretes Forschungs Projekt einen gewissen Datensatz, wo die Zwecke benannt sind zum Beispiel.

7 [0:08:01] **I:** Mhm, und im B to B Bereich als wie wichtig würdest du Vertrauen erachten also, wenn Unternehmen sich überlegen, ob sie Daten mit anderen Unternehmen teilen?

8 [0:08:17] **B17:** Na, also was heißt Vertrauen? Vertrauen auf einer persönlichen Ebene oder Vertrauen darauf, dass, dass die Interessen gewahrt bleiben? Und das ist der entscheidende Punkt also ob die sich kennen oder nicht, ist erst mal nachgelagert so. Es geht darum, die Verträge dann die Datenteilungsverträge im Zweifel so zu gestalten, dass. die Weitergabe zum Beispiel gesichert ist also zum Beispiel der Waschmaschinenhersteller A teilt seine Daten nicht mit den Waschmaschinenhersteller B, wenn er befürchtet, dass die Daten von A über B zu C weitergegeben werden. Was er nicht möchte oder dass irgendwelche Wettbewerber die daten erhalten.

9 [0:09:31] **I:** Und würdest du eher sagen, dass es ein subjektiver oder objektiver Entscheidungsprozess ist, wenn Unternehmen Daten teilen?

10 [0:09:42] **B17:** Ja, nur objektiv, nur objektive Parameter anhand deren einfach entschieden wird, ist es ein Wettbewerber, um welche Daten geht es, welche Informationen stecken da drin, was muss bereinigt werden? So der kriegt nicht einfach irgendwas, sondern der kriegt nen klar definierten Datensatz oder man fängt nur an überhaupt erst über einen klar definierten Datensatz zu sprechen, wo man klar ist OK. Die dritte Spalte, die muss da in jedem Fall raus, bevor wir hier anfangen, weiter zu sprechen, weil da stehen Summen drin Gelder zu, also Geldsummen zum Beispiel

Umsätze drin. Das ist völlig außer Frage, dass was weitergegeben wird zum Beispiel. Und danach geht's weiter in welchem Geschäftsbereich ist das Unternehmen tätig, warum auch immer was ist für diesen Geschäftsbereich überhaupt als Geschäftsgeheimnis relevant? Also es ist immer use Case abhängig, ne weil du hast die also ich weiß nicht ob das jetzt zu weit führt, sonst musst du mich bremsen die das Teilen von Daten da sprechen wir in unserem Data Governance Projekt von von dem sogenannten Wertrisiko Dilemma. Hast du davon schon mal irgendwas gehört?

11 [0:10:55] **I:** Ja, ich denk mal, wenn man Daten teilt oder dass man das Unternehmen dann abwägen oder zwischen dem was es ihm bringt, quasi Benefits und Risiken, oder?

12 [0:11:09] **B17:** Ja, genau das Problem ist, dass der Wert den Daten haben, nur sehr abstrakt zu beziffern ist, die Risiken aber sehr konkret von Anfang an. Der Wert zeigt sich eigentlich erst am Schluss ist extreme Use Case abhängig und du kannst eigentlich erst am Ende, nachdem du die Daten verarbeitet hast, kannst Du sagen OK das und das ist der konkret bezifferbare Mehrwert aber zu sagen ich tue Unternehmensdaten raus und damit probiere ich mein eigenes Geschäftsmodell. Die Risiken sind exorbitant. Häufig weißt du gar nicht, OK, der Datensatz, den ich jetzt rausgebe könnte der Wettbewerber oder die Firma XY, mit der ich das Teile in 20 Jahren auf die Idee komme, diesen Datensatz zu nehmen, ihn mit irgendeinem anderen Datensatz, von dem ich jetzt noch gar keine Ahnung habe zu kombinieren und schwupps fängt er an, mein Geschäftsmodell zu torpedieren oder zum Beispiel keine Ahnung bei öffentlichen Stellen und Mikromobilität irgendwie könnten die. Könnten die kommunalen Netzbetriebe irgendwann selber auf die Idee kommen, so ein, die die lukrativen Strecken mit zusätzlichen Bussen abzudecken und dadurch geht mein Geschäftsmodell in den Keller zum Beispiel ne also es kommt immer auf den Newecase drauf an die Risiken sind exorbitant hoch und der Wert, den so n, den die Weitergabe und damit auch Verarbeitung von Daten am Ende für irgendeine Firma hat sind eben ja, sind aber eben total abstrakt und das macht Unternehmen eben zurückhaltend im Datenaustausch.

13 [0:12:19] **I:** Aber würdest, würdest du dann sagen, das Unternehmen quasi die Risiken vielleicht etwas überschätzen und den Wert unterschätzen auch und deswegen vielleicht weniger?

14 [0:12:31] **B17:** Nee, auf keinen also, das glaub ich nicht.Weil die Risiken sind da. Ja, da ne wer könnte denn mit Sicherheit sagen, dass die die BVG in Berlin jetzt nicht in 5 Jahren anfängt irgendwie sich die sich die Daten mal genauer anzugucken und zu sagen okay auf der und der Strecke fahren? Ab 22:00 Uhr einfach super viele Leihroller durch die Gegend, woran liegt's naja und total nervig, weil die werden dann irgendwie auf den Bürgersteigen abgestellt. Und die älteren Leute stolpern da drüber und brechen sich ein Bein. Die Stadt hat ein Problem damit, und die

Verkehrsbetriebe kommen auf die Idee. Na ja, gut, dann machen wir mal ein paar Paar mehr Busse hierhin und verkaufen auch ein paar Tickets mehr und schwupps ist wird dem Geschäftsmodell das Wasser abgegraben.

15 [0:13:19] **I:** Ja OK. Und den Wert dann unter und den Wert unterschätzen?

16 [0:13:26] **B17:** Nee, ich glaube, da sind sich alle relativ einig, dass mehr Daten für alle grundsätzliche Mehrwerte bietet, so aber dieses Spannungsfeld navigiert werden muss zwischen dem Mehrwert für beide für Teilenden und Empfänger und den Risiken. Und da kommen dann eben solche Technologien ins Spiel, wo du sagst okay, naja eigentlich gebe ich die Daten gar nicht wirklich aus der Hand, sondern die werden eben irgendwo bei so einer Trusted Third Party irgendwie verarbeitet oder mit irgendwelchen weiß nicht incription Sachen, wo sichergestellt, technisch, technisch, organisatorisch eben schon sichergestellt ist, dass der andere gar nicht draufgucken kann, sondern der kriegt eben nur das Ergebnis und damit ist sozusagen diese gesamte Frage abgeräumt unter der Prämisse, dass ich dem Dienst vertrauen kann und das ist auch n bisschen schwieriger, wie wir gelernt haben ja.

17 [0:14:05] **I:** Und gibt es irgendwelche externen oder interne Einflüsse, die Unternehmen beeinflussen Daten zu teilen.

18 [0:14:15] **B17:** Externe Einflüsse wären zum Beispiel einfach Regulierung, also unter bestimmten Umständen sind Unternehmen ja gezwungen, Daten zu teilen und. Das kann zum Beispiel also wenn du, wenn du Praxisbeispiel haben willst so ne Sondernutzungserlaubnis sein, zum Beispiel für Mikromobilitätsanbieter in Berlin, die einfach dann gezwungen sind, das zu machen, weil sie ansonsten die Genehmigung nicht bekommen. Interne Faktoren natürlich quitt pro Quo also sozusagen Unternehmen A bietet Unternehmen B Daten im Gegensatz also im Tausch also gib mir deine Daten, ich geb dir meine Daten, da muss man dann irgendwie eine Übereinkunft finden. Unternehmen kaufen die ganze Zeit Daten von anderen Unternehmen ein. Das ist gängige Praxis und funktioniert auch insgesamt ganz gut so zwischen der Unternehmen, wenn die sich irgendwie eigentlich einig sind. Geschäftsentwicklung natürlich. Also wenn du aber das ist auch nichts Neues, weil wenn du ganz weit zurück gehst irgendwie in die Zeit der Versandkataloge und so weiter. Auch die haben sich schon bei irgendwelchen Adresshändlern bedient, so und da geguckt okay über. In dem und dem oder über über 70 in dem und dem Adressbereich, die kriegen die Kataloge für die Rentnermode so. Auch die hatten das ja schon ne deshalb also natürlich Geschäftsentwicklung Umsatzsteigerung, das sind die internen Faktoren.

