Multi-Party Computation as a Privacy-Enhancing Technology
Implications for Data Sharing by Businesses and Consumers

Agahari, W.

DOI
10.4233/uuid:83e0a9bc-f429-479a-ac2b-8a42e6da63f1

Publication date
2023

Document Version
Final published version

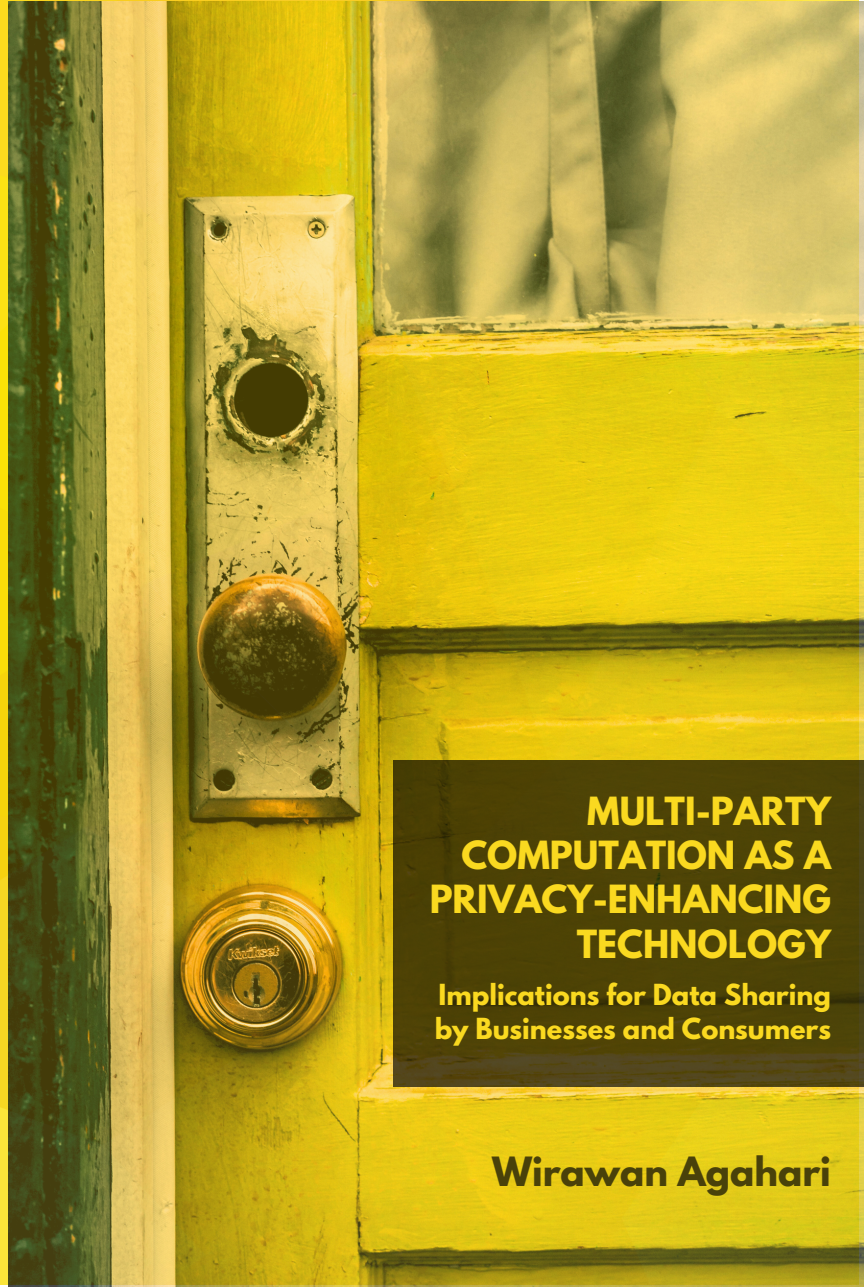
Citation (APA)
Agahari, W. (2023). *Multi-Party Computation as a Privacy-Enhancing Technology: Implications for Data Sharing by Businesses and Consumers*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:83e0a9bc-f429-479a-ac2b-8a42e6da63f1

Data sharing through data marketplaces, which rely on a Trusted Third Party (TTP), can benefit businesses and society. However, many companies and consumers are increasingly reluctant to share data due to mounting concerns over data control and privacy. Emerging privacy-enhancing technologies (PETs) like Multi-Party Computation (MPC), which enables joint computation to generate insights while keeping the input data private, could address data sharing barriers due to its differences with the traditional data sharing approach relying on a TTP. Thus, MPC could challenge the current understanding of why and how businesses and consumers share data. Nevertheless, whether businesses and consumers would be more willing to share data with MPC in place is unclear, as less attention is given to the socio-technical implications of MPC on data sharing decisions in data marketplaces and its antecedents. This research aimed to theorize the socio-technical implications of MPC on sharing through data marketplaces, by investigating how MPC potentially impacts data sharing antecedents by businesses and individuals. We do so through a mixed-method research design focusing on the automotive industry. Based on interviews with 15 MPC experts, which were structured using a Unified Business Model framework, we explored value propositions enabled by MPC use in data marketplaces. These value propositions allow MPC to potentially impact control, privacy, trust, and risks as antecedents of data sharing decisions in data marketplaces. Subsequently, we interviewed 23 automotive industry experts to investigate the potential impact of MPC use in data marketplaces on control, trust, and risks as antecedents of business data sharing. We then conducted an experiment via an online crowdsourcing platform with 1457 participants to investigate the potential impact of MPC use in data marketplaces on control, privacy, trust, and risks as antecedents of consumer data sharing. In this way, we contribute to the socio-technical understanding of MPC beyond technical perspectives. At the same time, we also demonstrate the relevance of MPC to practitioners by pointing out key aspects that should be considered while exploring the possibility of implementing MPC. Furthermore, this research provides a foundation for future studies on understanding the socio-technical implications of MPC on data sharing decisions.

Keywords: Multi-Party Computation, privacy-enhancing technologies, data marketplaces, data sharing, mixed-method research, automotive industry

MULTI-PARTY COMPUTATION AS A PRIVACY-ENHANCING TECHNOLOGY

Wirawan Agahari



MULTI-PARTY COMPUTATION AS A PRIVACY-ENHANCING TECHNOLOGY

Implications for Data Sharing by Businesses and Consumers

Wirawan Agahari

MULTI-PARTY COMPUTATION AS A PRIVACY-ENHANCING TECHNOLOGY:
Implications for Data Sharing by Businesses and Consumers

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus, Prof.dr.ir. T.H.J.J. van der Hagen,
Chair of the Board for Doctorates
to be defended publicly on
Friday 29 September 2023 at 10:00 o'clock

by

Wirawan AGAHARI

Master of Science in Engineering and Policy Analysis,
Delft University of Technology, the Netherlands
born in Jakarta, Indonesia

This dissertation has been approved by the promotors.

Composition of the doctoral committee:

Rector Magnificus,	chairperson
Dr.ir. G.A. de Reuver	Delft University of Technology, promotor
Dr.-Ing. T. Fiebig	Delft University of Technology, copromotor

Independent members:

Prof.dr.ir. M.F.W.H.A. Janssen	Delft University of Technology
Prof.dr.ir. P.H.A.J.M. van Gelder	Delft University of Technology
Prof.dr. T. Li	Erasmus University Rotterdam
Prof.Dipl.-Inf.Dr. S. Lindstaedt	Graz University of Technology, Austria
Dr.ir. C. Hernandez Gañán	Delft University of Technology

This research was funded by the European Union's Horizon 2020 research and innovation program under grant agreement No 825225.

Keywords: Multi-Party Computation, privacy-enhancing technologies, data marketplaces, data sharing, mixed-method research, automotive industry

Cover design by Shauma Lannakita Tamba and Mahira Shaliha Agahari
Cover photo by Tucker Good on Unsplash
Printed by Gildeprint – https://www.gildeprint.nl/
ISBN: 978-94-6384-480-2

An electronic version of this dissertation is available at https://repository.tudelft.nl/

Preface and acknowledgments

Doing a PhD is something I never thought I would do, let alone finish it. But here I am, after so many ups and downs, I completed my academic journey at TU Delft, one of the best universities in the world. What is important to note is that, despite being titled as an “academic” journey, I would say that doing a PhD is so much more than that, as it shapes me to become not only an expert in my field but (hopefully) also a better person, husband, and father. For that, I would like to give all praise and gratitude to Allah Ta’ala for His love, mercy, and never-ending guidance so that, after four years, I can finally cross the finish line of this amazing PhD journey.

Finishing a PhD is like completing a 1000-piece puzzle. It is complex, messy, and scattered. We often have no idea how to finish it, so we should focus on what we know: the corners and the four sides of the puzzle, then focus on specific parts of the puzzle as time passes. Once we finish the puzzle, it is definitely going to be very rewarding, but to achieve that, we need to be patient because setbacks often occur throughout the process. So, the support from everyone around us becomes crucial in finishing the puzzle, and it would be very tough to do so without them. Reflecting on my PhD journey, finishing it would be impossible without endless support from everyone who has become my support system and helped me through it, so I would like to dedicate this section to acknowledge their contributions and continued support.

Mark, I still remember our first interaction when I met you to discuss possible ideas for a master’s thesis back in 2015. I was rushed and underprepared, which I thought was not a good look then! I am glad this was not the case, and we maintained good contact even when I was already working in Indonesia. Words cannot express my gratitude to you for allowing me to pursue this opportunity. For four years, I always enjoyed our serious and not-so-serious conversations about PhD, academic career, and life in general. I like that you always asked me, “How’s life?” on top of everything

else, showing that you truly care about your PhD students. Also, thanks for all the eye-opening guidance, suggestions, feedback, comments, and lessons! I am truly honored to be supervised by you throughout this PhD journey, and I hope we can still collaborate in the future.

Tobias, in the early days of my PhD, I was initially nervous about working with you. However, as time passes, you are easily approachable for discussions and provide valuable input on my work when I ask you to. You always forced me to be critical while guiding me in addressing certain issues. What's more, I can see that you genuinely care about the well-being of your PhD students (myself included) and always stressed the importance of work-life balance. Many thanks for your support and guidance all this time, and hopefully, we can find ways to work together again in the future.

I would also like to express my sincere gratitude to the members of my defense committee. Thank you for investing your time in becoming part of this important milestone by providing valuable comments and suggestions to improve my dissertation. I appreciate your interest in my research and making yourself available to attend my defense. I am truly grateful and honored!

My PhD would not be possible without massive support from the people of the Safe-DEED project. To fellow project partners, especially Mihai, Patrick, Gert, Lukas, Mihnea, Alessandro, Alexandros, Ioannis, Evangelos, Michael, and Leonie, thank you for the fruitful discussions and valuable insights throughout the project. I learned a lot from all of you, and hopefully, we can collaborate again on future projects.

Big thanks to former and current colleagues in the ICT section. To Alexia and Klara, my first office mates in room B3.250, thanks a lot for your warm welcome! You helped me a lot to get used to the new office environment. To Marijn, Yao-Hua, Nitesh, Harry, Anneke, Aaron, Roel, Jolien, Nadia, Caroline, Jacobien, Marcela, Jacopo, Fernando, Boriana, Iryna, Hosea, Marcus, Ali Latifi, Ini, Sem, Wiebke, Thierry, Dewant, Eva, and

Elyas, I cannot thank you enough for every collaboration, conversation, lunch talk, or even simple interactions that we had during my PhD, which made me grow and develop as a researcher and a person. Also, big thanks to former and current secretaries of the ICT section for becoming the first point of contact regarding everything that I need (and everyone in the ICT section!). Laura, Jo-Ann, Ellen, Minaksie, and Fanny, I cannot appreciate all of you enough! Keep up the great work in arranging many social events for the ICT section!

To other TPM PhDs, postdocs, and colleagues that I have known during my time in TU Delft: Steven, Jan-Jelle, Brendon, Elsa, Anna, Vladimir, Ivan, Trivik (my PhD mentor who always welcomes me every time I need to talk), Nicolas (the very amazing TPM data steward!), and Montijn (although you are from TU/e :)), thank you for your help, support, and opportunity to interact and learn from you guys!

I am also grateful to have worked with several student assistants and master students throughout my PhD research. Bas, many thanks for supporting me in transcribing interview recordings. Your help is really essential for me to make things easier in conducting qualitative analysis. Iris, your contribution to the Safe-DEED project is enormous! Thank you for your detailed analysis and comprehensive writing for our user experiment report! To Riccardo, Masud, Christian, and Kenny, thanks for your interest in my research and for eventually conducting your master's thesis project within the MPC domain. Your work provided important insights for my research and helped me a lot in shaping this dissertation. To Evelien, although your master's thesis is not about MPC, I am grateful to be involved and learn something valuable from your work. I wish you guys a very successful career in the future!

Special thanks go to Antra and Gijs as my paranymphs. Antra, I genuinely enjoyed our conversation on PhD, embracing the future as academia or other less important stuff. I had a great time collaborating with you on papers, and you helped me a lot in a pilot interview for my qualitative study. It has been an honor to know you and work together

with you! Gijs, although we only met on Thursdays (sometimes on Fridays), so much that I learned from you in a short time. Your encouragement and support really matter to me, especially during the toughest times of my PhD. Your local wisdom is truly beneficial in making my life in the Netherlands meaningful and enjoyable. Also, thanks for participating in the pilot experiment and providing suggestions to improve the survey. I cannot ask for better colleagues and paranymphs than both of you. Thank you, and I wish you all the best in finishing your PhDs!

Living abroad is not an easy feat, but I am grateful to be surrounded by Indonesian communities that helped me and my family a lot in making our life in the Netherlands feel like home. Many thanks to fellow Indonesians in Delft (among others, mas Dhata and family, Albert Stroop and family, Wicak and family, Mikhta and family, mas Sigit and family, mas Nasikun and family, mas Marwan and family, mas Agung Wahyudi and family, mas Tofan and family, mas Sebrian and family, pak Luthfi and family, pak Aries and family, mas Arry Rahmawan and family, mas Habib and family), Indonesian PhDs in Delft (among others, Bramka, Aldy, kang Kusnandar, mas Arie Purwanto, mas Agung Indrajit, Albert Santoso, Alfan, mas Yogi, mba Etty, mba Liza, Fira, and mba Reni), PPI Delft (among others, Sasa, Panji, and Devano), KMD/Delft Muslim Family (among others, mas Gilang, mas Rifki, mas Dhoni, and Hanif), Islamic study group Al-Hikmah Mosque in Den Haag (pak Hasyim, mas Nunu, kang Dani, mas Ryan, mas Ridwan, mas Henk, mas Luis, Jetmir, Tom, and Sindu), and IA-ITB the Netherlands chapter (among others, mas Raymon, mba Intan, kang Rihan, pak Eko, pak Bambang, Ryvo, Vira, Amay, and mba Monik).

To my closest friends in high school (among others, Rachmat, Nana, Tasya, Lya, Gagas, Evin, Andra) and Institut Teknologi Bandung (among others, Aul Virnanda, Ijul, Sandy, Manaor, Aud, Caca, Farid, Hakim, Irsyad, Mushofi, Pedca, Sukma, Mo), thanks for keeping in contact even after more than ten years! I won't be here without meeting and interacting with you guys. I will always cherish those memories. To my former colleagues from CIPG (among others, mas Yanuar Nugroho, mas Irsan, mas Fajri, mas

Nardo, mba Mona, Teteh Ria, Inal, Zya, Daya, Thesa, Vitha, Natasha, Fildza), thanks for keeping me close to the research world. To Fuad, my best friend, it was no coincidence that my PhD defense was right after your birthday. Probably, it shows how connected we are! Thanks for all the intellectual discussions about, among others, how we can improve Indonesia's education system, which I think plays a key role in driving my PhD to the finish line. And thanks definitely for many laughs and silly talks! Let's stay in touch.

My special appreciation to my parents, Ibu Atiek Larasati (mom) and Ayah Dedy Permadi (dad). Many thanks for your love, support, and prayers throughout my life. To my parents-in-law, Tulang Murlan Tamba (father-in-law) and Nantulang Seri Maulina (mother-in-law), thank you for the warmest support and never-ending encouragement to finish this journey. To all my family: Thika, Kak Ica, Bang Ardi, Iki, Tasya, Atha, Arshaka, and the rest of the extended family, thank you for the unconditional support, prayer, and encouragement during my PhD.

Finally, my biggest love and hugs goes to *mijn lieve vrouw*, Shauma Lannakita Tamba (Kitty). Thank you for your endless love, care, and support in good times and bad times. I am grateful that you always believed in me throughout this journey and pushed me to go beyond my limits. So, thank you for doing this together, and I am excited to find out what the future has in store for us. To *mijn lieve dochter*, Mahira Shaliha Agahari (Aira), thank you for all the memories and happiness you created. You create meaning for our lives, and you show us that there is more to life than just work. Let's continue making beautiful memories, and may Allah always be merciful and continue to give grace to our little family. Last but not least, to our unborn baby, who passed away after six weeks in Bunda's tummy, I also dedicated this dissertation to you. I hope Allah can reunite the four of us together again in Jannah.

Delft, 30 August 2023

Wirawan Agahari

Table of contents

List of figures

List of tables

1 Introduction

1.1. Background

Data has become one of the most important resources in the world that drive our economy and society (European Commission, 2020). In 2025, the global data volume is predicted to reach 175 zettabytes, while the economic value of data is estimated to reach 829 billion Euros, which is 5.8% of the EU GDP (European Commission, 2020). This is possible thanks to recent advances in digital technologies that enable massive data generation via sensors and smart devices. With such an enormous amount and potential, data brings new opportunities for, among others, improving decision-making, streamlining business processes, developing new products and services, and addressing societal problems (Günther et al., 2017; Hartmann et al., 2016; Sorescu, 2017; Susha et al., 2020; Wixom & Ross, 2017; Zott & Amit, 2017).