19 [0:16:05] **I:** Für welche Use cases würdest du SMPC am ehesten sehen? Könntest du dir irgendwas vorstellen wo es sinnvoll wäre, wo es nicht sinnvoll wäre, vielleicht auch?

20 [0:16:20] **B17:** Ich was, was ich so gemerkt habe, ist große Interesse an einem an einem Zusammenschmeißen von verschiedenen Daten gibt es vor allem innerhalb von so Branchenverbänden. Ja, dass du sagst irgendein Automobilhersteller oder sonst was sammelt jetzt irgendwie die keine Ahnung streckendaten von den 10 größten Autobauern in Deutschland ein, die irgendwie mit Sensorik ausgestattet sind. Die schmeißen das zusammen und kriegen eine branchenübersicht, die sagen OK, jedes Auto ist im Schnitt soundsoviel zehntausendhunderttausend Kilometer gefahren, bevor es kaputt gegangen ist. Das sagt überhaupt nichts über einzelne Hersteller aus, bietet aber den einzelnen Herstellern die Möglichkeit, sich in diesen Durchschnitt zu verorten und zu sagen OK, die Wettbewerber sind so und so oder kannst du dir ungefähr rausrechnen, wo du da einfach stehst im nationalen Wettbewerb. So also das ja genau weil, da geht es ja erstens darum irgendwie zu aggregieren, also mehrere Daten zusammenzuschmeißen und dann eben zu sagen OK, aber wer sagt mir denn, dass der, dass der Typ aus dem aus dem Branchenverband nicht nächste Woche zu einem Hersteller zurückwechselt, so das ist ja auch sind ja auch Karussells und wie stelle ich sicher, dass diese Info nicht am Ende trotzdem irgendwo landet, wo sie nicht hingehört? Also das ist das unserer Erfahrung nach der tatsächlich interessante Case aktuell, dass du sagst irgendwie ja, es geht eigentlich darum, einen großen Überblick zu bekommen darüber wo stehen eigentlich die einzelnen Unternehmen? Und das, wo Unternehmen Daten von anderen wirklich gezielt einkaufen wollen, da gehen sie sowieso in den Dialog und gestalten auf ihre Teilungsszenarien so aus, dass es irgendwie für alle passt oder es klappt halt nicht ne, aber diese großen Aggregations und Verarbeitungssachen von den wirklich großen Mengen so das ist sind irgendwelche Aggregationen, zum Beispiel für Branchenübersichten.

21 [0:18:07] **I:** Und würdest du sagen dass oder glaubst du dass SMPC irgendwie das Datenteilverhalten von Unternehmen langfristig beeinflusst oder verändert? Oder glaubst du, dass also, dass es keine großen Auswirkungen hat?

22 [0:18:16] **B17:** Das das kann ich nicht so richtig beurteilen, weil ich das Verfahren jetzt nicht im Detail kenne. Die Frage, die sich stellt, ist einfach wie sicher ist das Ding? Und wie einfach ist es zu handeln?
Na, weil wenn es kompliziert ist, ist es nicht gut so, wenn du dafür irgendwie immer jemand externes brauchst, um einmal eine Verarbeitung zu machen auch nicht gut, weil kostet und deswegen aber für ja es kommt drauf an.

23 [0:18:44] **I:** Ja, und wie glaubst du verändert SMPC den Faktor Vertrauen?

[0:18:53] **B17:** Ja, also vielleicht muss ich das noch mal genauer ausschmücken also ich hab natürliches Vertrauen wichtig, aber ich glaube die Präzisierung die ich machen muss ist so ein blindes Vertrauen. Also ich sag nicht OK na ja gut also ja, der Felix ja der macht damit schon nichts so. Das passiert nicht ne, sondern es wird eben alles sehr genau geprüft und ein Risiko assessment gemacht und und das Ergebnis am Ende sagt ich vertraue, weil ich kann vertrauen oder ich kann nicht vertrauen, ne aber eben, weil es geprüft worden ist und weil die weil die Rahmenbedingungen klar festgelegt sind und das heißt dann eben okay unter diesen Bedingungen kann ich vertrauen und fällt ein Baustein davon weg nicht mehr. Vertrauen ist das Ergebnis also Vertrauen ist in dem Sinne eine Abschätzung auf Basis eines sehr intensiven Risiko Assessments.

24 [0:19:34] **I:** Könntest du dir irgendwelche neuen Risiken vorstellen, die entstehen, wenn die Technologie jetzt genutzt wird zum Datenaustausch.

25 [0:19:45] **B17:** Ja klar also neuer neues Glied in der Kette bedeutet immer. Potenzielle Leakage. So also wer wie wird sichergestellt, dass dieser Prozess sicher ist? Und das nichts rausgeht, wer haftet dann irgendwas passiert? Wenn du sagst dezentral OK, aber irgendwo werden die Daten mal übermittelt so was passiert da zwischen wem passiert das und wie ist das technisch, organisatorisch, irgendwie sichergestellt? Was passiert mit den Datensätzen in, also mit den meinetwegen auch verschlüsselten Datensätzen in 20 Jahren so wenn die auf irgendeiner Platte liegen, ja, oder? Vielleicht liegen sie auch nicht auf einer Platte lokal zumindest oder eben auf irgendeinem Server oder sowas ne keine Ahnung, aber was passiert damit? Wie lange werden die vorgehalten? Wer hat darauf Zugriff? Wie werden sie gesichert und was kann man damit vielleicht irgendwann alles anstellen? Also das, das sind ja immer die Fragen, die sich bei Innovation stellen ne also wer hat Zugriff? Wie sicher ist das ganze und wie einfach ist es zu händeln?

26 [0:20:47] **I:** Und was würdest du sagen, ist die größte Herausforderung für die Akzeptanz von solchen Technologien auf dem Markt?

27 [0:20:55] **B17:** Ja, du musst ein paar early Adopter irgendwie überzeugen von deinem Produkt und dann musst du irgendwann so eine kritische Masse erreichen. Das auch wiederum die, wenn sie Daten mit anderen Unternehmen teilen wollen, die Schlagkraft haben zu sagen na ja, wir machen das jetzt nicht so, wie wir es sonst gemacht haben, sondern wir nutzen dafür jetzt das und da müssen sich alle drauf einigen können deswegen, das ist einfach dann auch eine Marketing und

Verkaufssache. Und da muss man einfach clever sein also ich glaube natürlich, dein Produkt muss gut sein und sicher sein und alles erfüllen, was ich gerade gesagt habe, da musst du es clever verkaufen, aber klar, der Grundstein dafür ist, dass das Produkt selbst funktioniert, ne und da sinnvolle Ergebnisse liefert, dann auch irgendwie den Nachweis, dass das, was da berechnet wird, was wo aber keiner raufgucken kann richtig ist und wie man das zeigt, ist eine spannende Frage. Am Ende in der Lage sein, im Grundsatz zu erklären, wie diese Berechnung stattgefunden hat und wie das Ergebnis zustande gekommen ist, weil wenn zum Beispiel wenn zum Beispiel eine Verwaltung sagt, hier ist die ist die Luft zu schlecht, wir schließen die Straße, weil alle werden krank. Es übersteigt die Grenzwerte. Und am Ende kommt raus ist halt irgendein KI Modell berechnet basierend auf irgendwie 10 Luftaufnahmen. Über einen randomisierten Zeitraum oder sowas und ein Bürger klagt dagegen, dann kann die Verwaltung im Zweifel nur sagen naja, er hat halt das Modell errechnet, wie das genau und rennt dann zu dem Dienstleister, von dem von der KI Simulation oder so und die sagen können wir euch nicht sagen. Das Geschäftsgeheimnis proprietäre Daten ihr habt das Ergebnis eingekauft, aber den Rest verraten wir euch nicht hat die Stadt ein riesen Problem, aber sie muss erklären können und das heißt im Grundsatz muss dieses Modell auch wieder zu öffnen sein und also im Zweifel für bestimmte Anwendungsfelder muss es auch irgendwie in der Lage sein oder grundsätzlich möglich sein, das auch zu erklären und das ist dieses.

28 [0:22:09] **I:** Ja, das ist ein interessanter Punkt tatsächlich, weil ja eigentlich SMPC ist ja genau so ein Blackbox verfahren, weil man ja genau die Daten eigentlich nicht einsehen kann und somit eigentlich nie beweisen oder nachweisen kann, dass es wirklich das Ergebnis ist.

29 [0:22:17] **B17:** Genau ja, genau aber das. Ne, weil wenn ich am Ende n riesen Unternehmen bin und jetzt überlege mach ich jetzt die neue Fabrik in China auf oder nicht? Und steck ich da jetzt n paar Milliarden rein, da möcht ich schon wissen wie es funktioniert, also wie die wenn der wenn der wenn das Modell am Ende sagt Ja oder Nein die Verarbeitung und ich schmeiße alles Mögliche mit rein die ganzen Wettbewerbsdaten von allen möglichen Firmen in Deutschland, die jemals von dieser Frage standen angenommen, ich hab das alles ne schmeißt das alles zusammen, der Computer sagt Ja oder Nein, da möcht ich schon wissen wie das berechnet ist? Bevor ich da Milliarden in die Hand nehme, deswegen also das ist schon n kritischer Punkt, weil ansonsten wer stellt sicher, dass das funktioniert?

30 [0:23:01] **I:** Ja okay, das war es eigentlich schon vielen Dank für deine Zeit.

31 [0:23:06] **B17:** Cool, gerne gerne.

1 Interview B18:

2

| | |
|---|---|
| Interview-Nr. | B18 |
| Date of the interview | september 24, 2024 |
| Duration of the interview | 18:59 min |
| Interviewer | Felix Starnecker (I) |
| Interviewee | B18 (Netherlands) |
| Role | Founder \| Head of customer success |
| Sector | IT sector (Privacy Enhancing Technologies) |
| Specialities | No specialities |

3 [0:04:19] **I:** Yeah. So my first general question would be like from your experience with who do companies share their data? And for what purpose?

[0:04:43] **B18:** Yeah, I think you said this already. I think in the supply chain. To optimize their supply chain is a big one. In healthcare, of course for Treatment pathways to understand treatment pathways. But I think in general, if you abstract these two examples like supply chain or healthcare treatment path, you know in both cases it's to improve service where multiple entities are sort of working on the same on the same service. No, that's I think that the common denominator, if you are a diabetes patient. You visit 15 care providers in your patient life. Different types of care providers but or if you're manufacturing a complex piece of machinery, you know the manufacturing process is at least a 15 supply chain or 15 entities, and in both you just want to align those value chains. I think that's the that's the common type.

4 [0:05:47] **I:** And would you see trust as an important factor? That companies trust each other before they share data. And how is this trust created between them?

5 [0:06:14] **B18:** Yeah, I think, of course, the human side, the relationships, they matter a lot. Makes sense, right? I think if they have done previous business with each other. It helps a lot. And I think control, if they feel they are still in control during the collaboration.

6 [0:06:34] **I:** And like, from your experience, would you say that when companies think about sharing their data that it's more a subjective decision process or objective decision process? I don't know one person who just if he has a good feeling or she then then she decides to share the data. Or

is it like a long risk analysis until something is shared?

7 [0:07:02] **B18:** Yeah, I think you're right in both examples. One needs to have the courage and the progressive mindset to do it. Then this person needs to go through the pipeline of all the compliance checks before it can be set up. So you need both.
You need subjective trust. Courage, you know ambition. You want to drive report and then you have to do the whole compliance side. It both needs to be there.

8 [0:07:32] **I:** And what would you say are the biggest risks when data is shared for companies?

9 [0:07:43] **B18:** Which I think privacy penalties, competitive economic Disadvantage. They can be off like if you have trade secrets you want to you don't want to lose your trade secrets. If you have patient data, you know what you don't want to get penalties from the authorities. There are so many risks.

10 [0:08:07] **I:** And are there any external or internal influences?

11 [0:08:21] **B18:** I think in the healthcare it's regulators.

12 [0:08:24] **I:** So you would say it's sector specific then, yeah.

13 [0:08:25] **B18:** Yeah. Yeah.

14 [0:08:30] **I:** And then to the second part about privacy enhancing technologies. For which use cases would you see secure multiparty computation as most useful?

15 [0:08:58] **B18:** Now if you look at it on our website, it's the I think the domains we are most active in. So healthcare and sort of security and defense or public sector security and defense, I think those are for us the main focus area.

16 [0:09:11] **I:** OK. What would you say in general, that's the main topics where it could be useful.

17 [0:09:20] **B18:** It could be everywhere. You we have to focus, right? I just spoke to one of our competitors this morning and they he mentioned that there's a lot of also a lot of activity in the advertising space.

18 [0:09:45] **I:** Do you think that secure multiparty computation will change the data sharing behavior of companies in the future?

19 [0:10:04] **B18:** For sure, yeah. What we see is that companies that haven't that one are public sector organizations that wanted it for years are now finally doing it because of the security guarantees. They share more sensitive data.

20 [0:10:31] **I:** OK. And would you say trust is still needed also with secure multiparty computation?

21 [0:10:48] **B18:** Yes, but less so if. So if you're, if the technology is certified and you know audited and the company is audited, that helps trust the technology as soon as you can trust the technology and the technology enforces a lot of control. Remain in control as a partner in the consortium. Then I think you need less trust.

22 [0:11:21] **I:** And like what conditions are needed for companies to invest in privacy enhancing technologies. Or which type of companies are more willing to invest in privacy enhancing technologies?

23 [0:11:45] **B18:** It comes back to the question before, right? I think there where we and our competitors are active. Those are, I think, the ones that are now sort of you. If you look over the Internet, sort of what are the most live use cases? There's like this repository, I think from Princeton University. Have you seen it? There's websites from Princeton where use cases of MPC. It's not completely up to date because many of our use cases are not there, but you will see sort of advertising, healthcare.

24 [0:12:14] **I:** Mm, but you would you say it's like from the region, are there like countries where it's more like where the people are more open to the technology?

25 [0:12:29] **B18:** My from my own market research, I think the Netherlands is, is leading at the moment. There is the comparison. We have a German sales team and what we're seeing is that in Netherlands people are more willing to use innovative technology to set up a data collaboration, whereas in Germany the theme is more dozen shoots, right we need to protect and protection is critical. And collaboration is only an option. But we first need to protect and I think in the Netherlands and in the Nordics there's more traditional sharing data and the security is very important. But they're a little bit more willing. So I think the cultures per market differ a lot as I think the situation in the netherlands was favorable for a solution.

26 [0:14:22] **I:** And how do privacy enhancing technologies like especially secure multiparty computation change the risks of data sharing?

27 [0:15:01] **B18:** They reduce them. Change is a bit of a cliche. You're putting a new technology, so if in the German or in the in the GDPR, if you're applying an innovative technology for sensitive data at scale, it's a reason to do a data privacy impact assessment, so. And that's because it's an operative technology. So you have to weigh the risks of introducing technology on sensitive data. That's the only thing. But then if you make sure that without it and source code reviews and penetration tests, you know certification. We're trusting our ears to like other, you know, if the cryptography solution like the blockchain or the banks I think I come from cybersecurity background. These technology platforms are not always as secure as what we build. So of course, we're introducing risks if you're introducing new technology, but I think Technologies if they are being built with security in mind from day one on the ground up. They reduce risks, more so that they introduce risk. A lot of people who want who like kicking in open doors will say, yeah, but there's enough of this technology. So there are risks. Which I fully agree with. But you have to compare it with where we are with, for example, the banking system or advertising technology, which is a screaming privacy nightmare. So I would say there's only. There's only upside. There's only benefits for privacy and security. While you know if you want to sound clever, you would say, yeah, but there's always new risks to be perfected on us. I think if you have a decent technology that's built an audited authority, it's significantly reduces risk.

28 [0:17:43] **I:** Yeah. What do you see as the biggest challenge for the acceptance of privacy enhancing technologies on the market?

29 [0:17:54] **B18:** Communication because it's so difficult to explain. Simple and I think the solution I would say because what are you selling right? Are you selling a data collaboration tool or are you selling a secure database?
Or are you? What's the what's the story? What's the marketing slogan? We landed on encrypted data spaces or secure data spaces, which I think the situation where Europe is really trying to promote the concept of a data space is helping the MPC companies a lot. But this goes communicate, communicate with communicating that there is a solution for this problem. It's very difficult.