Data is even more valuable when shared with other parties (Elsaify & Hasan, 2021), as multiple data sources often need to be combined to generate meaningful and actionable insights (Koutroumpis et al., 2020; Susha et al., 2020; van den Broek & van Veenstra, 2018; Virkar et al., 2019). In this regard, there is an emergence of data marketplaces—a platform facilitating data sharing and monetization by businesses and individuals (Abbas et al., 2021; Koutroumpis et al., 2020; M. Spiekermann, 2019). Typically, data marketplaces rely on a Trusted Third Party (TTP), an intermediary that collects, stores, and aggregates different data sources (Susha et al., 2020; Richter & Slowinski, 2019). A TTP also distributes the aggregated datasets through a matchmaking process between parties interested in sharing and acquiring datasets (Susha et al., 2020; Richter & Slowinski, 2019). Thus, by sharing through data marketplaces that use a TTP, businesses can capture new opportunities to create value from data through personalization, automation, knowledge creation, and a new offering of products and services (Cichy et al., 2021; Schomakers et al., 2020). For instance, agricultural firms could share their data generated by smart farming devices

for benchmarking to improve cultivation practices (De Prieëlle et al., 2020). Individuals can also benefit from data marketplaces by directly selling their personal data for monetary compensation (Bataineh et al., 2020). This way, individuals would not miss out on a large sum of financial benefits often only enjoyed by big technology companies (Bataineh et al., 2020; Kirkpatrick, 2021). In short, data marketplaces could play a crucial role in establishing a single functioning market for data that allows it to flow freely across sectors for the greater good and for leveling the playing field of the data economy (European Commission, 2020).

However, businesses and individual consumers are often concerned and reluctant to share data, let alone through data marketplaces that use a TTP (Jernigan et al., 2016; Richter & Slowinski, 2019). Businesses often fear losing control over sensitive data that could benefit competitors (Jarman et al., 2016; Klein & Verhulst, 2017; Zrenner et al., 2019), making them distrust each other (Arnaut et al., 2018; Dahlberg & Nokkala, 2019; Kembro et al., 2017). Businesses are also concerned that sharing data could violate privacy regulations (Eurich et al., 2010; Khurana et al., 2011; Sayogo et al., 2014). Similarly, consumers feel they have lost control over their data and often have no idea how companies handle it (S. Spiekermann & Novotny, 2015), making them increasingly concerned about their privacy (Cichy et al., 2021; van Schaik et al., 2018). Consumers even fear possible data misuse by companies and unauthorized third parties, making them lose trust in business activities involving sharing personal data (Pal et al., 2021). These concerns and challenges might result in failure to tap the potential of the collected data, making them unused and ultimately leading to missed opportunities for society (Roman et al., 2021). Furthermore, the lack of a functioning market for data would cause most companies to struggle to compete in the digital world against big technology companies, which control substantial masses of data.

Against this backdrop, there is a need to rethink how data sharing should be conducted in a digital society, particularly to reduce dependence on a TTP in data sharing. In this regard, various privacy-enhancing technologies (PETs) have been

developed, which generally refer to technical means to protect sensitive data through, for instance, anonymization, encryption, and secure computation, without losing its functionality (Borking & Raab, 2001; Burkert, 1997; Heurix et al., 2015). The rapid emergence of PETs is timely, given the increasing need to balance the tension between protecting privacy, maintaining control, and utilizing data for value creation (Gast et al., 2019; Schwinghammer et al., 2022; Zöll et al., 2021). Thus, PETs could be valuable in giving back control over data and addressing privacy concerns for businesses and individuals while sharing through data marketplaces (Acquisti et al., 2016).

Our focus in this dissertation is one class of PETs called Multi-Party Computation (MPC), which has been present for some time (Yao, 1982) but has been experiencing rapid growth and commercialization recently thanks to advances in computing resources (Gartner, 2021). It is a cryptographic technique where multiple parties (individuals or organizations) perform a joint computation without disclosing the input data between any involved parties (Bestavros et al., 2017; Choi & Butler, 2019; Yao, 1986; C. Zhao et al., 2019). MPC allows computational analysis to generate meaningful insights while keeping the input data private (Balson & Dixon, 2020). It works by splitting input data into multiple parts, which are then distributed among participating parties. Subsequently, each party combines parts of the input data from all participating parties and then computes the results based on the agreed query. This way, the requesting parties will only receive the results generated from the computational analysis without learning anything about the input data from participating parties (Bestavros et al., 2017; Choi & Butler, 2019; Yao, 1986; C. Zhao et al., 2019). As such, MPC could reduce the role of a TTP in data sharing through data marketplaces, from collecting, storing, and processing data as well as matching between data providers and data buyers to matchmaking only (Bruun et al., 2020; Helminger & Rechberger, 2022). This change in the role of a TTP is possible thanks to the capability of performing distributed computation with MPC without revealing the

input data to a TTP. In this regard, the role of data storage, analysis, and processing is not relevant anymore for a TTP. Nevertheless, a TTP is still needed in the case of data marketplaces, particularly in facilitating matchmaking between data providers and buyers.

We illustrate the difference between sharing with a TTP and MPC in Figure 1. In the end, MPC could enable businesses and individuals to retain control of their data and reduce data sharing risks by safeguarding data privacy and confidentiality (C. Zhao et al., 2019).

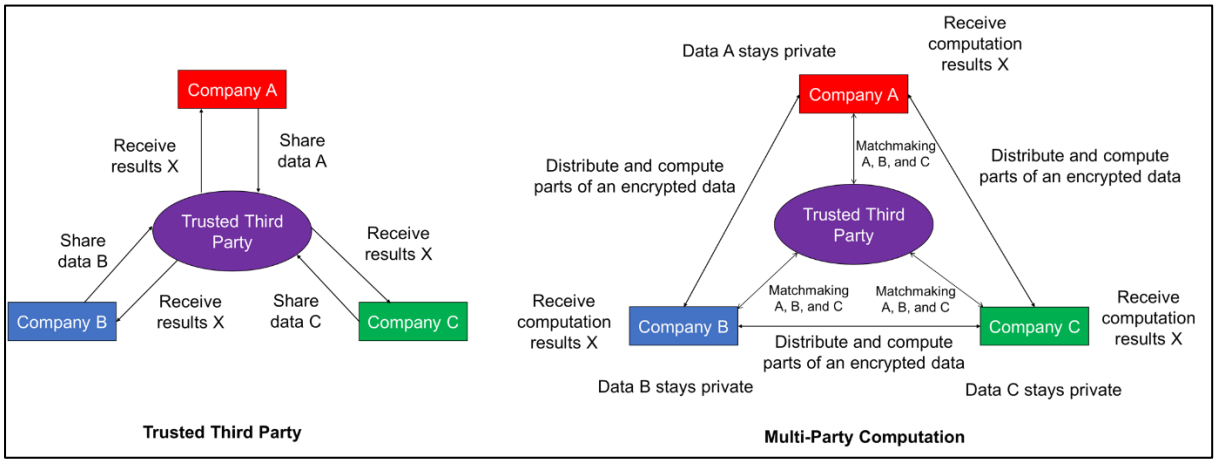


Figure 1 Comparison between TTP and MPC (adapted from Noble et al. (2019))

MPC might technically help change how businesses and consumers share data in a digitized society (Zare-Garizy et al., 2018). However, it is unclear whether MPC will also increase the willingness of businesses and consumers to share data. Also, the emerging nature of MPC might result in unexpected side effects not occurring in existing data sharing approaches based on a TTP, which might cancel out the positive impact of MPC. Hence, it is necessary to understand whether MPC would lead to a greater willingness to share data by investigating the potential impact of MPC use in data marketplaces on data sharing antecedents by businesses and consumers.

Based on our analysis, in this dissertation, we address the following practical problem:

Digitalization enables a massive generation of data that can be used for value creation, especially when shared with other parties. The emergence of data marketplaces, which rely on a Trusted Third Party (TTP), could play a key role in facilitating data sharing to generate meaningful insights and enable new opportunities for businesses and individuals. However, many companies refrain from data sharing due to, among others, fear of losing control over data, while consumers are also concerned that their privacy is compromised. These dynamics create urgency regarding how data sharing should be approached, particularly balancing data utilization, privacy, and control while reducing dependency on a TTP. Advances in privacy-enhancing technologies like Multi-Party Computation (MPC) might change how companies and individuals share data by enabling joint computation to generate insights without disclosing the input data. However, despite the promises that MPC offers in addressing data sharing barriers, it is unclear whether businesses and individuals would be more willing to share data. Thus, we must identify the potential impact of MPC use in data marketplaces on antecedents of data sharing decisions by businesses and individual consumers.

1.2. Scientific problem

Previous research on inter-organizational information sharing (e.g., Morrell & Ezingeard, 2002; Robey et al., 2008) primarily focuses on sharing data between two partners in a clear usage context, with clear arrangements on which companies are involved and how the data is being utilized (Elgarah et al., 2005; Narayanan et al., 2009; Praditya et al., 2017). Scholars working on this topic have identified various data sharing antecedents by businesses, such as control of data ownership (De Prieëlle et al., 2020; Priego et al., 2019; Sayogo et al., 2014), privacy (Jarman et al., 2016; Sun et al., 2018), trust (Asare et al., 2016; Pavlou & Gefen, 2004), and risk (Dahlberg & Nokkala, 2019; Johnson, 2009; White et al., 2007). Meanwhile, within the context of consumer data sharing, previous research in the information privacy literature (e.g., Popovič et al., 2017; Smith et al., 2011) mainly investigates consumer data disclosure in emerging phenomena like social media (Alashoor et al., 2017; Hajli & Lin, 2016) and

IoT services (Bélanger et al., 2021; Cichy et al., 2021). Antecedents such as perceived control (Dinev et al., 2013; Krasnova et al., 2010), privacy concerns (Cichy et al., 2021; Malhotra et al., 2004; Schomakers et al., 2020), perceived risks (Dinev & Hart, 2006; Xu et al., 2011), and trust (Ažderska, 2012; Kehr et al., 2015; Pavlou, 2003) were found to influence consumers' data sharing decisions.

However, both literature streams have yet to consider the emerging phenomenon of MPC, particularly when used in data marketplaces. MPC fundamentally differs from the existing data sharing approach that primarily relies on a TTP as an intermediary. MPC implies a distributed computation paradigm where the input data need not be shared and can be analyzed on the premise of data owners, which is also referred to as in-situ data rights (Van Alstyne et al., 2021). Because the computation is performed in a distributed manner, a TTP that acts as an intermediary that facilitates data sharing is no longer needed (Bruun et al., 2020; Helminger & Rechberger, 2022). MPC also emphasizes the importance of computation instead of the underlying details of the data (Van Alstyne & Lenart, 2020). This is because MPC only computes and shares the analysis results (data insights) with the requesting party while keeping the input data private (Bestavros et al., 2017; Choi & Butler, 2019; Elliot & Quest, 2020; C. Zhao et al., 2019). Hence, these radically new ways of sharing data with MPC could increase the willingness of businesses and consumers to share data, as it could establish new ways to retain control, enhance privacy, reduce risks, and eliminate the need for trust in other parties. At the same time, the emerging nature of MPC might create new risks and issues unknown in existing data sharing approaches, as it can be utilized unexpectedly beyond the initial purpose. Thus, these differences challenge the relevance of existing data sharing antecedents in the new context of MPC in data marketplaces.

Understanding how MPC challenges the relevance of data sharing antecedents in data marketplaces is even more important because of a scientific knowledge gap in MPC and data marketplaces literature. Despite its potential to enable a new data sharing

approach, most discussion on MPC focuses on technical aspects rather than business and societal implications (e.g., Guo et al., 2020; Volgushev et al., 2019). While academics are beginning to realize that MPC can create public value and address societal problems (e.g., Bestavros et al., 2017; Lapets et al., 2018), only a few studies consider how the technology can be used in a TTP like data marketplaces (e.g., Koch et al., 2021; Roman & Vu, 2019). Meanwhile, data marketplace studies are still dominated by pricing mechanisms (e.g., Balazinska et al., 2013; Fricker & Maksimov, 2017; Muschalle et al., 2013) and architectural design (e.g., Brandão et al., 2019; Mišura & Žagar, 2016; Özyilmaz et al., 2018; Ramachandran et al., 2018). Socio-technical aspects of data marketplaces, such as data governance, data ecosystems, and user studies to understand social implications, are yet to become the focus of academics researching this domain (Abbas et al., 2021). As such, we need to understand what MPC means for data sharing decisions by investigating if and how MPC impacts businesses' and consumers' antecedents of data sharing in data marketplaces. This way, we can theorize the potential impact of MPC use in data marketplaces on data sharing decisions by businesses and consumers.

Based on the elaboration above, we formulate the following scientific problem:

Existing literature on data and information sharing has studied various antecedents of businesses' and consumers' data sharing decisions. However, prior research on this topic is mainly based on the assumption that a Trusted Third Party (TTP) facilitates data exchange. As an emerging phenomenon, MPC is a different data sharing approach compared to sharing through a TTP such as data marketplaces, in which (1) the data stays with the owner, (2) the computation process is distributed without an intermediary, and (3) only the results are shared. These differences challenge the current understanding of why and how businesses and consumers share data. Establishing a new understanding is crucial because research on MPC and data marketplaces mainly emphasizes technical aspects while neglecting the business and societal impact. As such, it is necessary to theorize the potential impact of MPC use in

data marketplaces on data sharing decisions by businesses and consumers to enrich our understanding of the business and societal implications of MPC.

1.3. Research objective and research questions

Based on the practical and scientific problems identified in the previous sections, we formulate our research objective as follows:

Our research objective is to theorize the socio-technical implications of MPC on data sharing through data marketplaces, by investigating data sharing antecedents that are potentially impacted by MPC and the resulting impact of MPC on data sharing decisions by businesses and individuals.

From the objective above, we propose three research questions.

Research question 1: what types of data sharing antecedents could be impacted by MPC use in data marketplaces?

The first research question explores possible data sharing antecedents that could be impacted by MPC use in data marketplaces. This step is essential as there are various antecedents of data sharing decisions, but not all of them could be impacted by MPC use in data marketplaces. Thus, answers to the first research question help us to derive a set of data sharing antecedents of data sharing decisions that we should focus on in subsequent studies regarding business and consumer perspectives on MPC. We describe the study design, procedures, and key findings of the business model analysis in Chapter 3.

Research question 2: what could be the impact of MPC use in data marketplaces on antecedents of data sharing by businesses?

After understanding the types of data sharing antecedents that MPC could impact, we need to investigate what kind of impact MPC possibly creates on those antecedents.

Therefore, the second research question focuses on understanding the link between MPC use in data marketplaces and antecedents of data sharing decisions from the perspective of business actors. Specifically, answers to this question deliver a conceptual framework explaining the potential impact of MPC on antecedents of data sharing decisions in data marketplaces, as elicited from the first research question. In this way, we can establish an understanding of changes resulting from implementing MPC in the context of business data sharing in data marketplaces. We describe the study design, procedures, and key findings of the qualitative study in Chapter 4.

Research question 3: what could be the impact of MPC use in data marketplaces on antecedents of data sharing by consumers?

In the last research question, we turn our attention to data sharing decisions of individual consumers. In particular, we used the same antecedents identified in the first research question to quantitatively investigate the impact of MPC use in data marketplaces on consumers' decisions to share personal data. Answers to the third research question are beneficial to generate a comprehensive understanding of the potential impact of MPC on data sharing decisions, which covers both business and consumer perspectives. We describe the detail of the study design and results in Chapter 5.

1.4. Research approach

We follow a mixed-method research design as the main research method to answer our research questions and ultimately fulfill our research objectives (Venkatesh et al., 2013). It is defined as "an approach that combines quantitative and qualitative research methods in the same research inquiry" (Venkatesh et al., 2013). Those two methods can be used to either inform each other (i.e., sequential) or be independent of each other (i.e., concurrent) (Recker, 2021). Hence, mixed-method research design

differs from the multi-method approach (Mingers, 2001, 2003), which uses multiple research methods within a single worldview, either qualitative or quantitative only.

This method is suitable for our research given the emerging nature of MPC and the lack of prior knowledge on what this technology means for data sharing decisions. Specifically, we need to explore the types of data sharing antecedents impacted by MPC use in data marketplaces and the underlying mechanisms of the impact that MPC creates. Hence, we need diverse views from different key actors (i.e., MPC developers, businesses, and consumers) to establish a complete understanding of the implications of MPC. In this regard, this research problem can only be addressed with multiple data sources from both qualitative and quantitative research methods, which justifies the decision to employ mixed-method research (Creswell & Clark, 2017).

Mixed-method research is also valuable for our research because it can complement the strength of both qualitative and quantitative research methods while mitigating their weaknesses (Creswell & Clark, 2017; Recker, 2021; Venkatesh et al., 2013). On the one hand, the qualitative method enables us to understand causality in more detail while maintaining the possibility of discovering alternative explanations. In our research, this strength can be translated into understanding why and how MPC could impact data sharing antecedents by businesses. On the other hand, the quantitative method allows us to identify the correlation in our findings, like whether MPC uses in data marketplaces correlate with greater consumer willingness to share data. Therefore, integrating findings from both methods can strengthen inferences while providing rich, diverse, and comprehensive views of the phenomena under study, which is the impact of MPC on data sharing decisions in data marketplaces (Creswell & Clark, 2017; Recker, 2021; Venkatesh et al., 2013).

The decision to employ mixed-method research should be driven by specific purposes based on the research problem. Venkatesh et al. (2013) classify seven purposes of mixed-method research: (1) *compensating* the limitation of one approach; (2)

obtaining *complementary* views; (3) capturing a *complete* overview of a phenomenon; (4) *corroborating* findings from one approach by using the other; (5) using findings from one approach for *developing* research questions of the next approach (6) generating *diversity* of perspectives about a phenomenon; (7) *expansion* of understanding derived from one approach using another. Our research follows a combination of developmental and diversity rationale (Venkatesh et al., 2013). In particular, findings from the business model analysis serve as a basis for the qualitative study (semi-structured interviews) and the quantitative study (experiment). Meanwhile, all three studies focus on different perspectives, from MPC developers (business model analysis), businesses (qualitative study), and individual consumers (quantitative study). In this regard, findings from all studies offer diverse views on how MPC changes perceptions around data sharing decisions by businesses and consumers. We also follow the convergent parallel mixed methods research (Creswell & Creswell, 2017) by integrating findings from all three studies to fulfill the overarching research objective.