30 [0:18:46] **I:** Yeah. OK. Nice. That was it actually already.

31 [0:18:54] **B18:** Yeah, I hope it helped.

32 [0:18:59] **I:** Thank you very much for the interview.

Interview B19:

| | |
|---|---|
| Interview-Nr. | B19 |
| Date of interview | September 21, 2024 |
| Duration of interview | 32:15 min |
| Interviewer | Felix Starnecker (I) |
| Interviewee | B19 (Germany) |
| Role | Research Privacy Enhancing Technologies |
| Sector | Automotive industry |
| Specialities | Interview in german |

[0:00:00.0] **I:** Mit wem würden denn aus deiner Erfahrung Unternehmen generell Daten austauschen und wofür? So die Hauptanwendungsgebiete?

[0:00:11.6] **B19:** Ja, wir haben eigentlich immer so drei Fälle unterschieden. Also das erste wäre halt Cross Company Data Exchange. Ja, also gerade im Automobilbereich, Ich weiß nicht. Vielleicht hast du zufällig mal von Catena-X gehört? Ja, genau. Es ist halt ein Zusammenschluss aus Automobilherstellern mit dem Ziel, Daten über die Lieferkette auszutauschen. In dem Feld geht es halt hauptsächlich darum, dass Unternehmen mit Supplyern Daten austauschen. Aber perspektivisch gesehen. Im Prinzip auch Wettbewerber. Also dass man ein Use Case war, dass man zum Beispiel was hinsichtlich Batterieverschleiß, dass man die gemeinsame Auswertung macht. Weil es gibt ja häufig den Fall, dass die gleichen OEMs, also die gleichen Automobilhersteller, irgendwie ähnliche Probleme haben. Genau. Da ging es auch mal um solche Fragen, aber das war eher perspektivisch. Die Use Cases, die laufen, sind eigentlich meistens Lieferkette Use Cases.

[0:01:23.3] **I:** Ich habe sowas auch mal gelesen, dass jetzt die deutschen Autobauer vermehrt Daten auch miteinander austauschen, wegen der großen Konkurrenz aus China zum Beispiel.

[0:01:35.9] **B19:** Also dazu konkret, also zu der Motivation habe ich nichts mitbekommen. Okay, also ich weiß, dass die Use Cases, die laufen, die sind wie gesagt meistens über die Lieferkette. Also Use Cases, die immer mal wieder angedacht werden, sind auch unternehmensübergreifend. Es gibt ja auch im Bereich Multi Party Communication, ist ja auch ein klassischer Fall dieses Cross Company KPIs, irgendwelche KPIs, sagen jetzt, sei es jetzt Environmental Standards oder Gender Pay Gap oder solche irgendwelche KPIs, die man halt unternehmensübergreifend vergleichen kann, man das halt auf diese Art und Weise macht. Genau. Aber vielleicht also als erster Bereich Cross

Company Data Exchange, dann der zweite war Cross Country Data Exchange oder Data Analysis, also dass man dort halt teilweise Probleme hat. Also gerade im Financial Bereich hast du Kundendaten, meinetwegen in der EU und Kundendaten in China. Und wenn du die jetzt in irgendeiner Weise in einem Modell oder für statistische Auswertungen nutzen möchtest, ist es gar nicht immer so einfach datenschutzrechtlich die Daten über die Ländergrenzen hinweg auszutauschen. Genau da haben wir uns auch ein paar Sachen angeschaut, wie man Use Cases über Ländergrenzen hinweg datenschutzkonform ermöglichen kann. Und genau da gab es auch so ein paar Überlegungen, das Ganze zwischen Fahrzeug und BMW Cloud zu machen. Also da geht es ein bisschen Richtung Federated Learning, also dass du für bestimmte Anwendungsfälle halt sensible Daten hast, die im Fahrzeug aufgenommen werden. Zum Beispiel ist der Fahrer aufmerksam, was dann meinetwegen über eine Kamera erfasst wird und solche Daten müssten dann gar nicht unbedingt beim Automobilhersteller landen, weil das Modell eben auch angereichert werden kann mit Daten, die nicht im Klartext zur Verfügung stehen. Okay, das waren so die drei großen Bereiche, die mir so begegnet sind.

7 [0:03:57.0] **I:** Okay, und würdest du sagen, wenn jetzt Unternehmen abwägen, ob sie ihre Daten teilen, dass da Vertrauen ein wichtiger Faktor ist, mit welchen anderen Unternehmen man Daten teilt? Also dass man die irgendwie kennt, eine Beziehung zu denen hat.

8 [0:04:11.4] **B19:** Ich würde sogar sagen, dass es aktuell eigentlich fast ausschließlich über Vertrauen läuft, weil auch die Lösung, die bei Catena X eingesetzt wird, die aktuelle bietet keinerlei technisches Enforcement. Es gibt da zwar bestimmte Policies, die man festlegen kann, aber es hindert keinem aus dem Softwaretool, was da genutzt wird. Die Daten halt in den eigenen Space zu kopieren und im Prinzip alles damit zu machen, was man möchte. Also aktuell würde ich sagen, es ist eigentlich alles Vertrauen basiert bzw vertraglich dann halt irgendwie spezifiziert, dass halt wenn sowas herauskommt, dass irgendwo Daten missbraucht werden, dann halt. Ja, okay.

9 [0:05:00.0] **I:** Okay. Und würdest du dann sagen, wenn jetzt Unternehmen sich überlegen, ob sie ihre Daten mit einem anderen Unternehmen teilen und das ja auch auf Vertrauen basiertt, würdest du sagen, dass das dann ein subjektiver oder objektiver Entscheidungsprozess ist inwieweit Daten geteilt werden?

10 [0:05:30.3] **B19:** Ja, also ich ich würde schon eher vermuten, dass es subjektiv ist. Also ich glaube nicht, dass es da objektive Kriterien gibt zwischen denen, wo wirklich klassifiziert wird. Unter den und den Umständen teilen wir Daten mit anderen. Ich glaube, das wird auf individueller Basis entschieden und ich glaube, da ist viel Subjektivität dabei.

11 [0:05:48.8] **I:** Okay, und welche Hauptrisiken würdest du sehen beim Datenaustausch.

12 [0:05:58.3] **B19:** Na gut. Also einmal geht es natürlich auch um Betriebsgeheimnisse von Unternehmen. Das ist ein Punkt. Und ein anderer Punkt sind gegebenenfalls private Kundendaten, also dass man gegen Datenschutzverordnungen verstößt. Also ich würde sagen, das sind erstmal die zwei. So kann man erstmal zwei zitieren Und dann im Speziellen richtet es sich halt immer nach Use case, also wie hoch die Gefahr ist. Also dass die Daten, die man halt jetzt teilt, irgendwas vom Unternehmen preisgeben oder wie gesagt auf die Kundendaten Rückschlüsse zulassen.

13 [0:06:46.1] **I:** Ja. Und brauchen Firmen irgendwelche Voraussetzungen, um oder was wären die wichtigsten Voraussetzungen, damit Firmen Daten teilen können?

14 [0:06:59.8] **B19:** Also ich bin nicht hundertprozentig mit den rechtlichen Anforderungen vertraut, weil es gibt auf jeden Fall rechtliche Anforderungen und da gibt es dann Abteilungen, die sich drum kümmern. Aber vielleicht, weil du hier geschrieben hast, Voraussetzung für erfolgreichen Datenaustausch, also die Motivation, das über Cadillacs ein bisschen zu vereinheitlichen und eben Tools zu schaffen, um das zu automatisieren, wo man einfach noch mal einen ganzen Schritt zurückgehen muss, bevor man sich überhaupt mit Themen wie SMPC beschäftigt. Erstmal über Digitalisierung, weil viele dieser Prozesse, die nicht unbedingt digital abliefen zwischen mittelständischen Unternehmen und Automobilunternehmen. Und dann geht es erstmal um ganz andere Herausforderungen überhaupt. Also ja, dass Daten in einem einheitlichen Format sind, dass man eine vernünftige Schnittstelle hat, dass man die auch in Echtzeit abrufen kann und nicht irgendwie anfragt und dann kriegt man was. Wenn ich Ihnen auf einem Shared Drive in dem Fall und das ist halt überhaupt nicht vereinheitlicht. Also ich würde sagen, dass zum ersten Punkt sind erst mal so Voraussetzungen.

15 [0:08:24.7] **I:** So die unternehmensinterne Data Governance dann oder so ja.

16 [0:08:31.4] **B19:** Die Dateninfrastruktur hätte ich jetzt genannt, aber auch halt mit dem Hinblick darauf. Also große Unternehmen haben vielleicht solche Sachen standardisiert, aber gerade halt in diesen Lieferketten Beispielen, wo man mit Mittelständischen Unternehmen zusammenarbeitet. Da ist es halt nicht unbedingt gegeben. Genau das wäre das erste. Und und das andere, was dann auch immer eine Herausforderung ist, was halt das ganze Thema komplex macht, dass man immer mehrere Stakeholder hat. Also das heißt am Ende müssen alle Beteiligten im Endeffekt zustimmen über den Use Case und das macht es oft auch nicht so einfach, sich auf technische Lösungen zu einigen und also viel Abstimmungsbedarf.

17 [0:09:30.2] **I:** Das wird dann wahrscheinlich bei Secure Multiparty Computation auch ein großes Problem sein oder? Dass ja alle die Technologie quasi verstehen müssen und besitzen.

18 [0:09:39.6] **B19:** Genau. Also ich würde sagen, das ist sogar nicht mal gegeben bei den großen Konzernen. Ich würde es sogar so formulieren: Ich glaube, wir hatten sogar keinen Fall, wo wir auf eine Abteilung zugegangen sind, die in dem Bereich irgendwie Datenaustausch mit Zulieferern oder anderen Unternehmen hatte, die überhaupt wusste, was Multi Party Communication ist, weil das ist halt auch kein Thema, was so in den meisten Informatikstudiengängen vorkommt. Genau das ist auch noch mal eine ganz andere Herausforderung.

19 [0:10:25.5] **I:** Ich habe es auch bei meinen Interviews gemerkt, weil ich oft halt auch Leute einfach aus der IT Security Abteilung oder so angeschrieben habe. Und von denen kannten das Thema eigentlich auch die meisten gar nicht. Das hätte ich eigentlich schon erwartet. Also dass das zumindest mal gehört wurde oder so, aber ja, das ist auf jeden Fall nicht so bekannt.

20 [0:10:43.2] **B19:** Ist leider keine sehr bekannte Technologie.

21 [0:10:46.6] **I:** Ja und bei externen und internen Einflüssen. Würdest du da, würden dir da spontan welche einfallen, die Unternehmen dazu beeinflussen, ihre Daten zu teilen oder nicht zu teilen?

22 [0:11:00.7] **B19:** Also interne Einflüsse ist halt auch so ein bisschen das Thema, worüber wir gerade geredet haben. Also einfach wenig Leute sind mit den Capabilities erstmal überhaupt vertraut worden, man sagt moderner Kryptografie. Also das heißt, daran kann es schon scheitern. Die schreiben einfach Use Cases ab, nach dem Motto okay, ist nicht möglich mit dieser Anforderung, weil sie halt einfach nicht wissen, dass es eine Technologie gibt, die dazu passen könnte. Aber dann hatten wir auch teilweise den Fall, dass Abteilungen genau dieses Problem hatten. Also zum Beispiel, dass es irgendein Zulieferer gab, der hat eine Mail geliefert und wir haben da sensible Produktionsdaten. Wir möchten dem also nicht die Daten zur Verfügung stellen und möchten trotzdem irgendwas mit denen also der Zulieferer soll trotzdem was mit den Daten anfangen usw und dann ist trotzdem noch ein großes Misstrauen, würde ich sagen. Oder Ablehnung, halt wirklich so eine neue Technologie einzusetzen. Weil also erstmal du musst dir vorstellen, okay, du hast eine Abteilung, die macht seit Jahren irgendein Use Case, dann kommt jetzt eine andere Abteilung und sagt okay, es gibt da eine neue Technologie. Dann ist erstmal gar nicht unbedingt die Bereitschaft da, das erstmal zu akzeptieren, dass das die Technologie leistet, was sie verspricht. Aber auf der anderen Seite haben die Leute dann auch nicht unbedingt die zeitlichen Kapazitäten, sich einzulesen und das Thema zu verstehen. Also das heißt ja, im Endeffekt ist es dann auch eine Vertrauensfrage. Vertrauen

die jetzt darauf, dass sie ihre MPC funktioniert, weil sie die zeitlichen oder fachlichen Kapazitäten haben. Das überhaupt? Ja.

23 [0:12:51.0] **I:** Also Sie vertrauen dann quasi.

24 [0:12:55.5] **B19:** Genau. Also das ist das ist sicherlich ein interner Einfluss und dann auch also oft haben große Unternehmen wie [Unternehmen] halt viele Themen, die anstehen. Und wenn es jetzt heißt, okay, wir könnten Use Case XY mit Privacy Enhancing Technologies umsetzen, müssten uns irgendwie in das Thema einarbeiten und es wäre ein erheblicher Aufwand. Und so weiter und so fort. Dann wird halt oft auch gesagt okay, wir haben viele Themen und dann gehen wir halt das Thema nicht an. Also es ist selten, dass man einen Use Case hat, der so kritisch ist. Also es fallen immer Themen irgendwie weg. Und es ist selten, dass ein Use Case so kritisch ist, dass man den dann überhaupt angeht. Aber es gibt auch sehr viel Bequemlichkeit, würde ich sagen, dass man halt alte Technologien einsetzt, dass der Bedarf einfach nicht so hoch ist. Genau. Und dann im Prinzip muss man sich vorstellen, also oft haben die Abteilungen halt auch kein inhärentes Interesse, jetzt kryptographische Verfahren einzusetzen, sondern möchten eigentlich auch oft die Sachen. Einfach ist das teilweise auch lieber, die Sachen einfach vertraglich zu regeln, dann machen die halt irgendwie setzen die das vertraglich auf, dass man die Daten halt nur für einen bestimmten Zweck verwenden kann, weil das halt bekannt ist aus ihrer Sicht. Dann sind die halt aus ihrer Sicht abgesichert. Und das aus Konzernsicht vielleicht zu bevorzugen wäre, dass man halt die technologische Garantie hat, weil im vertraglichen Fall ist ja immer nur, dass die Garantie halt, wenn es rauskommt, dass dann irgendwelche Strafen fällig werden. Also dieses Interesse, das müssen die konkreten Abteilungen gar nicht unbedingt teilen, also aus deren Sicht ist es halt so okay, die haben alles richtig gemacht, wenn sie einen Datenschutz haben, der vertraglich abgesichert ist. Genau. Und wenn du sagst, externe Einflüsse, also da gibt es halt, würde ich sagen, vor allen Dingen rechtliche Bestimmungen. Aber wie gesagt, da bin ich jetzt nicht so hundertprozent vertraut mit. Aber das war halt immer bei den länderübergreifenden Use Cases vor allen Dingen der Fall. Aber das ist halt auch eine Herausforderung von Secure Multi Party Computation, dass zurzeit sich keiner so 100 % sicher ist wie das eigentlich hundertprozentig angeordnet ist. Und es gibt sogar Stimmen, die sagen also wir hatten da letztens auf einer Konferenz dann einen Vortrag von jemandem, der sich das rechtlich irgendwie angeguckt hat und der war sogar der Auffassung, was nicht heißt, dass das die allgemeine Auffassung ist, dass aktuell nach ich weiß nicht, ob er sich da DSVGO angeguckt hat. Daran kann ich mich leider nicht mehr so genau erinnern, aber dass aktuell verschlüsselte Daten nicht als Privatsphäre konform verarbeitet gelten. Also etwas, was nicht sinnvoll sein muss, aber was der rechtliche Stand wäre. Aber wie gesagt, es gibt einfach noch keine klaren Regelungen dazu. Ich glaube da kann man erstmal ein Fragezeichen hinter setzen. Also dass man halt sagt okay, intuitiv

gesehen würde man erwarten, dass es rechtlich als sicher gilt, aber in der Praxis kann das auch anders aussehen.