We specify our research context in the data marketplaces domain, where businesses and individual consumers can proactively share and sell their data for monetary compensation (Abbas et al., 2021; Schomakers et al., 2020). This context represents a unique setting with a high fear of losing control over data, low trust among actors involved, and high risks of data sharing, ultimately leading to limited adoption due to unwillingness to share (Koutroumpis et al., 2020; M. Spiekermann, 2019). We specifically focus on data marketplaces in the automotive industry, where key actors like original equipment manufacturers (OEMs), car insurance companies, and mobility service providers are very conventional and secretive when dealing with their car data (Docherty et al., 2018; Kerber, 2018). The sensitive nature of driving data also creates mounting concerns for consumers' privacy as they also perceive the lack of control over data (Docherty et al., 2018). Nevertheless, the growing trend of digitalization underlines the importance of data sharing in the automotive sector to develop novel

products and services, such as connected cars, usage-based insurance, and shared mobility (Athanasopoulou et al., 2019; Kaiser et al., 2021). Hence, it is relevant to investigate the impact of MPC on willingness to share within this context of data marketplaces in the automotive industry.

Our research design comprises four phases. In Phase 1, we review the core concept and current landscape of MPC (as our focal technology) and data marketplaces in the automotive industry (as our research domain). By starting with a literature study, we obtain a comprehensive overview of existing knowledge and research gaps, particularly on the possible impact of MPC on sharing through data marketplaces, which is currently lacking in the literature. In this regard, a literature study enables us to focus our effort on bridging existing gaps and embedding them into the scientific and practical relevance of the field.

In Phase 2, we identify data sharing antecedents that could be impacted by MPC use in data marketplaces. We use insights obtained in Phase 1 to conduct the first empirical study through exploratory interviews with MPC experts and practitioners. We opt for the qualitative approach in this phase as research efforts on MPC and data marketplaces so far focus on technological aspects while neglecting their businesses and societal implications. In doing so, we draw upon the Unified Business Model framework by Al-Debei & Avison (2010), which enables us to explore the business model implications of MPC in data marketplaces and identify which data sharing antecedents are impacted by MPC use in data marketplaces. From this phase, we derive a set of data sharing antecedents that could be impacted by MPC use in data marketplaces, which serves as a foundation for our further empirical research (Phase 3 and 4). We detail the research approach for this study in Section 3.2 (Chapter 3).

After identifying data sharing antecedents that could be impacted by MPC use in data marketplaces, we proceed with two follow-up studies in Phase 3 and 4 to determine the potential impact of MPC on these data sharing antecedents, focusing on two

perspectives, namely business and consumer data sharing. In Phase 3, we conduct semi-structured interviews with experts and practitioners in the automotive industry to gather a rich array of perspectives regarding how MPC changes businesses' decisions to share through data marketplaces. A qualitative study is suitable as prior studies lack insights into the business and societal implications of MPC, making our attempt to understand the impact of MPC exploratory in nature.

In doing the qualitative study, we look into the literature on business-to-business data sharing and develop a theoretical overview of data sharing antecedents based on our findings in Phase 2. We then construct theoretical propositions on how MPC qualitatively impacts data sharing antecedents in business data sharing through data marketplaces. We use these propositions as the foundation of our qualitative study, including the development of an interview protocol that can capture the potential impact of MPC on antecedents of data sharing decisions by businesses in data marketplaces. While findings in Phase 2 guide the data collection process, we also keep an open mind for new insights not identified before, allowing us to get a complete picture from multiple perspectives. These differences are possible as MPC experts might have a more positive view of the technology, while business actors might not completely agree with it. We then synthesize the qualitative interview data to obtain insights and refine the initial propositions. These propositions reflect how MPC qualitatively changes the dynamics of data sharing decisions by firms in data marketplaces. We detail the research approach for this study in Section 4.3 (Chapter 4).

In Phase 4, we conduct an experimental study to investigate the potential impact of MPC on data sharing antecedents by consumers in data marketplaces. This approach is suitable for our research since we aimed to explore the effect of a new instrument (i.e., MPC) in changing the current situation (i.e., consumers' data sharing decisions in data marketplaces). Hence, we have to ensure a high internal validity to achieve this

goal. In order to attain this high internal validity, we opted to use an experimental research approach (Verschuren & Doorewaard, 2010).

We conduct the experiment online using the crowdsourcing platform Prolific, which enables us to improve generalizability and internal validity, as the online settings provide access to a large sample size (N=1457). We create the experimental setup in which participants act as data providers and experience the data sharing process through a mock-up of MPC-enabled personal data marketplaces in the automotive domain. We randomly assign participants into one of the three groups (Trusted Third Party, MPC, and a fictitious technology called Data-Computation-Protection/DCP) to get the approximately same size for each group. We follow a post-test-only between-subject design (Campbell & Stanley, 2015), meaning that we observe each group once using a questionnaire after the treatment (i.e., post-test). We analyze the data by performing the analysis of variance (ANOVA) to compare the means of different groups. In this way, we can investigate whether MPC performs better than the existing condition (TTP) and the fictitious technology in impacting antecedents of data sharing decisions by consumers in data marketplaces. We detail the research approach for this study in Section 5.3 (Chapter 5).

Finally, in Phase 5, we synthesize findings from all studies to theorize the potential impact of MPC on data sharing decisions in data marketplaces from multiple perspectives. We present the main findings and implications for theory and practice in Chapter 6.

1.5. Contributions and relevance

This study delivers a conceptual framework that theorizes the potential impact of MPC on decisions to share data by businesses and consumers in the context of data marketplaces. Theorizing is a process of explaining, synthesizing, and generalizing empirical research findings, which results in deliverables like lists of variables,

propositions, hypotheses, and frameworks that establish a foundation and provide a direction for a particular research phenomenon (Weick, 1995). In our study, theorizing involves (1) synthesizing data sharing antecedents that could be impacted by MPC use in data marketplaces, (2) explaining the potential impact of MPC on those antecedents from the perspective of businesses and consumers, and (3) generalization of our findings from the three empirical studies to understand the potential impact of MPC on data sharing decisions by businesses and consumers. As such, our study does not claim to develop a theory but rather a foundation and possible direction on how MPC challenges the current understanding of why businesses and companies share data. This is important because, so far, antecedents of data sharing are based on the presence of a trusted third party as an intermediary. MPC differs in that data can be shared differently without relying on an intermediary through joint computation while maintaining that the data stays with data providers and is kept private. Thus, this research forms a basis for understanding the impact of MPC on data sharing decisions by businesses and consumers in the context of data marketplaces.

This study also contributes to the MPC literature. Research on MPC is dominated by attempts to improve its efficiency and scalability, which is understandable given the novelty of the technology. Nevertheless, we argue that understanding the socio-technical elements of MPC, such as business models, governance, and user evaluations, is equally essential to ensure its widespread adoption. In this regard, our research aims to close this gap by linking MPC with antecedents of data sharing decisions in the context of data marketplaces. This way, we theorize the potential impact of MPC on data sharing decisions as a fundamental basis for further user research on MPC.

Moreover, this study contributes to the data marketplaces literature, especially regarding its adoption challenges. Currently, businesses and consumers are still reluctant to participate in sharing through data marketplaces due to the lack of trust,

privacy concerns, and fear of losing control over data that might benefit competitors. Thus, through this research, we can investigate what MPC can offer when used in data marketplaces. Ultimately, this research provides important insights regarding how MPC can address the adoption challenges of data marketplaces.

Beyond theoretical contributions, our research also informs practitioners working in the domain of MPC and data marketplaces. Given the novelty of MPC, an understanding of its business value and implications for companies' business models are required to boost MPC adoption in the market. This research provides insights to MPC developers and service providers regarding the business value of MPC and how it should be promoted to prospective users. This research also informs data marketplace operators on incorporating MPC into data marketplaces and its possible implications for their business models. Further, our findings generally benefit businesses interested in moving towards privacy-friendly business models by showcasing how MPC can fulfill consumers' demand regarding greater privacy protection and control while still creating value from data.

1.6. Dissertation outline

This dissertation is structured as follows. Chapter 2 explores MPC as our technology in focus and data marketplaces as our research domain. We provide an overview of MPC definitions, building blocks, architecture, security requirements, use cases, and adoption challenges. Subsequently, we present the current landscape of data marketplaces, including its conceptual definitions, key roles, and open challenges.

Chapter 3 answers the first research question by exploring the business model implications of MPC in data marketplaces. We start by describing our framework for business model analysis, followed by explaining procedures for data collection and analysis. After that, we present our findings by elaborating on the value proposition, value architecture, and value finance of MPC in data marketplaces. We conclude this

chapter by presenting a set of data sharing antecedents impacted by MPC that will be the focus of subsequent studies.

Chapter 4 answers the second research question by discussing the business perspective on MPC and data sharing. This chapter begins by discussing data sharing antecedents identified in Chapter 3 within the context of business-to-business data sharing and specifies them to MPC as the focal technology. This resulted in initial propositions for further investigation in the qualitative study. Then, we describe our strategy to collect and analyze qualitative data through interviews with key actors within the automotive industry. We also explain how we present the complex concept of MPC and data marketplaces to interviewees during data collection. Subsequently, we present and discuss our interview findings by comparing statements, identifying patterns, and understanding differences. We conclude the chapter by refining further the propositions explaining the impact of MPC on data sharing decisions by businesses. We also outline limitations we encounter during the execution of the study.

Chapter 5 discusses the consumer perspective on MPC and data sharing. First, we focus on antecedents identified in Chapter 3 in reviewing existing literature on consumers' data sharing decisions based on information privacy theory. After that, we relate the antecedents with MPC and develop research hypotheses. Subsequently, we describe the experimental setup, including the mock-ups we used to compare three data sharing scenarios (trusted third party, MPC, and fictitious technology). We then discuss the measures, procedures, demographic profile of participants, and the analysis results. We conclude the chapter by discussing the findings compared to the literature and reflecting on the limitations.

Finally, Chapter 6 brings together all studies and reflects on the results. First, we present an overview of key findings to answer the research questions. Then, we discuss similarities and differences between the findings of the qualitative and

quantitative studies. After that, we elaborate on the theoretical and practical contributions of this study. We wrap up this chapter by outlining limitations and recommendations for future research.

We summarize the connection between each research phase, research objective, research questions, approach, and dissertation chapter in Figure 2.

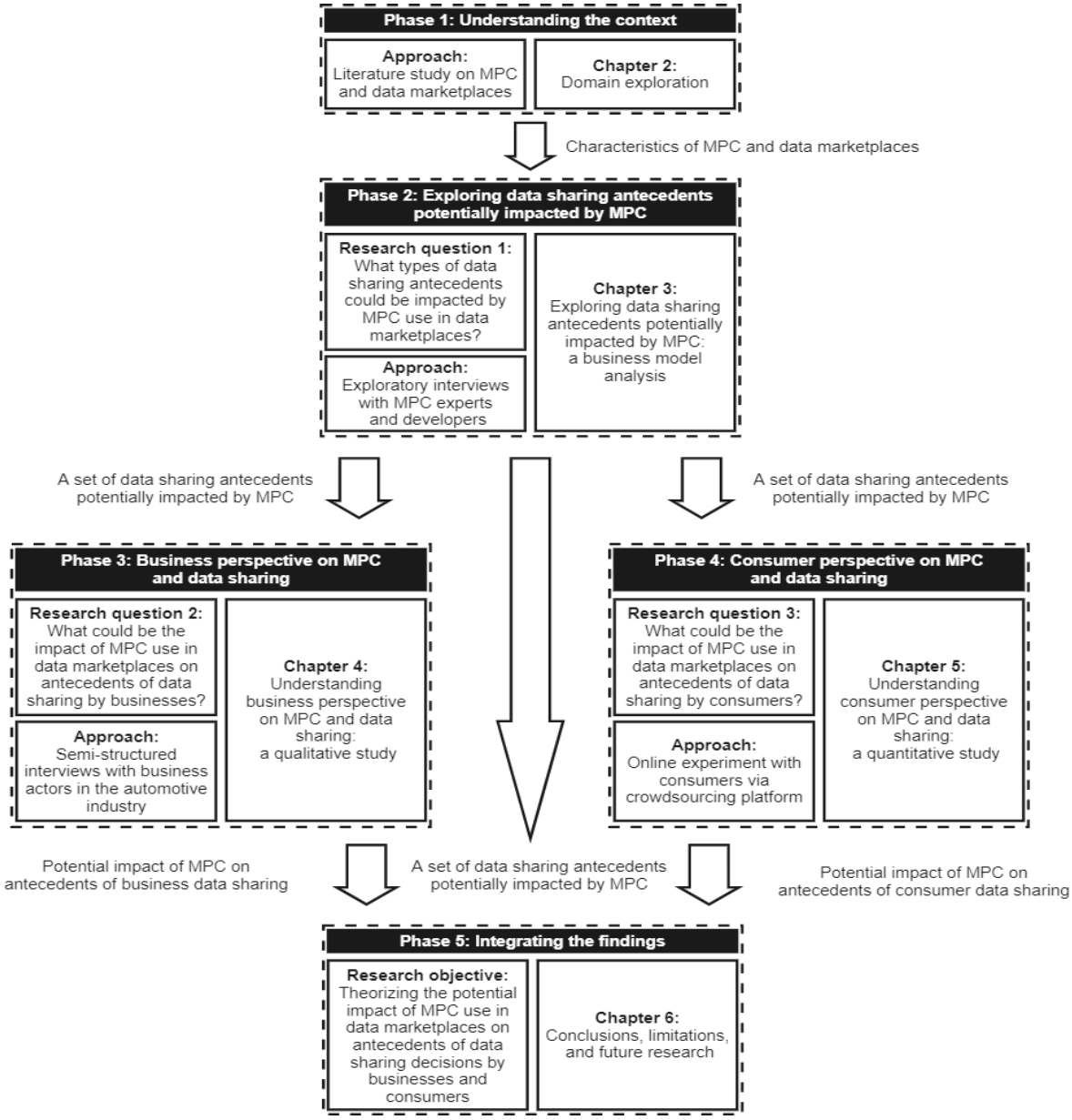


Figure 2 Dissertation outline and the connection with research phases, objective, questions, and approach

2 Domain exploration

In this chapter, we describe MPC and data marketplaces and delineate them within the scope of this work. This chapter is an essential piece in characterizing how MPC, as a novel technology, changes the existing data sharing situation through data marketplaces. To do so, we conduct a literature review on MPC and data marketplaces, which covers scientific articles, white papers, and grey literature. We first describe the definition of MPC, followed by the technical foundations of the technology and an illustration of how MPC works. We then elaborate on MPC's security requirements and robustness towards adversaries, followed by architecture configurations and examples of MPC use cases. We conclude our analysis of existing literature on MPC by outlining its adoption challenges.

Meanwhile, for data marketplaces, we start by providing the definition and comparison to similar terms, followed by the current state of data marketplaces in the automotive industry. Then, we describe key roles in automotive data marketplaces and summarize open challenges faced by data marketplaces. Finally, we conclude the chapter by discussing how MPC can address data sharing challenges in data marketplaces and fulfilling the need to rethink how data sharing should be conducted. This way, we can focus our investigation on data sharing antecedents that could be impacted by MPC use in data marketplaces.

2.1. Multi-Party Computation

2.1.1. Definitions

MPC is a cryptographic technique where two or more parties perform a joint computation that results in a meaningful output without disclosing the input provided by either party (Bestavros et al., 2017; Choi & Butler, 2019; C. Zhao et al., 2019). MPC enables computation on encrypted data since all parties only receive the output of a function while keeping the input data private (Archer et al., 2018; Choi & Butler, 2019).

MPC is particularly useful in a distributed computing scenario where multiple parties (organizations or individuals) would like to cooperate by computing a function together and obtaining more valuable information without leaking their confidential data (C. Zhao et al., 2019). Traditionally, this scenario can be performed using a Trusted Third Party (TTP) as an intermediary (see Section 1.1). However, this scenario is vulnerable because all data is stored and processed in one single point, creating an additional threat if the TTP turns out to be malicious itself or is compromised by a malicious entity. In this regard, MPC can be seen as a means to eliminate TTP and replace it with a cryptographic technique (Sousa et al., 2018). In this way, multiple actors can collaboratively analyze data to obtain meaningful insights without the need to share the underlying input data and establish mutual trust (Alter et al., 2018; Bestavros et al., 2017).

Nevertheless, Alter et al. (2018) and Lindell (2020) pointed out that people often misunderstand two things about this technology. First, MPC cannot detect and prevent adversaries from providing incorrect input data, meaning that any input data are allowed in MPC, even fictitious ones. To give an example, if two people (A and B) would like to use MPC to know who has the highest salary (without revealing their salary information), then person A could provide the highest possible number as input data, while person B could provide some random number that might be lower. Hence, MPC would compute two sets of incorrect input data and show that person A is the highest earner even though the number is not true. Given that this issue cannot be addressed as part of MPC algorithms, additional mechanisms are needed to force all participating parties to ensure data quality and correctness before participating in the computation, although it may incur high costs.

Second, despite the ability of MPC to secure the computation process, the output resulting from the computation may still reveal sensitive information. Returning to the same example of persons A and B, they can use MPC to compute their average salaries without disclosing their individual salaries. While nothing but the average will

be revealed, person B can still calculate the salary of person A using the MPC output and his/her salary. In this regard, all participants should also agree on what functions can and cannot be performed to prevent revealing sensitive information from the computation results. Alternatively, MPC can be complemented by other technical solutions like differential privacy, which protects data by adding random noise to allow disclosing results without giving away sensitive information (Dwork, 2006; Dwork & Roth, 2014).

2.1.2. Building blocks: garbled circuits and secret sharing

The theoretical foundation of MPC was first developed by Yao (1982), in which he developed a two-party secure computation protocol. This protocol was developed within a scenario called "the millionaire's problem," in which two millionaires would like to know which is richer without revealing their net worth to each other or to a TTP (Yao, 1986). Building from this two-party computation protocol, Goldreich et al. (1986) developed the protocol further to make it suitable for computation involving multiple parties. An important building block in both works is *garbled circuits* (Beaver et al., 1990), which leverage (a combination of) basic logic gates like AND, XOR, and OR to construct any functions. Garbling is described as an obfuscation process of the Boolean gate truth table (Yakoubov, 2017). Garbled circuits mainly rely on oblivious transfer (Rabin, 1981), in which one party (the circuit generator) produces two secret inputs, and the other party (the circuit evaluator) can select one of them. However, the circuit generator will not be able to learn the secret input selected by the circuit evaluator (Choi & Butler, 2019). See Table 1 for an illustration of the garbling of an OR gate.

(a) An OR gate *g*			(b) An OR gate with garbled keys			(c) A garbled OR gate G(*g*)
Input wire w_1 (*u*)	**Input wire w_2 (*v*)**	**Output wire w_3 (*u* OR *v*)**	**Input wire w_1 (*u*)**	**Input wire w_2 (*v*)**	**Output wire w_3 (*u* OR *v*)**	**Ciphertexts in a garbled OR gate (with random permutation)**
0	0	0	k_1^0	k_2^0	k_3^0	$\mathrm{Enc}_{k_1^0}(\mathrm{Enc}_{k_2^0}(k_3^0))$
0	1	1	k_1^0	k_2^1	k_3^1	$\mathrm{Enc}_{k_1^0}(\mathrm{Enc}_{k_2^1}(k_3^1))$
1	0	1	k_1^1	k_2^0	k_3^1	$\mathrm{Enc}_{k_1^1}(\mathrm{Enc}_{k_2^0}(k_3^1))$
1	1	1	k_1^1	k_2^1	k_3^1	$\mathrm{Enc}_{k_1^1}(\mathrm{Enc}_{k_2^1}(k_3^1))$

Table 1 Illustration of the garbling of an OR gate, adapted from Zhao et al. (2019)

Another key building block of MPC is the *secret sharing* technique (Shamir, 1979). In this approach, each party splits its input data into multiple encoded parts called secret shares, which are then computed and recombined to generate the final output. In this way, no information other than the computation results will be revealed to all parties, while input data will stay private (Alter et al., 2018). The secret sharing technique uses a threshold scheme, allowing only a certain number of participants to perform the computation. This scheme provides added control over what kind of participants are qualified to participate in the computation (Choi & Butler, 2019). Furthermore, compared to garbled circuits, secret sharing is more efficient and allows more parties to participate in the computation (Pedersen et al., 2007). Because of this advantage, most of the real-world MPC implementations make use of the secret sharing technique.

2.1.3. Illustration

To illustrate the generic MPC process with the secret sharing technique, we developed a hypothetical scenario in the automotive industry (see Figure 3). This scenario was developed based on prior studies (Bestavros et al., 2017; Bogdanov et al., 2015;

Bogetoft et al., 2009; Roseman Labs, 2022). In this scenario, three car manufacturers would like to calculate their average yearly sales figures without disclosing their own numbers. For simplicity, we assume that in 2022, company A sold 48.000 cars, while companies B and C sold 60.000 and 54.000 cars, respectively. One option would be to share those sales numbers with a TTP that will calculate the average yearly sales. However, sales numbers are considered sensitive data, making car manufacturers reluctant to share this data due to concerns about losing competitive advantages.

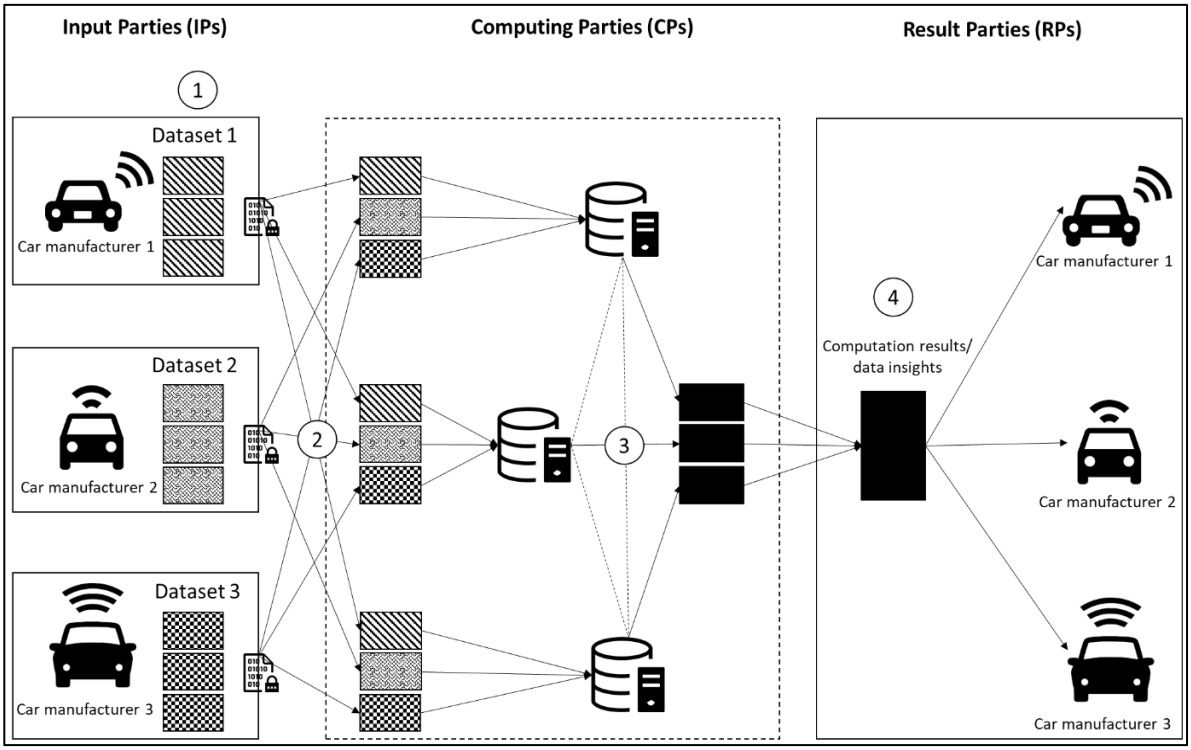


Figure 3 An illustration of the MPC process, adapted from Bestavros et al. (2017); Bogetoft et al. (2009); Bogdanov et al. (2015); and Roseman Labs (2022)

MPC could tackle those concerns by employing cryptographic techniques to replace a TTP and perform privacy-enhancing data sharing. In implementing MPC systems, three main roles should be present: input parties (IPs), computing parties (CPs), and result parties (RPs) (Alter et al., 2018; Archer et al., 2018; Lapets et al., 2019). *Input parties (IPs)* are individuals/organizations that provide sensitive data as input that will

be computed with MPC. Meanwhile, *computing parties (CPs)* are entities that deploy and maintain computing resources to compute input data with MPC and generate computation results. Further, *result parties (RPs)* are entities that request and obtain computation results with MPC, which can be IPs or completely different individuals/organizations.

The MPC process based on the secret sharing technique comprises four steps. First (Step 1 in Figure 3), as IPs, each car manufacturer locally encrypts and splits their yearly sales data into three parts (or three secret shares). When combined, these secret shares resulted in the original yearly sales data, but when looked at individually, each share is a meaningless number that reveals nothing about the sales number. See Table 2 below for the numerical illustration of splitting yearly sales data into three secret shares.

Car manufacturer	**Secret share 1**	**Secret share 2**	**Secret share 3**	**Yearly sales (in thousands)**
Car manufacturer A	13	57	-22	48
Car manufacturer B	35	18	7	60
Car manufacturer C	-77	23	108	54

Table 2 The secret sharing process

Subsequently, these secret shares are distributed to multiple CPs managed by multiple partners and completely independent of each other (Archer et al., 2018). Together, these CPs compute the encrypted and partitioned data according to the requested function (Step 2 in Figure 3). Each CP calculates parts of data received from different IPs to form partial results that reveal nothing about the input data (Step 3 in Figure 3). We numerically illustrate the calculation in the secret sharing process in Table 3.

CP 1	CP 2	CP 3
13	57	-22
35	18	7
-77	23	108
Total: -29	**Total: 98**	**Total: 93**

Table 3 The calculation of the secret sharing process

Finally, those partial results are recombined to form average yearly sales data from those three car manufacturers as agreed prior to the computation (Step 4 in Figure 3). By adding the results from the three CPs and dividing them by three (number of CP), we can calculate the average yearly sales of the three companies, which equals 54.000 cars. The results will be the same if each car manufacturer shares its yearly sales data with a TTP to calculate the average. However, in the computation with MPC, the input data will not be revealed to a TTP or other car manufacturers. Instead, only the computation results will be revealed to RPs. Thanks to the secret sharing technique, the input data provided by IPs is encrypted and split into meaningless parts that will not reveal anything about the input data (Shamir, 1979). Further, the computation result is received by all IPs, which also act as RPs in our example. However, alternative scenarios are possible where other entities, like regulatory bodies or government institutions, act as RPs that request the same computation.

2.1.4. Security requirements and robustness toward adversaries

Lindell (2020) and C. Zhao et al. (2019) describe five security requirements of MPC. First, *privacy* implies that each party should only obtain information dedicated to them, not the output intended for other parties. Second, *correctness* means that each party must receive an accurate output. Third, *independence of input* stipulates that each party's input must be kept secret from other parties. Fourth, the *guarantee of output* ensures that honest parties receive their output without intervention from the

corrupted parties. Finally, *fairness* mandates that corrupted parties only receive their input when honest parties also receive theirs.

Prior reviews on MPC (Choi & Butler, 2019; Lindell, 2020; C. Zhao et al., 2019) also outlined three types of adversaries' behavior based on actions they may be able to do while deviating from the MPC protocol: semi-honest (or honest-but-curious), malicious, and covert. These classifications determine the robustness of MPC towards adversaries in the form of security models. In the s*emi-honest adversaries model*, adversaries participate in the computation correctly while passively observing protocol execution to gather information from other cooperating parties that should remain private. Although this adversary model is relatively weak, it ensures no unintentional data leakage while representing many practical scenarios. To illustrate using the same auction example, all bidders would like to maintain their reputation by not behaving maliciously, such as cheating. However, it is possible that each bidder would like to obtain sensitive information from other bidders as much as possible to possess competitive advantages.

Meanwhile, in the *malicious adversaries model*, adversaries actively deviate from the agreed protocol, aiming to manipulate the computation results or obtain other parties' private input. This is a strong security model that can guarantee protection from any adversarial attacks, but it reduces the efficiency of the protocol. This model is ideal for a scenario involving competitors that would like to perform joint computation tasks since participants might seek to maximize their benefits by behaving maliciously to cause an error in the computation (despite the possibility of being caught cheating). Further, in the *covert adversaries model*, adversaries behave maliciously like in the malicious adversaries model. However, there is a probability that this malicious behavior can be detected. In other words, adversaries in this model are willing to cheat as long as they are not caught cheating. Therefore, their possibilities to deviate from the agreed protocol are limited. This model is comparable to the business, financial, or political contexts because ensuring honest behavior among organizations is not

possible. However, they cannot risk damaging their reputation, financial penalty, or legal actions if caught cheating.

2.1.5. MPC architectures

MPC is currently computationally intensive, impractical, and might incur high costs for businesses/organizations interested in adopting the technology (C. Zhao et al., 2019). Hence, scholars are working on various deployment scenarios based on how CPs are set up (Alter et al., 2018), namely *private servers* and *cloud-assisted* (either *single cloud* or with *multiple cloud providers*). Each MPC architecture poses design trade-offs regarding trust requirements, security guarantees, complexity, and resource provision. In the following paragraphs, we review these three different MPC architectures and their implications on IPs regarding the aforementioned trade-off aspects. It should be noted that many other design trade-offs are out of the scope of MPC architecture and, therefore, not included in this review.

In the *private servers* model, all IPs also act as CPs, meaning that MPC software and computation servers are installed locally in each participant (Alter et al., 2018). This model represents an exemplary MPC implementation since no other entities (i.e., cloud servers) are involved as CPs to execute the computation. This model also provides a robust security guarantee, meaning that all IPs only need to trust other IPs not to collude and behave maliciously. However, this model requires extensive computational resources for executing the MPC protocol that each participating entity must provide. The communication cost is also expensive since each IP has to communicate with other IPs to perform the computation, making it necessary to have high-bandwidth, low-latency network connections. For these reasons, it is impractical to scale up this model to accommodate many IPs.

Given the drawbacks of the private servers model, there are attempts to leverage cloud providers to outsource the computation securely to increase the practicality and

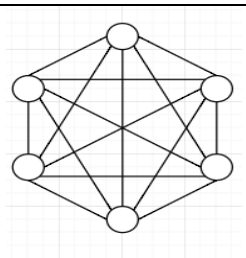
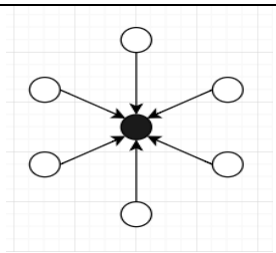
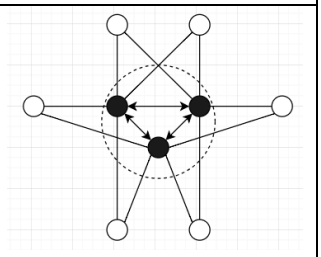
scalability of MPC. This model is called the *cloud-assisted* model, which treats cloud providers as CPs (rather than as TTP) while maintaining security guarantees like in the private servers model (C. Zhao et al., 2019). The use of cloud infrastructure to execute MPC protocol is justified because, although MPC requires massive resources, it is not necessary to use them all the time constantly. For this reason, this model is suitable for entities that would like to collaborate through privacy-enhancing data sharing and analytics but lack resources and expertise. Therefore, they can outsource the deployment and management of the MPC infrastructure to MPC service providers (Lapets et al., 2018). Nevertheless, delegating the expensive part of the computation to MPC service providers must be done while ensuring that the security is not compromised (Lapets et al., 2018).

The cloud-assisted model can be deployed with a single cloud or multiple cloud providers (Alter et al., 2018; C. Zhao et al., 2019). In the *single cloud* model, only one cloud provider is used as a CP that executes computation on encrypted data based on homomorphic encryption (HE) (Gentry, 2009; Naehrig et al., 2011). In this regard, MPC protocol is not the core of the architecture but rather complementary (yet important) in generating a public-key/private-key pair for the single cloud-assisted model (Alter et al., 2018). Specifically, each IP participates in the MPC protocol to generate a public key and a secret sharing of the private key for the single cloud model. Then, IPs encrypt their input data using the public key and upload the encrypted data to the cloud. Hence, when designated RPs make a query from the encrypted data, they can only decrypt it by engaging in another MPC protocol using the secret sharing of the private key they obtained before. This model is relatively robust despite adding another entity as a CP, as the input data and results stay encrypted during the computation and can only be decrypted by the designated RPs. Also, this model eases the computational burden for IPs since they do not have to provide their own MPC infrastructure, making it easier to scale up and add more IPs. A concern might be that the one cloud provider represents a single point of failure, which is vulnerable if

something goes wrong. Performance efficiency is another limitation of this model, as HE is also computationally intensive, making it challenging to perform complex functions.

Meanwhile, in the *multiple cloud providers* model, CPs comprise multiple independent providers that manage and deploy the cloud. This model leverages the secret sharing technique (see Section 2.1.2) by splitting the input data into parts and distributing them into different CPs. Subsequently, each CP communicates, executes the computation, and then shares the result with RPs. In this regard, RPs will only receive the result from the agreed computation function among parties, not the input data each IP provides. This model can be seen as a middle ground that balances the strengths and weaknesses of the private servers and single cloud models. On the one hand, using cloud servers would reduce (or even eliminate) the burden for IPs in deploying and maintaining costly computational infrastructure on their premise. This responsibility belongs to the cloud providers, guaranteeing high computational performance and efficiency. Not only that, but this model is also easier to scale up in terms of adding more IPs without creating more burdens for other IPs. On the other hand, using multiple cloud providers would make CPs less vulnerable to attack. Nevertheless, each IP has to distribute trust toward CPs so that each CP will not collude and behave maliciously. In other words, this model requires a greater trust requirement due to the involvement of multiple entities as CPs.

We summarize the comparison between three different MPC architectures in Table 4.

Characteristics	Private servers	Single cloud	Multiple cloud providers
Illustration			
Trust requirements of IPs	No trust toward CPs, but need to trust other IPs and trust that the computation runs correctly	IPs only need to trust one CP that it will not see and reveal the data	IPs distribute trust on multiple CPs deployed by multiple independent entities
Security guarantee	Robust security guarantee as CPs are placed in each IP (no separate CPs are involved)	Data stays encrypted during the computation and can only be decrypted by the designated RP, but CP is vulnerable to attack	CPs are less vulnerable to attack, but security is guaranteed as long as there is at least one honest CP
Complexity and resource provision	Each IP needs to provide extensive resources and execute the computation itself, making it challenging to scale-up	Easy to scale up, IPs do not need to provide resources, and CP can be deployed by other entities, but the homomorphic encryption is	Easy to scale up, IPs do not need to provide resources, and CPs have to be deployed by multiple entities to prevent collusion and malicious behavior

		computationally extensive than MPC	

Table 4 Comparison between three MPC architectures (adapted from Alter et al., 2018)

2.1.6. MPC use cases

Thanks to advances in computational power and efficiency, the theoretical concepts of MPC that have been around for some time (Yao, 1982) can now be translated into real-life applications (Lindell, 2020). Given the emerging nature of MPC, its usage is expected to grow massively in the coming years (Gartner, 2021). Here, we provide a non-exhaustive overview of some MPC use cases from various domains to show how MPC is being implemented presently.