25 [0:17:58.3] **I:** Okay und für welche Anwendungsfälle würdest du denn SMPC jetzt am sinnvollsten sehen? Oder glaubst du vielleicht auch, dass es nur eine schöne Technologie ist, aber die wirklichen Anwendungsfälle nicht praktikabel nutzbar sind?

26 [0:18:19.7] **B19:** Also ich denke, es ist eigentlich sehr praktikabel nutzbar. Also ich glaube, die Komplexität von der Anwendbarkeit wird häufig überschätzt, weil inzwischen Compiler, die einem nahezu ermöglichen oder auch andere Interfaces, es müssen nicht unbedingt immer Compiler sein, die einem nahezu ermöglichen, seinen Code in nativ zu schreiben und der dann halt durch die entsprechenden sicheren Instruktionen ersetzt wird und man eigentlich nur grundlegende Sachen wissen muss, wie zum Beispiel. Okay, jetzt definiere ich halt einen Datentyp als einen Sekretär und jetzt wähle ich einen Char und dann halt darauf achten, dass ich nur die Daten, die ich am Ende regieren möchte. Also das ist meiner Meinung nach inzwischen ziemlich simpel und auch von der Laufzeit her gibt es eigentlich wenig Anwendungsfälle, die an der Laufzeit scheitern würden. Also wenn man jetzt nicht gerade große Modelle trainieren möchte oder Influence betreiben möchte, dann gibt es wenig Use Cases, die man aktuell nicht machen kann in praktikabler Zeit. Und für welche Anwendungsfälle das sinnvoll ist. Also ich denke, das sind im Prinzip erstmal die drei Sachen, die ich beschrieben habe. Also einmal Cross Company, Data Change. Also da gibt es schon prinzipiell Use Cases und auch Use Cases, die man besser lösen könnte als heutzutage zum Beispiel. Also ein klassischer Use Case, der halt gemacht wird und der auch rechtlich, glaube ich inzwischen gefordert ist, ist CO2 Aggregierung. Also dass man den CO2 Ausstoß beim Herstellen eines Fahrzeugs über die gesamte Lieferkette auf aggregiert. Und da gibt es irgendwie so ein Up on down Principle derzeit. Also dass man einfach sagt okay, mach das halt so in der Lieferkette, dass immer der nächste Supplier, also an dem ein Teil gesendet wird, der Krieg auf dem CO2 stand. Und da wäre es halt trivial eigentlich bei so einem einzelnen Fall zu sagen, macht es halt über Secure Multi Party Computation und dann hast du am Ende den CO2 aggregierten CO2 Stand und bist dann halt gesetzeskonform, aber zu keiner Stelle muss jetzt ein Unternehmen seinen konkreten Wert mitteilen. Jetzt kann man natürlich noch die zusätzliche Frage stellen okay, möchte man vielleicht sogar diesen CO2 Ausstoß von den einzelnen Lieferketten Teilnehmern vielleicht auch wissen? Klar, dann gibt es diesen Gedanken nicht. Aber ich denke, grundsätzlich lassen sich innerhalb der Lieferkette schon Use Cases finden, wo das grundsätzlich anwendbar ist. Also es muss da ein Use Case Bedarf geben. Und genau bei den anderen Sachen, also bei diesen Cross Country Themen. Also man ist innerhalb eines Unternehmens und möchte über Ländergrenzen hinweg etwas auswerten. Also diesen konkreten Fall, gegen diesen konkreten Bedarf, den gibt es tatsächlich. Also dass man sagt ja, zum

Beispiel im Bereich von Finanzdaten, man möchte länderübergreifende Modelle machen, gerade wenn man irgendwelche Länder hat, wo jetzt vielleicht die Datenbasis nicht ganz so groß ist, dass man sagt, man reichert das irgendwie mit den Daten an.

27 [0:21:54.4] **I:** Und gibt es irgendwelche Bedingungen, die erfüllt sein müssen für Unternehmen, damit der Datenaustausch mit Secure Multiparty Computing sinnvoll ist? Wahrscheinlich dasselbe, was du vorhin schon gesagt hast mit Datenqualität Schnittstellen und so.

28 [0:22:12.8] **B19:** Das dauerte aus technischer Sicht. Aber ich sehe eigentlich, die technische Sicht sekundär. Also das erste ist, die Leute müssen sich auf die Technologie einlassen und du hast wie gesagt mehrere Stakeholder, also Daten. Also ich würde sagen alle Use Cases, die Datenaustausch unternehmensübergreifend beinhalten, sind komplex, weil man mehrere Stakeholder hat usw und da ist dann halt einfach noch mal die Bereitschaft neue Technologien einzusetzen eh gering. Man hat eine geringe Awareness von CDU und FDP. Hat die Kommunikation auch, sogar Misstrauen teilweise, dass es wirklich das einhält, was es verspricht? Genau. Und du musst mehrere Stakeholder überzeugen. Also ich sehe diese Sachen wie Dateninfrastruktur und welches Framework man dann benutzt und kriegt man das effizient aufgesetzt. All das sehe ich teilweise eigentlich sogar eher als sekundär.

29 [0:23:06.3] **I:** Okay. Und glaubst du Secure Multiparty Computation wird den Datenaustausch verändern in der Zukunft?

30 [0:23:22.3] **B19:** Also ich würde sagen, das Potenzial ist da. Also ich denke, es gibt da zwei unterschiedliche Gründe. Einmal meine Betriebsgeheimnisse und einmal man schützt persönliche Daten. Also ich glaube, bei persönlichen Daten gibt es eigentlich nur das Szenario, dass der Gesetzgeber irgendwas vorschreibt oder irgendwie secure multiparting computation als sichere Lösung da akzeptiert, weil ein Unternehmen hat erst mal wenig Anreize, über gesetzeskonform Lösungen einzusetzen, um Daten zu schützen. Bei den Betriebsgeheimnissen sieht es anders aus. Also da braucht es einfach Vorreiter, denke ich. Also technologische Vorreiter, die sagen aus Unternehmenssicht okay, wir gehen jetzt mal so einen Use case an, der wichtig genug ist. Und deshalb, ja, deshalb würde ich sagen, es hat auf jeden Fall Potenzial, den Datenaustausch zu verändern. Aber ob der Fall eintritt, ist denke ich eine andere Frage.

31 [0:24:40.8] **I:** Und also so secure Multiparty Computational ist ja eine Technologie, wo man den Datenaustausch kontrollieren kann. Würdest du sagen, dass Vertrauen dann trotzdem noch ein wichtiger Faktor bleibt zwischen den Unternehmen oder kann der Faktor ganz rausfallen? Weil

wenn jetzt ein Akteur im im System nicht ehrlich agiert, vielleicht Daten manipuliert oder so und oder Daten schlechter Qualität in den Algorithmus gibt. Ob sowas vielleicht auch ein Risiko ist oder ob man auf sowas vertrauen muss, dass das der andere Partner nicht macht.

32 [0:25:31.9] **B19:** Hm, ja sicherlich. Also ich denke, aus der Pflicht kann man auch in zwei Fälle unterscheiden. Also du hast eigentlich so was wie entscheidungsgetriebene Use Cases. Das wäre sowas wie private Voting. Da ist jetzt das Problem, was du angesprochen hast, nicht relevant, weil wenn man ein Voting durchführt oder ein anderer Fall wäre private Actions. Wenn man ein Voting oder eine Auktion durchführt, dann hat man halt die Situation, wenn jemand seine Daten nicht ehrlich übermittelt, dann ist es halt praktisch die eigene Schuld. Weil wenn ich irgendwas vote, wofür ich nicht voten möchte oder wenn ich für irgendwas einen anderen Preis angebe, wie ich eigentlich zahlen möchte, dann bringt mir das erstmal wenig. Ja, also ich denke diese Art von Anwendungsfälle sind sicher. Aber gerade bei diesen ganzen unternehmensübergreifenden Daten Austausch geht es eigentlich darum, immer Daten zu verarbeiten. Klar, da besteht das Risiko, dass wenn man irgendwelche Daten in schlechter Qualität eingibt, dann gibt es da, glaube ich. Also es wird bestimmt Ansätze geben, aber ich glaube nicht, dass es eine verlässliche Art und Weise gibt, das wirklich zu unterbinden. Genauso aus aus der Hinsicht würde ich sagen, ist das natürlich was, was SMPC nicht garantiert. Also man kann zwar die Korrektheit garantieren bei bis zu -1 ist es die sag ich mal Adversis. Aber ja, also dass jemand Daten liefert, kann vorkommen und da braucht man irgendwo noch Vertrauen.