One of the first large-scale and real-life applications of MPC was *auction-based pricing* for Danish sugar beet production contracts (Bogetoft et al., 2009). A concern in the traditional auction is that submitting bids to a TTP might reveal sensitive information about farmers' financial conditions that can disadvantage them. Hence, this use case leveraged MPC as a virtual auctioneer to replace a TTP in determining the market clearing price (i.e., the optimal price per unit where total supply equals total demand). In this use case, a (private) bid is submitted by each buyer (i.e., how much they are willing to buy at each price) and each sugar beet farmer (i.e., how much they are willing to sell at each price). MPC then computes the market clearing price and shares the result with all participating parties as a basis to sell/buy the agreed amount of sugar beet.

In the health domain, MPC has been used to conduct large-scale *genome-wide association studies* (GWAS) (Cho et al., 2018; Jagadeesh et al., 2017; Kamm et al., 2013). Such studies are essential for discovering and treating previously unrecognized chronic and rare diseases. However, individual genomes are personal data that might reveal sensitive information about, among others, one's health condition, resulting in

privacy concerns and reluctance to share such data. MPC enables collaborative computation of genomic data between multiple medical institutions to generate meaningful insights while keeping individual genomic data private. In another use case, MPC has also been used to *improve future prediction* of health risks, prognosis, and optimal treatment selection (Spini et al., 2022; van Egmond et al., 2021). This is done by combining sensitive healthcare-related data possessed by hospitals and health insurance companies, resulting in the prediction model for the involved parties without revealing anything about the input data. Thanks to MPC, it is possible to perform such analytics from sensitive and highly regulated data that are scattered across different actors.

In economics and finance, MPC enables collaboration between financial institutions, governments, tax authorities, and law enforcement agencies to *detect criminal activities* such as money laundering and tax fraud (Bogdanov et al., 2012, 2015; Sangers et al., 2019). MPC has also been used to *address economic inequalities* in the Greater Boston Area through collaborative analysis of salary data from 114 companies to identify the gender wage gap (Lapets et al., 2018, 2019; Qin et al., 2019). In both cases, each entity typically possesses relevant data that, when analyzed jointly, might be beneficial in identifying suspicious transactions, money flows, and aggregated salary differences across demographics. However, public and private institutions are often unable or unwilling to share such data due to privacy regulations, fear of losing a competitive edge, and reputational damage. With MPC, those actors can cooperate in performing privacy-enhancing analytics to generate meaningful insights for tackling financial crimes and reducing the pay gap without compromising privacy and confidentiality.

Finally, there are several unique and innovative use cases of MPC. For instance, Hemenway et al. (2016) and Kamm and Willemson (2015) used MPC in the space domain to enable satellite operators to jointly calculate collision probabilities without revealing sensitive information about satellites' orbit. In the energy sector, Zobiri et al.

(2022) leverage MPC to develop a privacy-enhancing demand response market so that consumers can get compensated for reducing the energy usage of their household devices without disclosing sensitive household electricity usage data. Moreover, in the automotive sector, scholars are exploring how MPC can be used in the connected autonomous vehicle (CAV), particularly to facilitate vehicle-to-vehicle (V2V) communication (T. Li et al., 2019) and enable cooperative object classification (Xiong et al., 2022) without leaking private information. Furthermore, MPC has also been used for social good by enabling sexual assault survivors to report perpetrators and identify repeat offenders while keeping their personal information private (Rajan et al., 2018).

We summarize our overview of MPC use cases in Table 5. It is important to note that these examples are non-exhaustive and not intended to be representative, especially due to the rapid development of MPC. This means there might be more innovative MPC use cases that have emerged recently and are not covered in this overview.

Use case	**Description**	**References**
Auction-based pricing	MPC is used in an auction to determine the market clearing price for Danish sugar beet production contracts	Bogetoft et al. (2009)
Genome-wide association studies	Collaborative genomic analysis for rare and chronic disease treatment without revealing sensitive health information	Cho et al. (2018), Jagadeesh et al. (2017), Kamm et al. (2013)
Health risk prediction and treatment	Patients' health prediction model based on joint analysis of hospital and insurance data without revealing sensitive information	van Egmond et al. (2021), Spini et al. (2022)

Fraud detection	Identifying money laundering and tax fraud activities by jointly computing data from private and public institutions	Bogdanov et al. (2012, 2015), Sangers et al. (2019)
Economic inequalities	Collaborative analysis of salary data from companies in Boston to identify the wage gap	Lapets et al. (2018, 2019), Qin et al. (2018)
Satellite collition prevention	Collaborative analytics on collision probabilities without revealing sensitive information about satellites' orbit	Hemenway et al. (2014), Kamm & Willemson (2015)
Energy transition	Privacy-enhancing demand response market for compensating household consumers without revealing electricity usage data	Zobiri et al. (2022)
Autonomous vehicle	Vehicle-to-vehicle communication and cooperative object classification without revealing information about the car and the driver	Li et al. (2019), Xiong et al. (2022)
Reporting sexual offenders	Enabling victims of sexual assault to report perpetrators and identify repeat offenders while keeping their personal information private	Rajan et al. (2018)

Table 5 Examples of MPC use cases

2.1.7. Adoption challenges

Despite recent advances and various real-life implementations of MPC that might bring value to businesses and society, it has not been widely adopted yet. Several technological, organizational, and legal factors hinder the large-scale adoption of MPC. From the technological perspective, MPC remains impractical due to high computational overhead, which is very extensive and limits its performance (Borking, 2011; Choi & Butler, 2019). MPC also still suffers from scalability issues, making it challenging to involve more participants and execute a more complex computation (Tõldsepp et al., 2012). Aware of these limitations, researchers have been developing various approaches and architectures to increase the practicalities of MPC, such as outsourcing parts of the computation to the cloud (see Section 2.1.5). It remains to be seen if such development would make MPC more attractive to end-users from a technical point of view.

Usability and transparency are other technological challenges hindering MPC adoption, which has also been a long-standing problem for encryption technologies in general (Gerber et al., 2018; Sheng et al., 2006; Whitten & Tygar, 1999). MPC is deemed highly complex for non-experts to understand, with some even describing MPC as "computation in the dark" that works like magic (Agrawal et al., 2021; Bruun et al., 2020). This lack of transparency makes it difficult for end-users to trace the connection between the input data and the computation results (Bruun et al., 2020). Despite this concern, raising awareness and understandability of prospective users regarding what MPC is and how it works is still not seen as a priority for developers (Agrawal et al., 2021; Kanger & Pruulmann-Vengerfeldt, 2015). In this regard, Evans et al. (2018) suggest that future MPC development should consider prospective non-expert users while building confidence in the technology through proper communication.

From the organizational point of view, the low maturity of MPC leads to uncertainty regarding the economic risks and costs of adopting MPC (Zöll et al., 2021). This is an essential consideration as companies often have limited resources for investing in new technologies without clear benefits, as in the case of MPC. Coupled with the presence of a functional substitute like a TTP, companies might decide against adopting MPC and changing their business practices (Kanger & Pruulmann-Vengerfeldt, 2015). Hence, addressing this organizational barrier requires an attempt to properly communicate the working principle of MPC and various business cases of which companies might be unaware (Evans et al., 2018; Kanger & Pruulmann-Vengerfeldt, 2015; Zöll et al., 2021).

Finally, from a legal perspective, the position of MPC is yet to be described clearly in light of data protection regulations (Choi & Butler, 2019; Zöll et al., 2021). Concerning this, researchers have attempted to show that companies using MPC could comply with the General Data Protection Regulation (GDPR) while benefitting from privacy-enhancing computation (Helminger & Rechberger, 2022). Nevertheless, further clarity is needed on whether using MPC complies with data protection regulations (Walsh et al., 2022). Hence, to help inform policymakers, Walsh et al. (2022) propose a framework for assessing MPC concerning data privacy laws based on (1) whether MPC deals with personal data, (2) whether executing MPC represents data disclosure activities, and (3) clarifying liability if something goes wrong. The expectation is that policymakers will be more aware of MPC in the coming years and start to look into its relevance in data protection regulation, which ultimately could boost its adoption.

2.2. Data marketplaces

2.2.1. Definitions

Data marketplaces can be broadly defined as multi-sided platforms facilitating data sharing and exchange among its participants (Fruhwirth, Rachinger, et al., 2020; Koutroumpis et al., 2020; M. Spiekermann, 2019). In data marketplaces, participants

can store, maintain, search, access, and exchange data from various sources, which are governed based on a wide range of standardized or negotiated licensing models (Alvsvåg et al., 2022; Schomm et al., 2013; Stahl et al., 2016). Both static and dynamic data can be exchanged in data marketplaces and accessible through different means like individual file downloads, web interfaces, or Application Programming Interfaces (APIs) (Fricker & Maksimov, 2017; M. Spiekermann, 2019). On top of that, data marketplaces also offer complementary applications and services to leverage the value of the data products (Alvsvåg et al., 2022; Koutroumpis et al., 2020).

There are various concepts in the literature that overlap with data marketplaces, such as data spaces, data platforms, and data collaboratives. The concept of data spaces, first introduced by Franklin et al. (2005), is initially described as a process in which different data sources are related (or associated) with each other. Recent work has taken a broader perspective in describing data spaces as an ecosystem that enables parties to share data based on agreed decision-making rights and processes to achieve shared goals (Beverungen et al., 2022; Otto, 2022; Scerri et al., 2022). Meanwhile, data platforms are a class of digital platforms whose main offerings revolve exclusively around data (de Reuver et al., 2022). This means that, in data platforms, data is viewed as a core product offered proactively by platform users and not the by-product or traces of platform usage. Data platforms comprise interoperable and extensible software components to support continuous, coordinated, and seamless data flow (Curry et al., 2022; Scerri et al., 2022). This way, data platforms facilitate data exchange and monetization between businesses while enabling third-party innovation in developing complementary services (de Reuver et al., 2022). Another similar term is data collaboratives, defined as collaboration initiatives among public and private organizations from different sectors to collect, share, or process data to address a societal challenge (Susha et al., 2017). A key enabler for data collaboratives is a trusted data intermediary (TDI) that ensures data availability,

accessibility, and usability for public purposes based on a wide range of business models (Susha et al., 2020).

Considering these similar conceptualizations of terms, we follow de Reuver et al. (2022) in positioning data marketplaces as specific instances of data platforms that enable data exchange among different entities. Subsequently, we follow Scerri et al. (2022) in positioning data platforms as part of a broader ecosystem of data spaces comprising various entities like businesses and data-driven service providers. Meanwhile, data collaboratives can be positioned as specific instances of data spaces formed based on public-private partnerships to achieve shared goals by addressing societal challenges (Susha et al., 2017). Data platforms, or data marketplaces in particular, can then be a TDI in the context of data collaboratives (Susha et al., 2020). Further, as data spaces can be distinguished between industrial and personal data spaces (Curry et al., 2022), we argue that data marketplaces can be uniquely positioned as intermediaries connecting various data spaces.

2.2.2. Data marketplaces in the automotive industry

As described in Section 1.4, we focus on the specific context of the automotive industry. In this industry, various data marketplaces like Caruso, Otonomo, and Automat are emerging (Bergman et al., 2022; Kaiser et al., 2021) due to the growing trends of digitalization that drive original equipment manufacturers (OEMs) like Bosch and Continental to open up access to their data and exchange data with other actors (Günther et al., 2017; Hartmann et al., 2016). The data exchanged can be *vehicle data* (i.e., data produced by sensors and electronic control units when vehicles are in use) or *context data* (i.e., additional data about the environment like geodata, weather, traffic, or map data), which can also be processed by OEMs before the exchange (Kaiser et al., 2021). This paradigm change enables new ways for OEMs to monetize their data by stimulating the development of innovative services beyond their main

products (Drees et al., 2021; Kaiser et al., 2021), such as connected cars, usage-based insurance, and shared mobility (Athanasopoulou et al., 2019; Kaiser et al., 2021).

Data marketplaces in the automotive industry operate based on diverse business models. According to Bergman et al. (2022), automotive data marketplaces can act as *aggregators* that recombine data from various data sources before reselling them to interested parties. Examples of automotive data marketplaces that follow the aggregator model are TomTom and INRIX, which provide tailored map data for their customers. Automotive data marketplaces can also complement their aggregator model with *an additional brokering service,* as in the case with HERE. Specifically, this data marketplace facilitates transactions between buyers and sellers on top of aggregating and recombining data into the customized map service. Moreover, automotive data marketplaces can apply the *consulting* business model in which they align the needs and preferences of participants in terms of data needs and price preferences. Caruso is an example of an automotive data marketplace that follows this consulting model. Specifically, Caruso provides a personalized service to mediate the needs and preferences of all participants interested in transacting their automotive data. Furthermore, *facilitating* data marketplaces such as IOTA and Ocean Protocol focus entirely on matchmaking between buyers and sellers through data brokering services.

2.2.3. Roles in the automotive data marketplaces

Generally, roles in the data marketplaces ecosystem comprise data marketplace operators, data providers, data buyers, and third-party service providers (Fruhwirth, Rachinger, et al., 2020; Kaiser et al., 2021; Koutroumpis et al., 2020; M. Spiekermann, 2019). **Data marketplace operators** match the supply and demand side of the market while also facilitating infrastructure provision for transporting data, contracts/licenses, and payments (Bergman et al., 2022). Then, on the supply side, **data providers** offer (access to) their data through data marketplaces to get monetary

compensation. In the automotive industry, this role of data providers can be filled by OEMs, individual vehicle users, or contextual data providers (i.e., companies that collect data about surroundings like traffic or weather). Meanwhile, on the demand side, **data buyers** are interested in acquiring (access to) the data they need and, therefore, perform queries in data marketplaces to look for available data. In the automotive industry, data buyers can also comprise OEMs, individual vehicle users, or data-driven automotive service providers that aim to improve customers' experience while driving their vehicles. Finally, **third-party service providers** offer complementary applications and services that can add value to data products in data marketplaces, such as data anonymizations, valuation, visualization, and analytics (Mucha & Seppala, 2020; M. Spiekermann, 2019). In supporting third-party service providers, data marketplace operators provide an environment for third-party development while also establishing an app store that stores third-party applications and services. This way, data marketplace operators ensure the availability and accessibility of those complementary services for data providers and data buyers. We illustrate roles in the ecosystem of automotive data marketplaces in Figure 4.

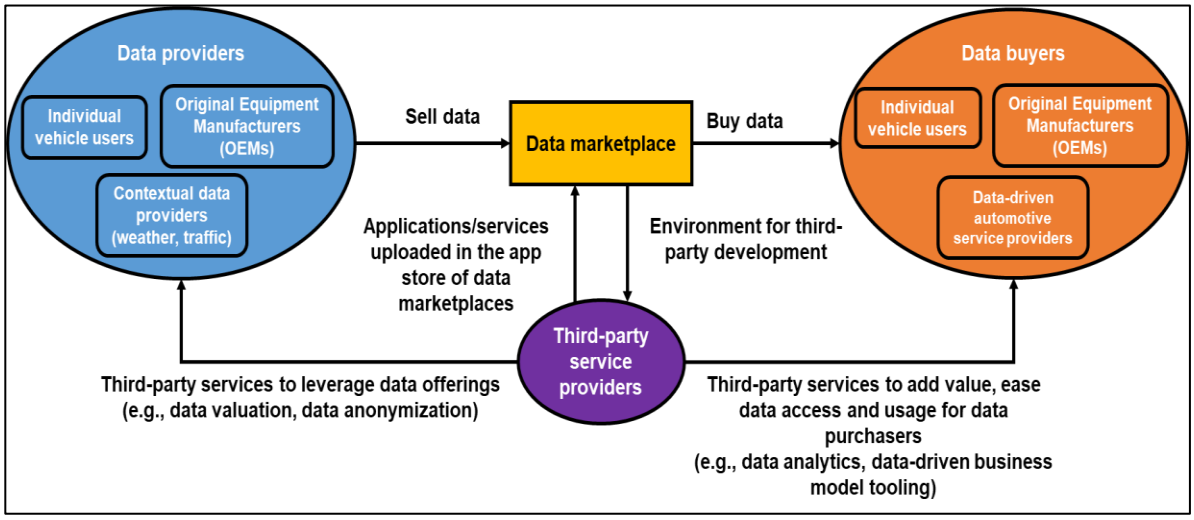


Figure 4 Roles in the automotive data marketplaces ecosystem (adapted from Kaiser et al. (2021) and M. Spiekermann (2019))

The presence of data marketplaces can create value for data providers, data buyers, and third-party service providers in different ways. Data providers can explore new revenue sources by selling their data assets in exchange for financial compensation. Data buyers, especially companies like OEMs, can quickly discover, access, and leverage the value of relevant external data to improve their existing business processes or develop new business offerings (Alvsvåg et al., 2022). For both actors, data marketplaces lower transaction costs for buying and selling data as it can help match supply and demand (Koutroumpis et al., 2020). Moreover, by providing an environment for developers, data marketplaces also stimulate innovation by third-party service providers to create innovative solutions that can benefit data providers and buyers in leveraging the data products (M. Spiekermann, 2019). Furthermore, data marketplaces can generate network effects in a way that more data providers can attract more data buyers and third-party service providers and vice versa (Koutroumpis et al., 2020).

2.2.4. Open challenges

Although data marketplaces are massively emerging, they are still struggling to maintain a strong position in the market due to the reluctance of data providers to share data and make it accessible to other parties (Jernigan et al., 2016; Richter & Slowinski, 2019). One of the challenges is the nature of data as experience goods, meaning that the value of data cannot be estimated without disclosing it. However, data is also non-rivalrous goods, meaning that it is easy to distribute data once it is disclosed, leading to the diminishing value of the data. This phenomenon is also called Arrow's Paradox (Arrow, 1972). Moreover, data providers often do not understand the cost related to data quality management. As a result, the value of data is even more difficult to assess. Furthermore, data characteristics as intermediate goods (i.e., needs further processing before use) lead to a lower willingness to pay for data than other goods and services.