33 [0:27:21.7] **I:** Ja. Irgendwie ist es immer noch im Spiel, auch wenn wahrscheinlich weniger Vertrauen gebraucht wird als davor. Bezüglich Unternehmen, die du am ehesten siehst oder Sektoren, in denen es am meisten Anwendungsfälle gibt oder Unternehmen, die am meisten, am ehesten darin investieren würden. Was würdest du da sagen?

34 [0:27:48.7] **B19:** Ja, also ich würde sagen grundsätzlich je größer und je digitalisierter der Konzern, desto höher die Wahrscheinlichkeit. Also zurzeit ist es nicht realistisch von mittelständischen Unternehmen, dass sie halt so eine Technologie von sich aus einsetzen würden. Und zusätzlich wie gesagt, ist halt auch das Risiko, dass große Unternehmen auch, dass selbst bei großen Unternehmen gar nicht unbedingt da irgendeine besondere Kompetenz da ist. Also erstmal je größer und je digitalisierter, desto besser. Und an sich würde ich sagen, ist eigentlich Medical am interessantesten. Also dieses klassische Beispiel von man hat mehrere Krankenhäuser und die möchten Patienten und möchten jetzt gemeinsame statistische Analysen machen. Ich denke, der Anwendungsfall, der ist schon extrem relevant und den kann man auch beliebig ausweiten, wo es glaube ich für statistische medizinische Analysen extrem sinnvoll wäre und wo es viele private Daten gibt. Das einzige

Problem ist, dass da auch wieder, also der Medical Bereich typischerweise kein Bereich ist, wo Konzerne besonders digitalisiert sind. Genau deshalb würde ich sagen also aktuell beschäftigen sich eher Techkonzerne damit und am meisten Potenzial hätte es aber wahrscheinlich im medizinischen Bereich.

35 [0:29:47.6] **I:** Okay, das wäre natürlich sehr interessant. Ja und wie verändern sich die Risiken des Datenaustausches durch Secure Multiparty Computation? Würdest du neue Risiken sehen, die entstehen?

36 [0:30:05.0] **B19:** Also man müsste halt irgendwann sich den Aufwand machen, ein meinetwegen zertifiziertes Framework zu etablieren, das einfach da von der Implementierung sichergestellt ist. aber im Prinzip hast du auch bei den Protokollen kein Risiko. Ich glaube, da gibt es inzwischen genug, sowohl informationstheoretische als auch kryptographisch sichere Verfahren, wo nicht zu erwarten sind, dass die ja, dass die praktisch geknackt werden. Ich würde sagen, eher eine kleine Herausforderung, die es noch gibt, was eigentlich eher Aufwand ist, dass halt Leute sich wirklich mal damit beschäftigen, dass man halt wie gesagt die Frameworks irgendwie zertifiziert und danach. Also ich würde sagen, also es entstehen keine neuen Risiken, unbedingt. Aber man muss halt im Hinterkopf haben, dass sie ihre Multi Party Computation sicher teilt Input Privacy ab. Danach muss man sich überlegen, wie man Output Privacy sicherstellt. Und das ist halt ein orthogonales Problem, was man dann mit Differenzial, Privacy oder anderen Technologien lösen kann. Also dass man, wenn man Analyseergebnisse mit Secure Multiplication erhält, dass man da keine Rückschlüsse darauf ziehen kann auf die Inputs. Ja, genau. Zumindest spontan fallen mir erst mal keine neuen ein. Nicht einfach, dass man sich den orthogonalen Risiken noch bewusst sein müsste.

37 [0:32:01.6] **I:** Und die letzte Frage wäre jetzt noch was deiner Meinung nach die größte Herausforderung für die Akzeptanz von Secure Multiparty, Computation oder anderen Privacy Enhancing Technologien auf dem Markt ist?

38 [0:32:15.7] **B19:** Also eigentlich würde ich sagen, die größte Herausforderung ist Awareness also, wenn du auf das Problem ansprichst, dass du IT Security Leute fragst und nicht mal die haben von der Technologie gehört. Also aktuell ist alles, was Kryptographie und Datenschutz angeht halt im Prinzip auf den Stand von mehreren Jahren und man benutzt vielleicht irgendwelche simplen oder vielleicht sogar naiven Anonymisierungsverfahren, aber nichts, was eigentlich so Stand der Technik ist. Und auch solche Sachen wie Differenzial Privacy oder fully homomorphic Encryption sind ja auch Technologien, die nicht wirklich eingesetzt werden. Also deshalb würde ich sagen, die größte Herausforderung ist eigentlich Awareness.

1 Interview B20:

2

| | |
|---|---|
| Interview-Nr. | B20 |
| Date of the interview | september 25, 2024 |
| Duration of the interview | 25:56 min |
| Interviewer | Felix Starnecker (I) |
| Interviewee | B20 (Netherlands) |
| Role | Lecturer and Researcher |
| Sector | Research (Privacy Enhancing Technologies) |
| Specialities | No specialities |

3 [0:03:05] **I:** Yes. So my first question would be like with whom do companies share data? And for what purpose?

4 [0:03:11] **B20:** That's pretty broad question, but I think it really depends on the what do you say the case, the purpose or the you know the under agreement. But like for instance if you think about supply chain relationship then probably the company will share data with partners within the supply chain, for instance. It depends on really a use case and then it really depends on the context and the purpose is I think. Can be also really brought because let's say you could you do that a sharing for monitoring purpose, for instance monitoring the progress or the populations. But you can also do data sharing to unravel new opportunities, for instance, new streams with companies so that you need to share data you could you need to update data from other sources.

5 [0:03:49] **I:** OK. And would you consider trust as an important factor when companies share the data like that? They trust each other.

6 [0:03:58] **B20:** Well, I asked this question also back then when I did my PhD and in what I understand and I think I did when I'm at this point now. They would really look more into the business failure business benefit of doing the sharing, so trust control options would be something that nice to have. But I think if they don't see a benefit of sharing data, what's in it for them, then it's it already.

[0:04:32] **I:** Uh, But if they like, if they have benefits and sharing data and they don't do it because they are scared of the risks of sharing data, would it be then helpful that maybe like the top

managers know each other or they are like some relationships because of earlier projects together or something that they knew each other before, would it be helpful?

7 [0:04:55] **B20:** If you if you put it that way, then that could play a role because usually if companies tend to do businesses with partners that they already know, so they have some initial trust at the beginning. So that's that will play a role and control options. If you have this question here like contracts agreement it's a form of control mechanism as well to make sure that Data sharing or that exchange can still be performed within the corridor that is agreed. So I think that could play a role, or if you put it differently, you could also see that in principle, companies don't trust each other. And then control options are there to make sure that with the conditions of no trust, business can still be run as usual. So yeah, because in in, in many ways you can always see one company and another as some kind of competitor, although maybe it's not really obvious where the competitiveness, but I think companies can always say that they are really careful with each other to build the relationship. So the starting points that they're not really touch each other then then that's where the contracts agreements comes in as a control option.

8 [0:06:17] **I:** And would you say that it's a subjective or objective decision process when companies, companies think about data sharing? Like is it more risk analysis or is it like one or two managers who like just if they have a good feeling they say yes.

9 **[0:06:36] B20:** That's interesting. I would say it's more objective. So it's not because of the person, not always about the person. I think the could be driven by that, but I think the more dominant role would be the objective fun is that what's the risk and what's the benefit? View so that's of course top management support can play a role, but I think in the end it goes back to the business decisions like whether it's it will bring better revenue, higher revenue for instance or it will reduce some costs in that by doing that the sharing? So I think it's driven more by objective rather than the individual.

10 [0:07:37] **I:** And what would you say are the biggest risks in data sharing? And would you say they're overestimated by companies?