Another challenge is the fear that individuals and companies could lose control over their data (M. Spiekermann, 2019). This challenge is also linked back to the nature of data as non-rivalrous goods. Once the data is disclosed, it can be easily shared and used relatively cheaply, making it even more challenging to track who got access to the data and the purpose of usage (Koutroumpis et al., 2020). Moreover, especially in the case of business data, companies also fear that their data might benefit other stakeholders (or even competitors) if they find a way to further process the data through analytics or aggregation with other data sources (Koutroumpis et al., 2020; M. Spiekermann, 2019). In addition, individuals are concerned about their privacy in data sharing because they fear their data might fall into the wrong hands and use it for purposes beyond their consent (Cichy et al., 2021; van Schaik et al., 2018). As a result, companies might worry about sharing through data marketplaces because it could harm their business interest. Individuals are also unwilling to share data as they want to protect their privacy and do not want companies to generate revenue from their personal data (Schwinghammer et al., 2022).

Furthermore, there is a lack of clarity regarding the legal framework for data sharing through data marketplaces (Alvsvåg et al., 2022; M. Spiekermann, 2019). Companies and individuals are often unaware of the extent to which data sharing is allowed. In this regard, companies particularly do not want to risk violating data privacy regulations, which might further harm their reputation. Hence, it is challenging to establish trust in data sharing, which is particularly important in a complex ecosystem of data marketplaces (Arnaut et al., 2018; Dahlberg & Nokkala, 2019; Kembro et al., 2017; Richter & Slowinski, 2019).

2.3. Conclusions

In this chapter, we have described the fundamentals of MPC, which is a cryptographic technique utilizing—on a high level—either garbled circuit or secret sharing techniques that enable joint computation between multiple parties, resulting in a meaningful

output without giving input data. We have outlined that MPC can be implemented in different architectural paradigms ranging from private servers to cloud-assisted models, which come with their trade-offs regarding security requirements and robustness toward adversaries. Recent progress in cryptographic research around MPC increased efficiency and reduced resource needs for implementing MPC. As such, a wider range of applications becomes computationally feasible, such as in the healthcare, energy, and financial sectors. Nevertheless, large-scale adoption of MPC is still limited due to technical, organizational, and legal challenges.

We have also demonstrated the conceptual foundation of data marketplaces and their position among interrelated concepts like data spaces, data platforms, and data collaboratives. Through synthesizing existing literature on data marketplaces, we identified data marketplace operators, data providers, data buyers, and third-party service providers as key roles within the ecosystem of data marketplaces. We then discussed various open challenges that hinder the large-scale adoption of data marketplaces: fear of losing control over data, privacy concerns, data valuation difficulties, and legal uncertainty.

Analyzing our findings, we argue that data marketplace operators need to rethink how data sharing should be conducted to address data sharing barriers. Then, drawing from the potential of MPC in addressing data sharing barriers in various use cases, we also see the potential of MPC in breaking the tension between sharing data, protecting privacy, and maintaining control in data marketplaces. Put differently, incorporating MPC into data marketplaces might address data sharing barriers, ultimately enabling further utilization and value creation from data to generate meaningful insights and stimulate innovation. In this regard, we expect that MPC could enable new architectural approaches for data marketplaces and change the value proposition for actors in the ecosystem. Nevertheless, how MPC would change the business models of data marketplaces is unexplored. As a result, we have little insight into antecedents of data sharing decisions that could be impacted by the use

of MPC in data marketplaces. Therefore, in the next chapter, we investigate the business model implications of MPC in data marketplaces to understand data sharing antecedents that are impacted by MPC use in data marketplaces.

3 Exploring data sharing antecedents potentially impacted by MPC: a business model analysis[1]

In this chapter, we answer the first research question: *what types of data sharing antecedents could be impacted by MPC use in data marketplaces?* In Chapter 2, we have identified challenges data marketplaces face in stimulating data sharing, indicating the need to explore new ways to increase willingness to share without compromising participants' wishes regarding privacy and control over data. Then, based on what MPC can offer, we found that MPC could address data sharing barriers in data marketplaces by sharing data insights while keeping the input data private. This way, MPC can break the tension between data privacy, data control, and data value creation (cf., Gast et al., 2019). Hence, to substantiate this finding, we now empirically investigate the business model implications of MPC in data marketplaces to understand how MPC addresses previously outlined challenges in sharing through data marketplaces. We then translate these challenges addressed by MPC as antecedents of data sharing decisions that MPC potentially impacts in data marketplaces. By identifying these antecedents, we can observe the implications of MPC in data marketplaces and narrow down our research regarding the impact of MPC on data sharing decisions by businesses and consumers throughout this dissertation.

---

[1] This chapter is based on: Agahari, W., Dolci, R., & de Reuver, M. (2021). Business model implications of privacy-preserving technologies in data marketplaces: The case of multi-party computation. In Proceedings of the *29th European Conference on Information Systems (ECIS 2021)*, A Virtual AIS Conference.

Because we need to understand which antecedents are impacted by MPC, the chapter first reviews the concept of business models and the Unified Business Model framework by Al-Debei & Avison (2010) that we used as our analytical lens. Subsequently, we present our interviewee selection strategy, data collection, and analysis approach. Then, we outline our findings on the value proposition, value architecture, and value finance of MPC in data marketplaces. Finally, we provide conclusions to this chapter by discussing the relevance of our findings in light of the rest of the dissertation.

3.1. Business models as an analytical framework

The business model construct was first introduced in Information Systems (IS) literature during the advent of the Internet (e.g., Alt & Zimmermann, 2001). Currently, business models are commonly used to understand the impact of digital technologies (e.g., Athanasopoulou et al., 2019; Bouwman et al., 2019). We use the business model construct to understand the link between digital technology on the one hand and value creation on the other (cf., Baden-Fuller & Haefliger, 2013; Chesbrough & Rosenbloom, 2002). Digital technologies can be a driver as well as an enabler for new business models (Bouwman et al., 2008; de Reuver et al., 2009).

We see a business model as a description of how firms create, deliver, and capture value (Osterwalder & Pigneur, 2010; Teece, 2010). For this chapter, we build upon the Unified Business Model Framework by Al-Debei and Avison (2010) as our analytical lens for two reasons. First, this framework is structured based on existing business model frameworks, making it comprehensive and sufficiently broad to capture relevant business model aspects. Second, other studies have been implementing this framework to analyze business models of data intermediaries (e.g., Janssen & Zuiderwijk, 2014; Ranerup et al., 2016; Susha et al., 2020), of which data marketplaces are an instance.

The framework comprises four dimensions. The *value proposition* dimension refers to business logic for creating value by providing products and services for targeted segments (i.e., data providers and data buyers). These values could enable utility, gains, benefits, opportunities, and possibilities for data providers and data buyers. The *value architecture* dimension can be defined as technological and organizational architecture like software, digital infrastructure, and data management systems required to provide products and services. Meanwhile, the *value finance* dimension is an arrangement of cost and revenue streams, including pricing strategies like fees, charges, subscriptions, and usage costs. Finally, the *value network* dimension covers a set of actors involved in creating value in data marketplaces. Our focus in this chapter is to analyze the value propositions, value architecture, and value finance of MPC for the three main actors in the value network of data marketplaces: data providers, data buyers, and data marketplace operators.

3.2. Methodology

In this chapter, we employed a qualitative approach through semi-structured interviews with experts and practitioners in the privacy and security domain. Given that we wanted to investigate MPC from a business model perspective within the context of data marketplaces, which is exploratory in nature, the qualitative approach is suitable for this chapter (Verschuren & Doorewaard, 2010).

3.2.1. Interviewee selection

To select our interviewees, we followed the judgment sampling strategy (Sekaran & Bougie, 2016), focusing on those with expertise in privacy and security in general and MPC in particular. We chose this sampling approach since we investigated new research areas, namely business model implications of MPC in data marketplaces (cf., Etikan et al., 2015). We started by using the references from Section 2.1 as the primary source to identify key scholars researching MPC. As a complement, we also look into

relevant reports and white papers to gather additional business actors. Moreover, we used our personal network by, for instance, selecting experts from European projects on MPC. In addition, we employed the snowball sampling approach by asking interviewees to suggest additional experts. From this sampling approach, we interviewed 15 experts from academia, research institutions, and businesses who averaged nine years of experience in the privacy and security domain (see Table 6). Only one of our interviewees is female. The interviews were conducted online from March until June 2020 in collaboration with a master's student (Dolci, 2020).

ID	Category	Role	Experience
I-01	Academia	PhD researcher in cryptography	2 years
I-02	Academia	PhD researcher in cryptography	2 years
I-03	Academia	PhD researcher in cryptography & cybersecurity	5 years
I-04	Academia	Research scientist in cryptography	5 years
I-05	Academia	PhD researcher in cryptography	2 years
I-06	Academia	Assistant professor in computational privacy	14 years
I-07	Academia	Professor in cryptography	20 years
I-08	Academia	Postdoctoral researcher in cryptography	6 years
I-09	Research Institution	Senior research scientist in data management	17 years
I-10	Research Institution	Cryptography specialist	7 years
I-11	Research Institution	Senior cryptography engineer	8 years
I-12	Research Institution	Senior research scientist in information security	15 years
I-13	MPC service provider	Chief Science Officer & Co-founder	15 years
I-14	MPC service provider	Chief Product Officer & founder	15 years

I-15	MPC service provider	Software Developer	2 years

Table 6 An overview of interviewees

3.2.2. Interview procedures and questions

For the interview, we developed a short presentation based on our synthesis in Chapter 2, covering (1) how data marketplaces and MPC work, (2) their use cases, and (3) a possible use case for MPC-based data marketplaces. To increase the understanding of the presentation, we used a fictitious scenario inspired by an explainer video developed by Boston University (see here: https://youtu.be/l25jcolQW6Q). We imagined that the transportation authority would like to identify popular pick-up spots for ride-sharing companies to develop policies to reduce traffic congestion. The authority would like to request this information from ride-sharing companies, but they might face resistance due to the sensitive nature of the data. Therefore, we proposed an MPC-based solution to facilitate privacy-enhancing data sharing between ride-sharing companies and the transportation authority. We present an excerpt of this presentation in Figure 5.

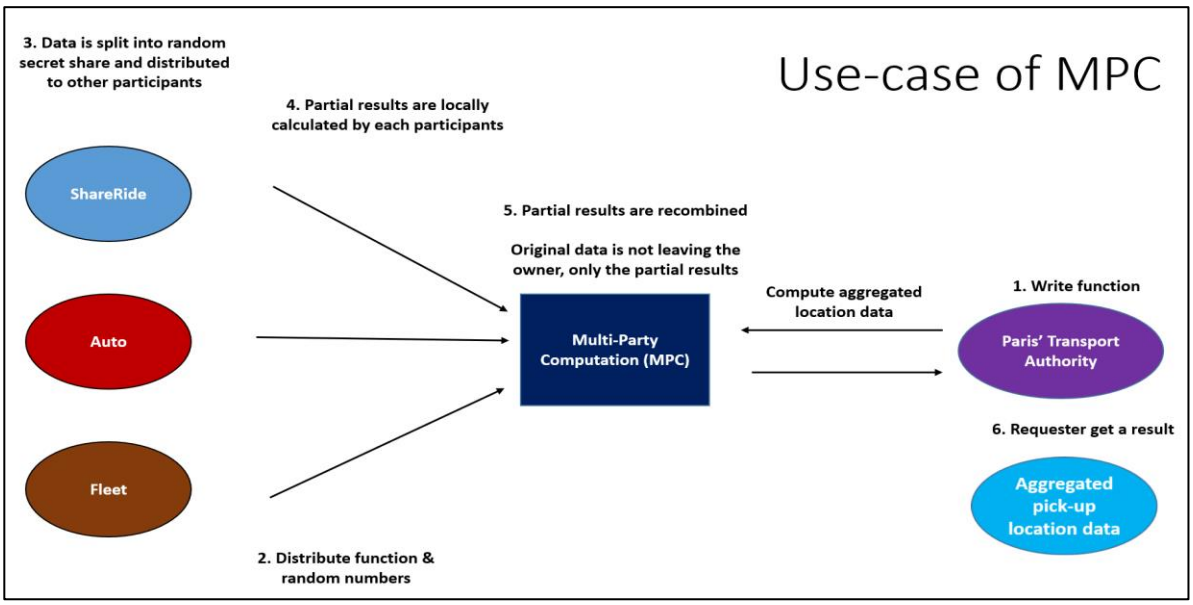


Figure 5 An excerpt of the presentation used for the interview with MPC experts

We then pitched the presentation at the beginning of the interview, followed by discussions with interviewees to gather feedback. We adapted the wording and illustration of use cases if necessary. This way, we ensure our data marketplaces and MPC descriptions are scientifically correct and easy for non-experts to understand. We also ascertained that interviewees had a similar understanding of data marketplaces and MPC. This is important as we used the same presentation to investigate the business perspective on MPC and data sharing decisions (Chapter 4). The complete presentation is accessible at the 4TU research data repository (see Appendix D).

After discussing the presentation on MPC and data marketplaces, we asked questions based on the interview protocol (see Appendix A), which consisted of two parts. For the first part, we asked how MPC works, what it can and cannot do, and its comparison with similar technology like homomorphic encryption. We did this to validate our understanding of MPC, which is essential when discussing the potential impact of MPC on business data sharing in the next phase of the research (Chapter 4). For the second part of the interview protocol, we derived questions based on the Unified Business Model framework we described in Section 3.1. We asked about the value that MPC could offer to businesses, how MPC could be implemented in data marketplaces, changes that might happen in the architecture of data marketplaces with MPC in place, and examples of real-life use cases. By asking these questions, we can obtain the perspective of MPC experts and explore the implications of MPC on business models of data marketplaces.

Each interview lasted around one hour. After informed consent, we recorded each interview and transcribed it anonymously. We also created notes to outline key insights and surprising remarks from the interviewees to support the analysis. We stored unique identifiers for each interviewee in a separate database, including sensitive information like name and company. To increase the validity of our findings (cf., Brink, 1993), we sent back transcripts to interviewees for approval.

3.2.3. Data analysis

We used Atlas.TI 8.0 to individually code and analyze each interview transcript (Bryant & Charmaz, 2007). We first performed open coding based on an initial set of codes derived from the Unified Business Model framework to guide the analysis. However, we kept an open mind for additional insights, which were added as additional codes. Next, all codes were combined, resulting in a long list of often similar and overlapping codes. After that, we compared findings between interviewees to assess consistency. Then, we merged and grouped commonalities into high-level concepts using axial coding. For instance, we grouped codes like *privacy-by-design*, *protecting their own data*, *sharing without sharing*, and *data will not be revealed* into one broader category of **privacy**. Finally, in selective coding, we referred to the Unified Business Model framework (see Section 3.1) to structure each category into three overarching themes: value proposition, value architecture, and value finance. See Appendix D for information on accessing the final coding list at the 4TU research data repository.

3.3. Results

3.3.1. Value proposition

Following our elaboration in Section 3.1, we outlined the value propositions of MPC for three main actors in the value network of data marketplaces: data providers, data buyers, and data marketplace operators.

Data providers

Most interviewees described **privacy** as one value proposition of MPC for data providers in data marketplaces. By default, MPC will protect input data from data providers from being revealed to anyone. As a result, individuals' privacy (i.e., end-users of data providers) will not be harmed. According to one interviewee (I-14):

> *"When the data is at rest, it is already encrypted. Once we want to use it, we apply a business function to it, and only the pre-agreed outcomes are public. So, you keep everything encrypted for as long as possible."*

Still, data buyers can obtain valuable insights from the computation. In this way, data providers are, in theory, not at risk of losing their data to, for instance, competitors. To illustrate this, one interviewee (I-13) mentioned:

> *"What is different about MPC is that you do not decrypt. You still encrypt the data before you share it, but you are able to not decrypt it at any point, and yet still … you get the insights, you get the results, you get the value from it. And neither party will need to decrypt anything. You just get the result at the end."*

Hence, data providers could share more sensitive data that would not be possible to share before.

> *"The main thing that MPC could be useful for is to allow competitors, in a sense, to share the data for meaningful insights without actually sharing the data."* (I-05)

Next, MPC also allows data providers to keep **control** of their data. Data providers receive strong security guarantees on how their data is used since they need to approve any computation function that runs through MPC. Put differently, data providers can decide what kind of queries can be performed by data buyers. One interviewee (I-06) stated:

> *"MPC is about privacy for individuals to be in control of their data. So that is definitely providing privacy, but it means control over the data. What can be shared? What can be seen? What can be processed?"*

Similarly, another interviewee (I-12) indicated that:

> *"In traditional data marketplaces, … you have the data you provided, and you lose control over it. With MPC, you can have sole control of what it is used for. You can allow only certain kinds of computations, and you can have some control over what aggregated values are revealed. So you remain in control."*

When it comes to trust, MPC enables **trustless** computation. Interviewees argued that mutually distrusting parties could collaborate using MPC to achieve a shared goal, which is performing joint computations to generate insights together. Put differently, data providers do not have to trust data buyers or other parties involved in the computation and can still get the output they need. Thanks to the robust security guarantee of MPC, data providers can maintain the secrecy of their data while taking part in the computation. As illustrated by the following quote (I-11):

> *"MPC is basically a computation of data between mutually distrusting parties. The aim is to achieve a computation on sensitive data among parties who do not trust each other without leaking any information on the data or affecting the integrity of the result based on a publicly known function to compute."*

Another interviewee (I-07) also supported this statement, saying that:

> *"The goal of MPC, essentially, is to achieve the same functionality without needing such a trusted party. In short, the aim is to execute a protocol among parties that do not necessarily trust each other."*

Finally, MPC **reduces data sharing risks** by facilitating a distributed exchange of insights. Data providers no longer need to transfer their input data centrally to data marketplace operators because the computation is performed directly between data providers and data buyers. In this regard, data providers remain in control of their data and approve each query. Further, as data stays with data providers, it is less likely that data will be transferred to another party for other purposes. One interviewee (I-02) argued that:

"[MPC] has a lot of potential because it gives the opportunity to share your data without sharing it essentially. So you can circumvent all these privacy-preserving limitations from the law and still use combined data."

Another interviewee (I-05) also expressed the same argument, saying:

"You can compute functions without needing to disclose the inputs or any intermediate results to the other parties."

Data buyers

One value proposition for data buyers is **privacy**. The inquiry or insights that data buyers intend to gain could be sensitive and might reveal their competitive advantage or even the private information of their end users. With MPC, data buyers can get results from the computation without revealing their query to data providers, preventing reverse engineering by data providers in the process. According to one interviewee (I-03):

"… maybe the buyers do not necessarily want to reveal that they are really interested in … because their competitors might also realize that [it] is important. … I think if some data market realized that you could … allow the companies to maintain privacy over their data … I feel like that is something that both buyers and sellers would want, especially if you had it at this meta-level where the buyers do not even have to reveal what it is they necessarily want."

The second value is **data availability**. With MPC in place, data providers may be willing to share more data through data marketplaces. As a result, data buyers would have access to more data and insights than before. As illustrated by one interviewee (I-13):

"[MPC] extends the opportunities in the marketplaces because now I have the option of both controlling my data and allowing people to use it."

Ultimately, more data buyers would use data marketplaces, creating network effects for both sides of the market.

> *"MPC could increase the revenues, the number of clients of data marketplaces. They have to create a trusted environment within the infrastructure of the marketplace and an encrypted channel to access this environment."* (I-14)

Data marketplace operators

We found three value propositions of MPC for data marketplace operators. First, MPC allows data marketplace operators to offer "data insights" based on **privacy-preserving analytics** instead of traditional data exchange. This approach would transform the core offering of data marketplaces from one dataset to an aggregation of multiple datasets. As illustrated by one interviewee (I-12):

> *"…the value is not that much, I think, in the data itself. But the value that you can extract from combining multiple data sources. … it is actually the combination of one data provider and another one that actually gives you a value."*

Second, MPC could be valuable for data marketplace operators in **reducing data sharing risks**. Typically, data marketplaces comprise a centralized architecture, meaning that data providers upload their data and store it in a central repository of the operator. Such data transfer creates liability for data marketplace operators for storing data they do not own. With MPC, data no longer needs to be transferred from data providers to data marketplace operators, as the computation is performed directly between data providers and data buyers. As a result, there will be no transfer of liability as well. Data providers remain in control of their data and approve each query.

> *"It does not matter how large or sophisticated a company you are, you may lose the data that's in your possession. So a marketplace provider that has a lot of*

sensitive data in its possession has a lot of liabilities. And if they use this [MPC]technology, they can start reducing those liabilities." (I-13)

Third, MPC enables new offerings of **privacy-preserving applications and services**. Data marketplace operators could use MPC as building blocks to develop innovative applications and services without compromising privacy. Put differently, MPC enables new offerings of privacy-preserving applications and services. In this way, data marketplace operators can strengthen their position as a platform that keeps data private. As one interviewee (I-13) put it:

"if I am a data market provider, I have my own services that are MPC compatible, and I give my customers software that can operate with those MPC services. And then they can use MPC to interact with my various offerings or products."

3.3.2. Value architecture

We discussed three main themes on how MPC affects the value architecture of data marketplaces: new roles for data marketplace operators, different computation processes, and deployment scenarios. We also outlined how these three themes are related.

New roles for data marketplace operators

Based on our interviews, we identify two new roles of data marketplace operators with MPC in place. In the first role of **data brokering**, data marketplace operators no longer facilitate data exchange since data does not have to leave the premises of data providers and data buyers.

"[Data marketplaces] could only provide the MPC protocols and the matchmaking service and never receive any actual data but just deploys the MPC protocols to the different parties, assigns them to different roles, and so on. … In this option, you believe that [data providers] compute the aggregated data and

then directly send it to [data buyers] without going through [data marketplaces]. This would be technically possible." (I-01)

As a result, data marketplaces no longer have to store data, allowing them to focus on coordination functions like matchmaking (e.g., mediating between data providers and data buyers) and governing which use cases and functions are allowed.

"So, you have the data exchange, and you have the market. And the market is being able to find each other and knowing what the other has. That is, of course, a function that can easily be centralized." (I-12)

Data marketplace operators will not be able to see the data, as the computation will be done in a peer-to-peer manner.

"It is important that even the data marketplaces cannot access the private data, but it remains an intermediary that makes the transaction happen. It is important that the data provider remains in full control [of] what [they] can show or not." (I-08)

Computation processes

Interviewees suggested two possible approaches to implementing MPC in data marketplaces. In the first approach, called **synchronous** computation, all participating parties (i.e., data providers and data buyers) need to be online simultaneously to conduct the computations. This scenario is the most common implementation of MPC.

"In the MPC case, you share the data, and you do the computation, but you have to communicate with the other nodes to perform this computation. In the end, you reveal the result. … In the MPC world, the data provider always has to be

online and available since they do part of the computation. Otherwise, it does not work." (I-04)

Synchronous computation only works if all participants have sufficient resources.

> *"If the companies who want to do MPC with each other are large organizations with IT departments and teams, and of course, they can all schedule something to happen at the same time."* (I-13)

For data providers that lack resources and computing power, MPC computations can run **asynchronous**ly. In this scenario, each data provider can participate (i.e., submit their data) at a different time, depending on their availability. A coordinating party in the center is needed to *"make it possible to do things in a different order or in a different schedule or not all at the same time."* (I-13). This coordinating party, which could be the data marketplace operator, does not see the data during the computation, meaning that they do not have to be trusted. As argued by one interviewee (I-13), they are *"only sending back and forth encrypted data or random numbers that are completely meaningless"* due to the secret-sharing protocol employed by MPC.

<u>MPC deployment scenario</u>

Most interviewees distinguish different architectures for deploying MPC, with different roles for data marketplace operators. In the first scenario of **peer-to-peer**, the MPC software is installed at data providers and data buyers to allow distributed computations.

> *"… everybody installs a program, … [then] all [of them] start this protocol together, and it takes [as] many rounds [as] it takes to get to completion. Then, they all get an output, and there was no intermediary whatsoever."* (I-03)

Since participants conduct the computations themselves, data marketplace operators do not see the data nor participate in the computation. Instead, their role could be *"… to provide the whole setup, the expertise on how to deploy the MPC protocols, [and] how to design them."* (I-01)

In the second scenario, intermediaries assist in coordinating the MPC process. Here, data marketplace operators provide both MPC software and computational infrastructure (i.e., computing server) as services offered to data providers and consumers. This computational infrastructure could involve **a single computing server** offered directly by the operator.

> *"You have one centralized person … [then] everybody has to mask their input in some way. Basically, this person is serving as a router. They are really running this thing, but it is all being run through this central router, but it is all encrypted. So, the central router does not learn anything."* (I-03)

Alternatively, it is also possible to have **multiple independent computing servers** deployed by multiple entities. In this scenario, data marketplace operators establish a consortium in which each company provides a server for the computation. As described by one interviewee (I-15):

> *"It could also be the case where a data marketplace is set up by different companies, and one company owns one of the engines, and another company owns the other one. So, they can create a sort of solid data foundation together, which is the MPC engine."*

In this way, we *can "split [the] trust, and you do not have to trust everyone fully, but if at least one of these [computing servers that] you trust behave honestly, then [it will be] fine."* (I-01) One interviewee (I-02) illustrated that:

"Let us say you have three computing servers, and the guys who want to share the data do a secret sharing that they share the input data amongst all of these three computing servers. And then you do not have to be present during the computation but only the three computing servers to do the computation."

Interrelation between roles, computation processes, and deployment scenario

There is an interrelation between MPC deployment scenarios, roles of data marketplace operators, and computation processes. For instance, all three MPC deployment scenarios suit the data brokering model. As described in this section earlier, this model focuses on providing a matchmaking service between data providers and buyers rather than facilitating data exchange between parties. Hence, data marketplace operators could opt for peer-to-peer architecture and provide technical expertise to install MPC protocols on the client side. This way, data is exchanged directly between data providers and buyers without involving the operator. However, data marketplace operators could also offer computational infrastructure as a service to ease the burden for data providers and buyers. In this regard, the intermediary architecture (either single or multiple servers) would be more suitable. While this architecture requires data marketplace operators to be involved in the computation, they would not be able to see the data as it remains encrypted throughout the process, and only data buyers can access the computation results.

Meanwhile, a peer-to-peer architecture is best suited for the data aggregator model. This model implies that data marketplace operators already own a wide range of data collected from various data providers. In other words, data marketplace operators are transforming into "data providers" that monetize their data. To do this, data marketplace operators could deploy MPC protocols on their side and offer technical expertise for the data buyer side. In this way, both parties could perform MPC to generate meaningful insights sold to data buyers.

Regarding the computation type, synchronous computation is most compatible with peer-to-peer architecture. MPC protocols generally require all parties to be online and present simultaneously. Peer-to-peer architecture would make this possible, as the MPC protocol would be installed at all parties, allowing them to be connected and present during the computation. The synchronous computation can be organized independently without having a trusted third party in the middle. Nevertheless, it is possible to implement multiple computing servers as intermediaries to facilitate synchronous computation. In this setting, all participating parties do not need to be present simultaneously, but only the multiple servers in the middle. For asynchronous computation, intermediaries have to be present to coordinate the computation process between all parties to participate at different points in time. For this reason, the intermediary architecture (either single or multiple computing servers) is the most suitable approach for synchronous computation.

We summarize this interrelation in Table 7.

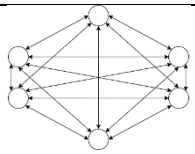
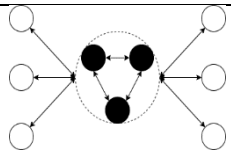
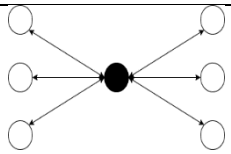
Aspect	**MPC deployment scenario in data marketplaces**		
	Peer-to-peer	**Intermediaries with multiple computing servers**	**Intermediaries with a single computing server**
Illustrations			
Role of data marketplace operators	Data broker & Data aggregator	Data broker	Data broker
Computation process	Synchronous	Synchronous & Asynchronous	Asynchronous

Table 7 The interrelation between roles of data marketplaces and computation process in three MPC deployment scenarios

3.3.3. Value finance

Interviewees suggested that, with MPC in place, data marketplace operators could generate revenue by offering MPC-as-a-service and MPC-based services. In the **MPC-as-a-service** model, data marketplace operators can offer technical expertise to data providers and buyers since the technology is relatively new for businesses. Examples include MPC software installation support and leasing the computational infrastructure to data providers and data buyers.

> *"… the data market [could] also offer an MPC node as a service. You pay the data market, for example, to also provide computing resources. And this is also a place where you could install an MPC node …"* (I-04)

This approach would benefit data providers and data buyers as they can reduce the cost of deploying the infrastructure.

> *"I think there has to be something like an automated setup that allows these smaller companies to quickly rent the capabilities from a cloud provider. … the data market needs to ensure that this all happens smoothly and that the smaller companies providing the data do not need their own infrastructure but can rent it on-demand …"* (I-05)

In the **MPC-based services** model, MPC is used as a building block to develop privacy-preserving applications and services, either by data marketplace operators or third-party service providers. In this regard, data marketplace operators could offer not only *"a simple exchange of data, but [could also offer] a service on exchange data that create new data."* (I-09)

One possible service mentioned by interviewees for the MPC-based services model is **privacy-preserving data valuation**. Data providers and data buyers could use MPC to explore whether there is a value for both parties to collaborate (e.g., combining

datasets). The computation would only reveal the newness of the insights in the form of a yes/no answer rather than revealing the datasets.

> *"What we are doing is essentially helping them take pieces of their workflow, like, for example, this notion of identifying whether or not there is value in data between two participants in the marketplace."* (I-13)

Another possible service is **privacy-preserving analytics** that delivers aggregated data insights without giving away the input data. One interviewee described that:

> *"the aggregation of the data could be interesting for a data buyer. … [MPC] can ensure that the buyer can receive only the aggregation without knowing the original inputs."* (I-10)

This possible service is highly relevant as current data marketplaces typically only allow data buyers to choose between acquiring all data or not acquiring anything at all. With MPC in place, data buyers can pay only for the insights/aggregation from multiple data providers instead of paying for an entire dataset.

> *"Now, there is an option in between everything or nothing. Right? They can buy this aggregated pickup location that is in the combination of three companies rather than just having to pay exorbitant amounts to get everybody's full database. … If you are a buyer, you want to be able to pay for just the insight and not the entire database. And I think this is a really good way to do it."* (I-03)

3.4. Discussion

Our findings show that, when used in data marketplaces, MPC enables new *value propositions* for data providers by facilitating a **distributed, trustless, privacy-enhancing data sharing** that maintains data **control** and reduces **risks**. This is important as data providers must ensure no leakage of end-users' personal data or

other sensitive data while participating in data marketplaces. In this way, we extend the work of Conger (2009) and Bonazzi et al. (2010), who argued that (1) end-users are pushing companies for strong privacy protection, and (2) companies are morally obliged to adopt privacy-enhancing technologies to protect the privacy of end-users. We demonstrated that adopting MPC could allow data marketplace operators, data providers, and data buyers to fulfill this moral obligation of protecting the privacy of end-users while still being able to create value from data. Furthermore, we complement the work by Bonazzi et al. (2010) and Conger et al. (2013) by exploring how MPC could address privacy problems in data sharing and enable privacy-friendly business models in data marketplaces.

Regarding the *value architecture*, MPC could substantially transform data marketplaces in terms of deployment scenario into either a **peer-to-peer model** (i.e., distributed architecture with MPC deployment in all data providers and data buyers) or **an intermediary model** (i.e., a centralized architecture where MPC is deployed centrally to orchestrate the computation). MPC could also change the role of data marketplaces into either **data brokers** (i.e., only as a matchmaker between data providers and data buyers) or **data aggregators** (i.e., only as a reseller of data collected from various sources). Further, implementing MPC in data marketplaces could be done through either **synchronous** (i.e., all parties should be online simultaneously) or **asynchronous** computation (i.e., only the computing party should be present all the time).

Each MPC architecture poses design trade-offs between trust requirements, complexity, and security guarantee, which are important for data marketplace operators. Thus, while other design trade-offs exist, they are not discussed within the scope of MPC architecture. In this regard, any architectural decisions will affect the viability of the platform and incentives for participants to join (Tiwana, 2014). On the one hand, a peer-to-peer model could provide a robust security guarantee and lower trust requirements at the expense of higher complexity and effort for all parties

involved. On the other hand, a single intermediary model could compromise security guarantee and trust requirement due to lower cost and complexity since the MPC protocol and infrastructure needs to be deployed centrally. A multiple-server architecture is a middle-ground alternative where each computing server is offered by an independent and unrelated entity that acts as an intermediary that performs MPC protocol. This approach could ease the onboarding process for data providers and data buyers but increase the complexity for data marketplace operators. Hence, we expand the work of Alter et al. (2018), who review three different MPC architectures (i.e., single cloud, multiple cloud providers, and private servers) and their trade-offs regarding trust requirements, performance, involvement of data providers, and scalability. We demonstrated that these MPC architectures are, in theory, also applicable within the context of data marketplaces. Nevertheless, since most MPC implementations in data marketplaces are still in the proof-of-concept phase, further research is required to investigate whether (1) those architectures are indeed applicable in practice and (2) the trade-offs are valid.

Regarding *value finance*, two important findings emerged. First, MPC enables new revenue sources for data marketplaces in the form of **MPC-as-a-service** (i.e., leasing MPC infrastructure and software) and **MPC-based services** (i.e., applications and services based on MPC to increase the value of data). Data marketplace operators can use a **subscription** model (i.e., monthly or yearly payments) or a **pay-per-use** model (i.e., only pay when needed) to generate revenue from these new offerings. Second, MPC shifts the core value of data marketplaces **from offering data to insights** (i.e., combining multiple datasets), allowing data buyers to look for something different (e.g., “what kind of insights are available in data marketplaces?” or “what kind of query that can I request in data marketplaces?”). In this way, data buyers could better assess data quality based on the usefulness of the insights, protecting them from the risk of purchasing data with unclear quality (Koutroumpis et al., 2020). Furthermore, like the decision on architectures, data marketplace operators should

carefully choose their monetization strategy as it will determine the participation and viability of the platform (Cusumano et al., 2019).

An unexpected finding was that MPC could be implemented in a wide range of services and functionalities within data marketplaces, suggesting that the value of MPC could go beyond the obvious use-case of data exchange. Specifically, MPC enables data marketplace operators to facilitate the creation of new privacy-preserving service offerings by providing an environment for third-party development. In this regard, MPC could be viewed as a boundary resource in the privacy-preserving data marketplaces, which, according to Ghazawneh & Henfridsson (2013), is "the software tools and regulations that serve as the interface for the arm's-length relationship between the platform owner and the application developer." Put differently, MPC could make data marketplaces extensible in a way that resembles other digital platforms (Tiwana et al., 2010). By leveraging MPC as generic and reusable components, data marketplaces could become "real" digital platforms that go beyond matchmaking features, creating values for participants through innovation from third-party service providers in an unforeseeable way (Kallinikos et al., 2013; Tilson et al., 2010). Such a mechanism would ultimately attract data providers and data buyers to participate in data marketplaces, creating network effects.

3.5. Limitations

A limitation of our exploratory study is that our samples are limited to MPC experts, making it possible that our findings only tell one side of the story. Hence, further research should incorporate the view of (prospective) data providers like companies that are already actively participating in data marketplaces. Another limitation is that our interviews were based on a thought experiment, given the lack of real-life implementation of MPC in data marketplaces. We suggest that scholars extend our study by exploring the value propositions of MPC based on a working prototype (or even a real-life application) of privacy-preserving data marketplaces based on MPC.

3.6. Conclusions

In this chapter, through business model analysis, we demonstrate the type of data sharing antecedents that could be impacted by implementing MPC in data marketplaces. Our interviews with MPC experts and practitioners demonstrate that MPC has characteristics that allow it to change the architecture and financial considerations of data marketplaces. These changes enable data marketplaces to facilitate a **distributed, trustless, privacy-enhancing data sharing** that maintains data **control** and reduces **risks**. In this regard, implementing MPC in data marketplaces could impact four types of data sharing antecedents: perceived control over data, privacy concerns, trust, and perceived risks. Hence, we focus on these data sharing antecedents for the subsequent studies. In the next chapter, we emphasize the business perspective to investigate the potential impact of MPC implementation in data marketplaces on these antecedents.