11 [0:07:49] **B20:** I think at least couple of years ago when I did this study, what I observed was three parts. So first there's a fear that companies will lose competitive advantage if they share data. So especially if the good data is pretty sensitive in that regard and. There's also concern that because usually if you share data, it's also involve consumer data, customer data. If you share it and then there there's a fear that those data will be leaked and there's possibility of misuse and in the end, it will harm their reputation in in broader in, in the, in the long run. Because what if there's some kind

of leak? And then customers are affected? So. So those implications are more creating some kind of. Make the other companies hold back in doing this data sharing. So if you look at whether it is overestimate of underestimated, that's a really interesting point of view because. I think in some ways there's it is. I'm not sure if the if those cancer rates are really big enough compared to, let's say if the if your if the factory of the company like OEM for instance if the factory are burned by fire, for instance. Those kinds of phrase are more tangible. Risks of the data is more intangible.

12 [0:10:17] **I:** So you say that you can't estimate the risks, right?

13 [0:10:24] **B20:** Yeah, yeah, I think. I think that would be. That would be the answer. Another way to see another way to look at it is that even when data sharing is not performed, companies still face a risk of data breach, right? There's, there's always those risk as well. Then that might create some kind of fear if companies participate in data sharing that there is more greater risk. I kinda think that is not something that you should really fear about, because if you don't stimulate the sharing, then you will lose more opportunities in the future to innovate. So I think there will be some kind of missed opportunity there. And that's also a kind of risk.

14 [0:11:41] **I:** And do you see any external or internal influences that encourage or decourage companies to share the data?

15 [0:11:55] **B20:** I think companies will really look at success, success stories like what are the successful use case and then they might follow those early adopters. For instance, I think data spaces are especially in German communities. I think they started to really take off like lots of initiatives piloting, although we're not, we don't know yet if it's going to really something big, but at this at the start, lots of initiative by ex Catana X you, you, you name it. So I think if they can create some kind of proven use case that it works.

16 [0:12:44] **I:** Okay now to the second part about like secure or privacy enhancing technologies and especially secure multiparty computation. For which use cases would you see secure multiparty computation useful?

17 [0:13:07] **B20:** I think it's yeah. MPC is an interesting technology, but one thing that I remember when I presented this work is that. It's like a solution looking for a problem. So some people would say that, well, it's interesting. It's interesting fascinating thing, but what is actually the problem that we want to solve? But yeah you could. You could look at it in different ways. I think in general MPC would be useful in the scenario when you have multiple organizations, multiple entities, you

typically don't collaborate with each other. But if you see but there's a need for better collaboration between those parties in order to address some kind of pressing common problems, then MPC could be suitable to breach those collaborations and then produce some kind of insights. Which, like you can generate something but you still can protect the data off the data owners. So I think it should be multiple parties, so it cannot be like homomorphic encryption which only one to one interaction. I think ideally I'm busy with, so if you have multiple potential collaborators. And they have sensitive data and they want to protect the way data is being processed or analyzed because I think it's important to know to be aware that MPC it only works on the. How do you say process or analytics layer? So do it protects the way you process the data by performing encryption and then splitting the data in those multiple parts. But it's not going to protect the input data that you provide, so it's not the scope of MPC, right? So if you put garbage in, for instance with the data of the companies, then you will get the garbage out. So the quality of the data still going to be poor even though you use MPC to analyze data. So that's also the issues of the limitations of MPC I would say. And then you need to have agreement at first on how so we want to use MPC. Do you agree to use it to collaborate together so those agreement to make collaboration is something that's beyond the technology itself, and very hard to get, because of multiples of different companies. So you still need complements in terms of like governance mechanisms to agreement. So that's why I think there are some exploration in the Netherlands with data sharing coalition. If you are familiar with that, they are exploring the idea of combining this data spaces with privacy enhancing technologies, which I think makes I think it makes sense because in data spaces you typically have a close group, right? Close sets of groups. Close sets of companies or organizations who already agreed to join and they have. Yeah, they pass some kind of test to join the data spaces. So then there's an agreement that they want to collaborate together in some way or some form, and then that's where prices and handling technology sort MPC get the leverage to like obtain some insights. It could be there's a competitor between those dataspaces, but they all they have some agreements in in some ways then it's more suitable to use it and the starting conditions for PETs are easier like this.

18 [0:18:09] **I:** And you said that you that the important thing is that you have multiple players who collaborate together, what are like conditions that must be fulfilled when like how do you get these multiple players together?

19 [0:18:26] **B20:** Conditions to make sure that that's interesting, alright. I think because the technology itself is still relatively novel, so probably those players are still not fully understand the technology. So but they had, they are interested in solving some problems together, but they don't really know the details on it. So I think the big challenge is to convince them that this is a technology. How it works? And we can show it that this is the technology that is relevant to solve their problems. What

are the limitations? What can it do and what can it what? What MPC can do and cannot do and what? What are the capabilities? And then what is the? How do you say? Yeah, the more of it. The governance mechanism, who can do what? So I think I think more about knowledge understanding the knowledge of how MPC works, let's say.

20 [0:19:40] **I:** And would you say trust is still an important factor when like MPC is involved? Or is it less important?

21 [0:19:51] **B20:** I think yeah, the important aspect here is that those people in the cryptography domain, they always say that MPC eliminates the need of trust because initially, if you have, if you want to perform collaboration, you typically have this middle man as trusted or trusted party. You'd say so with MPC we try to remove those trusted third. Parties. So you can still collaborate without having to trust. Each other. Important differences compared to this traditional ways that it's true that you can still collaborate with a trust or trust between organizations might not be relevant anymore, but I would say there's an increasing relevance of the trust in the technology or algorithm or whatever you wanna call it. So the it's more about the way this technology works. Or you cannot fully verify if. The input data that you provide is really not traceable to you, right? So I mean, because the promise of MPC is that you can generate insights, you can collaborate with others and then get something out of it, but others cannot see your data and you cannot see other data. So that's something that we don't really know how to verify that. I think there's also, but there's also growing. A field of verifiable MPC thing, but that's beyond what I research so far. Yeah, but still, that's a big. I think a big question and that might play a role in the trust decision, so maybe not trust in the other players, but more about how this technology works, whether it fulfill its promise or not.

22 [0:21:55] **I:** Yeah. And would you say there like coming or like MPC creating new risks?

23 [0:22:04] **B20:** Yeah. I would say there's a risk in terms of. I don't if this is a good illustration of risk, but. As I mentioned before, MPC has only handles the data processing layer. So how the data's being processed in the middle but not in terms of preparing the data from the start? So there's a risk of the input quality of the input data is compromised because I can still input fake data, not real. There's a risk that relates to the trust issues that I mentioned before risk that the data that is the result is produced is not high quality data because the input data is not controlled so. Manipulation and I think there's also possibilities of Reverse engineering. You could say that because MPC is based on the queries or questions that you want to answer, right? You want to solve some issues and then we need the input data and then the MPC perform computation that. But the queries that are the questions that we ask, it can say something about our companies or some informations or some kind

of resource that we have? Yeah. In some ways it can say something about the requester. So then it will create some kind of possibilities to reverse engineering and then based on that maybe they need this. They need that and then they can somehow figure out, although I'm not. I don't know if this is theoretically possible, but it could be there could be.

24 [0:24:11] **I:** OK. And what would you see like as the biggest challenges for the acceptance of privacy enhancing technologies on the market?

25 [0:24:19] **B20:** Yeah. I think the biggest challenge so far is still the lack of use cases I would say, and so far most of the projects are still very much in the pilot so-called pilot or trial or still in the in the small scale.

26 [0:24:45] **I:** So last question, would you think that like these technologies are changing the way of data sharing in the future?

27 [0:24:57] **B20:** Well, the way I see it, it's changing. It changed the way it can change the way we see data sharing because she may know it works, but yeah, versus MPC it works by allowing companies to still obtain something beneficial for them without having to give anything away. So that's in that way, it changed the way we share data because usually we need to reveal our input data and then others can also know the the our input data. But now with MPC we can we still have to provide input data but the result is not something that can say something about our company. So it changed the way we perform data sharing, but whether that is compelling enough for big companies or to create more impact then that's something completely different.

28 [0:25:47] **I:** Yeah. Yeah. OK. And then thank you very much for the interview and sorry that it took a bit longer.

29 [0:25:56] **B20:** Nah, that's OK. I also have to organize my thoughts about this. But it's interesting. It's interesting topic.