4 Understanding business perspective on MPC and data sharing: a qualitative study[2]

In this chapter, we answer the second research question: *what could be the impact of MPC use in data marketplaces on antecedents of data sharing by businesses?* In Chapter 3, based on the perspective of MPC experts, we found control, privacy, risks, and trust as data sharing antecedents that could be impacted by MPC use in data marketplaces. Hence, we investigate the perspective of business actors as (prospective) data providers to substantiate whether these antecedents are indeed impacted by MPC use in data marketplaces in the context of business data sharing. We then explore the potential impact of MPC usage in data marketplaces on these business data sharing antecedents.

Since this chapter is about the business perspective, we first review the literature on the organizational perspective on data sharing. We focus our review on control, risks, and trust as data sharing antecedents by businesses, as privacy is generally discussed within the context of end-users at the individual level and not about business-to-business relationships (Bélanger & Crossler, 2011; Smith et al., 2011). From this review, we establish a conceptual background on data sharing antecedents by businesses in data marketplaces. Then, we specify these antecedents to the MPC domain, which results in initial propositions on the potential impact of MPC use in data marketplaces on antecedents of data sharing by businesses. We subsequently

[2] This chapter is based on: 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*, 1–26.

introduce our methodology to evaluate our propositions, including our interview selection, procedures, and data analysis. As a next step, we present and discuss our findings concerning how MPC use in data marketplaces could impact antecedents of data sharing by businesses. We also synthesize boundary conditions under which MPC use in data marketplaces could impact antecedents of business data sharing. Finally, we discuss our findings, outline the limitations of this chapter, and answer the second research question in the conclusion section.

4.1. Organizational perspectives on data sharing: a review

Literature on organizational willingness to share data mainly draws from three theoretical perspectives on perceived control, trust, and perceived risks. In the following sub-section, we describe each theoretical stream and draw interrelations between them to better understand antecedents of business data sharing.

4.1.1. Perceived control over data

One stream of research explaining organizational (un)willingness to share data draws from control theory (J. Li et al., 2006; Stefansson, 2002). Generally, control refers to any attempt to ensure that the controlee (the target of control) behaves according to the objective of the controller (the source of control) (Tiwana et al., 2010; Wiener et al., 2016). Control is essential in the interaction between the controller and the controlee because their interests will likely be divergent. Hence, the controller typically exercises control via various mechanisms, such as technical artifacts, rules, and incentives, to create convergent goals between the controller and the controlee (Goldbach et al., 2018; Tiwana, 2014).

Current literature differentiates control mechanisms into two distinct types: formal and informal control (Goldbach et al., 2018; Mukhopadhyay et al., 2016; Tiwana, 2014). Formal control can more broadly be considered the visible aspects of control and is further divided into input, process, and output control. In *input control*, the controller

implements various selection and acceptance criteria that need to be fulfilled by the controlee before both parties interact. Meanwhile, *process control* focuses on aspects like rules, guidelines, and specific methods that the controlee needs to follow to ensure their behavior aligns with the controller. Furthermore, *output control* broadly includes specifications expected to be fulfilled by the controlee to maintain interaction with the controller.

Regarding informal control, two categories can be identified: self-control and relational/clan control (Goldbach et al., 2018; Mukhopadhyay et al., 2016; Tiwana, 2014). *Self-control* relies on the controlee's commitment to monitor their own behavior independently. Although the controlee implements it, the controller can provide tools and guidelines to strengthen the capacity of the controlee and encourage self-control. Meanwhile, in *clan/relational control*, all controlees are engaged in shared norms and values that can be encouraged by the controller. Ultimately, this could lead to mutual beliefs and common goals among the controlees in producing desirable outcomes that align with the controller's primary objective.

In the context of data sharing, control over data can be referred to as data providers' ability to define data usage by data buyers (Otto et al., 2019). Studies in this literature stream have shown that control over data plays a key role in the data-driven society as firms need to find a balance between protecting their data and sharing data to stimulate innovation (Gast et al., 2019; Otto et al., 2019; van den Broek & van Veenstra, 2018; Vimercati et al., 2021). Lack of control over data could result in firms' reluctance to share data, as they fear losing sensitive information that might benefit competitors (Arnaut et al., 2018; Richter & Slowinski, 2019). Hence, firms need to maintain control over who gets access to which data and for what purpose (Koutroumpis et al., 2020; Mosterd et al., 2021; Reimsbach-Kounatze, 2021). Another way is through a centralized structure since control can be exercised over who uses the shared information (Samaddar et al., 2006). In this way, firms can protect their valuable assets and maintain an advantage over competitors (Kembro et al., 2017; Nokkala et

al., 2019). Only after firms are able to control data usage and flow would they be more willing to share data with other firms (Dahlberg & Nokkala, 2019; De Prieëlle et al., 2020; Opriel et al., 2021).

4.1.2. Trust

The second stream explains the organizational willingness to share data by drawing on social-relational concepts such as trust, commitment, reciprocity, and values (e.g., Chen et al., 2014; Kolekofski & Heminger, 2003; Zaheer & Trkman, 2017). Studies in this stream draw on theories such as information sharing (Constant et al., 1994) and social exchange theory (Cropanzano & Mitchell, 2005; Emerson, 1976). For instance, using social exchange theory, Hall and Widén-Wulff (2008) found that the degree of social integration of firms based on trust with other partners is more important in influencing firms' decision to exchange information than financial incentives. Also, organizations are more likely to share data if they trust and have a committed and reciprocal relationship (Zaheer & Trkman, 2017). Moreover, trust also plays a vital role in firms' willingness to share data with other firms in the industry 4.0 context (Müller et al., 2020). Further, the willingness to exchange information is further strengthened as trusted collaboration grows between organizations (Du et al., 2012).

According to Mayer et al. (1995), trust is defined as the extent to which one party (i.e., the trustor) is willing to be vulnerable to the actions of another party (i.e., the trustee). Trust reduces tendencies for opportunistic behavior by firms (Morgan & Hunt, 1994). In data sharing, trust is central as a foundation to sustain interaction between firms (Chen et al., 2014; Richter & Slowinski, 2019; M. Spiekermann, 2019). Data providers need to trust that data buyers are committed to the agreement for data usage. Otherwise, data providers will refrain from sharing data (Kembro et al., 2017; Müller et al., 2020).

Prior research has identified various mechanisms that can be used to establish trust in data sharing between firms. One mechanism is technical solutions, as proposed by Ratnasingam et al. (2002). Examples include digital signatures, encryption, and authorization, which can be implemented as protective measures to ensure reliable data sharing transactions between firms. Another mechanism is screening and review (Richter & Slowinski, 2019; Son et al., 2006; Subramanian, 2017). Such mechanisms can help inform firms about the reputation of prospective data buyers before deciding to participate in data sharing. Finally, Noorian et al. (2014) proposed a data use agreement that clearly states the purpose of data usage, including the penalty that will be enforced in the event of a violation.

4.1.3. Perceived risks

The third stream of literature considers issues related to perceived risks in the online environment (Nicolaou & McKnight, 2006; Pavlou & Gefen, 2004). For example, in the healthcare domain, organizations are reluctant to share data due to various standards, regulations, and a lack of integration across healthcare systems (Azarm-Daigle et al., 2015; Harris et al., 2007). Such risks are evident in this context due to issues like information security and integrity (e.g., Shen et al., 2019) and standardization (e.g., Harris et al., 2007).

Following Pavlou and Gefen (2004), we define perceived risk as a firm's subjective belief of suffering a loss from the occurrence of an uncertain event. Unlike physical goods or other services, data characteristics might pose a higher risk in the context of data sharing for several reasons. First, competitors may use the data in ways that harm data providers' business interests. Through reverse engineering or de-anonymization, data buyers may identify critical business processes, harming the competitive advantage of data providers (M. Spiekermann, 2019). Second, the possibility to re-sell and re-share data at no cost once exchanged may create risks as unauthorized third parties can use the data in unforeseen ways (Koutroumpis et al.,

2020). Third, the possibility of combining the data and the ability to apply algorithms to the data may result in the de-anonymization of personal data and create privacy harm (H. Li et al., 2020).

4.1.4. Interrelations between perceived control, trust, and perceived risks

The three streams of literature provide an overview of concepts relevant to understanding inter-organizational data sharing. The first stream outlined the importance of control in preventing collaborating firms in data sharing from pursuing their self-interest alone. The second stream emphasizes the role of social and relational aspects, such as trust as a factor influencing organizational willingness to share data. However, the second stream does not inform us how trust could be established between firms without prior business relationships. Furthermore, the third stream focuses on risks stemming from data characteristics. It recognizes that even in the presence of control and trust, willingness to share data might be affected by the perceived risk associated with a transaction. However, the third stream does not discuss how the perceived risk of data sharing could be impacted by trust among partners or control to influence the willingness to share data. In this regard, when viewed separately, the three streams cannot comprehensively explain why firms share data with other firms.

Analyzing the literature, we argue that despite being three distinct streams focusing on different concepts of perceived control, trust, and perceived risks, each stream complements our understanding of firms' willingness to share data. Specifically, the three streams are inherently related in a way that trusts and risks are seen as consequences of (lack of) control. For instance, firms struggle to maintain control over what and how data buyers might use data once it is shared (Asare et al., 2016). This lack of control could create risks for data providers if they engage in data sharing, like becoming vulnerable to losing competitive advantage or harming the privacy of their end-users. In this regard, trust among organizations could reduce the tendency

for opportunistic behavior by firms using the data in the presence of relatively limited control (Emsley & Kidon, 2007; Kagal et al., 2001). Similarly, having control mechanisms in place is also essential in data sharing since it could reduce risks and establish trust between data providers and data buyers in data marketplaces (cf., Bons et al., 1998, 2012). Thus, the interrelation between perceived control, trust, and perceived risks suggests that they cannot be separated in investigating the potential impact of MPC use in data marketplaces on antecedents of business data sharing. In the next section, we specify these concepts of perceived control, trust, and perceived risks to the MPC domain to understand the expected mechanisms through which MPC could impact antecedents of business data sharing in data marketplaces.

4.2. Specifying conceptual background to the MPC domain: initial propositions

So far, we have derived insights from MPC literature regarding its characteristics and capabilities in computing multiple data sources while keeping them private (Chapter 2). We have also identified that implementing MPC in data marketplaces could impact perceived control, trust, and perceived risks as antecedents of business data sharing (Chapter 3). Subsequently, we synthesize conceptual background based on existing literature on organizational perspective on data sharing (Section 4.1). From this knowledge, we can now specify our conceptual background to the MPC domain in data marketplaces setting. Specifically, we derive initial propositions that propose the impact of MPC use in data marketplaces on antecedents of business data sharing.

First, as defined by scholars researching MPC in Section 2.1.1, MPC enables joint computation between multiple parties and only shares the results (data insights). Put differently, MPC allows distributed data sharing without storing it centrally. MPC could also eliminate the need for Trusted Third Parties (TTP) that perform data analysis and processing. As a result, data marketplaces would only act as a broker that performs matchmaking between data providers and buyers. The computation will be performed

automatically, resulting in aggregated insights (instead of datasets) that restrict how data buyers utilize the data. As we found in the business model analysis in Chapter 3, this approach represents a change in how data is stored and processed, allowing firms to regain control while sharing data through data marketplaces. In short, given the characteristics of MPC in exercising control for data providers in data marketplaces, our study should consider perceived control as a relevant antecedent of business data sharing. As such, the first initial proposition is:

P1. Perceived control over data is more relevant for data providers while sharing through data marketplaces that use MPC.

Second, the characteristics of MPC (see Section 2.1) makes it possible for data providers and buyers to perform computation together and generate insights while keeping data secure. In this regard, in line with our findings in Section 3.3.1, there is no need to establish trust between both parties. Trust in intermediaries is also eliminated since MPC removes the need for a TTP to perform data processing and analysis. Thus, implementing MPC in data marketplaces could impact trust in data sharing in a way that trust might be less relevant. Therefore, our second initial proposition of this study is:

P2. Trust is less relevant for data providers while sharing through data marketplaces that use MPC.

Third, as our domain exploration (Chapter 2) and business model analysis (Chapter 3) show, implementing MPC in data marketplaces allows data buyers to only receive computation results and not the original datasets. This way, data providers could have more control over how the data is processed and utilized by data buyers. Hence, data providers might no longer feel at risk of sharing data because their data stays with them during the computation. In this regard, MPC use in data marketplaces could

impact perceived risks in data sharing in a way that perceived risks might be less relevant. Therefore, we propose that:

P3. Perceived risks are less relevant for data providers while sharing through data marketplaces that use MPC.

Table 8 summarizes the initial propositions based on the theoretical concepts, which are evaluated and refined based on the empirical findings of our study.

Concept	**Initial proposition**
Perceived control over data	P1. Perceived control over data is more relevant for data providers while sharing through data marketplaces that use MPC
Trust	P2. Trust is less relevant for data providers while sharing through data marketplaces that use MPC
Perceived risks	P3. Perceived risks are less relevant for data providers while sharing through data marketplaces that use MPC

Table 8 Initial propositions

4.3. Methodology

In this chapter, we opt for a qualitative approach, given the exploratory nature of the first research question (Verschuren & Doorewaard, 2010). Specifically, MPC is a relatively novel technology with limited practical deployment in the context of automotive data marketplaces. Thus, while research has been done on control, trust, and perceived risk as antecedents of business-to-business data sharing, empirical studies on the impact of MPC are scarce. As a result, there is a limited understanding and prior work in automotive as an instance of our general perspective. Hence, the qualitative approach allows us to investigate the “why” and “how” concerning the possible link between MPC and implications for perceived control, trust, and perceived risks, as well as conditions for these implications to materialize (Recker, 2021). We

opted for semi-structured interviews with experts and practitioners as our data collection strategy. This technique is beneficial to be used in our study as it offers both rigidity (i.e., guided by pre-defined general questions) and flexibility (i.e., allows improvisation based on interactions with interviewees) (Kallio et al., 2016).

4.3.1. Interviewee selection

We used a judgment sampling approach to recruit our interviewees (Sekaran & Bougie, 2016). By leveraging our networks, we selected experts and practitioners in the automotive and mobility industries with data-related roles in their companies. To expand our interviewee candidates, we consulted grey literature such as reports and white papers to gather perspectives from the industry and get an overview of currently active business actors. We also searched LinkedIn using keywords like “data sharing”, “connected cars”, “automotive”, and “mobility” to identify potential interviewees. Besides that, we also target academic experts who work on data marketplaces, data platforms, automotive data sharing, and connected cars. We then created a shortlist of 54 potential interviewees, which we contacted via e-mail and LinkedIn. After each interview, we employed a snowball sampling approach by asking our interviewees to recommend other potential interviewees. We stopped searching for interviewees once the last round of interviews did not provide new information.

Table 9 presents an overview of our interviewees, organized based on the order of the interview. Twenty-three interviews with automotive experts and practitioners were conducted online from June to October 2020, with sixteen of them coming from businesses. From this number, three interviewees worked in relatively new companies in the automotive and mobility sector, while the rest worked in established companies. All of our interviewees are men and hold positions at a senior management level with an average of nine years of experience.

ID	Organization	Type	Profile	Experience (in years)
A01	Research institution	Expert	Project manager and doctoral researcher (B2B digital platforms)	7
A02	Not-for-profit research and consulting institution	Expert	Researcher and project manager (data spaces in the mobility sectors)	5
A03	Platform integrating shared mobility services	Newcomers	Head of partnerships & business development	3
A04	Research institution	Expert	Scientific director (IoT and business model innovation)	10+
A05	Insurance company	Established players	Fraud investigation specialist	10+
A06	Technology advisory and consultancy service	Established players	CEO	10+
A07	Mobility software and data analytics service provider	Newcomers	Business development consultant (transport and mobility)	10+
A08	Innovation lab for data marketplaces technologies	Expert	Initiator and digital connectivity lead	9
A09	Research institution	Expert	Senior scientist (transport and urban mobility)	10+

A10	Payment provider	Established players	Head of connected car & IoT	7
A11	Automotive R&D company	Established players	Product line manager (data intelligence)	10+
A12	Mobility service provider	Established players	Senior product manager (dynamic services)	10+
A13	Car OEM	Established players	Function owner for privacy management	4
A14	Advisory and consulting	Established players	Associate director & advisory (mobility & automotive)	7
A15	Automotive supplier	Established players	Senior manager (IoT business model innovation)	8
A16	Corporate mobility consulting service	Newcomers	CIO	5
A17	Automotive R&D center	Expert	Senior researcher (connected car)	8
A18	Fleet management software provider	Established players	Product manager (connected car)	7
A19	Platform integrating connected car services	Newcomers	Co-founder and head of data transformation	10+
A20	Car OEM	Established players	Business development manager (connected car)	10+
A21	Car OEM	Established players	Project manager (connected car)	3

A22	Car OEM	Established players	Product owner (car app store)	4
A23	Automotive bodyshop association	Expert	Public affairs & communication	4

Table 9 Overview of interviewees

4.3.2. Interview procedures and questions

We used a short presentation on the conceptual foundation of MPC and data marketplaces we developed and validated with MPC experts, as described in Section 3.2.2 in Chapter 3. We did this to ensure interviewees had the same baseline understanding of data marketplaces and MPC so that they could reflect on the potential impact of MPC on data sharing antecedents by businesses. We gave the presentation at the beginning of each session of one interview, in which we offered interviewees an opportunity to clarify and discuss each concept to reach a common understanding. We refined the presentation based on any feedback we received, which was shown in the subsequent interview. This way, we can improve interviewees' understanding of MPC definitions and use cases. Although this approach could lead to a potential bias due to interviewees' reliance on our explanation of MPC in answering interview questions, it is justified since MPC is still in an early stage of development and adoption by businesses, with few known implementations of MPC in data marketplaces. Further, our presentation was developed based on the domain exploration (see Chapter 2), validated with MPC experts (see Section 3.2.2 in Chapter 3), and kept constant in each interview, thus limiting the potential bias resulting from our approach. Figure 6 shows an excerpt of the presentation, while the complete presentation is publicly accessible at the 4TU research data repository (see Appendix D).

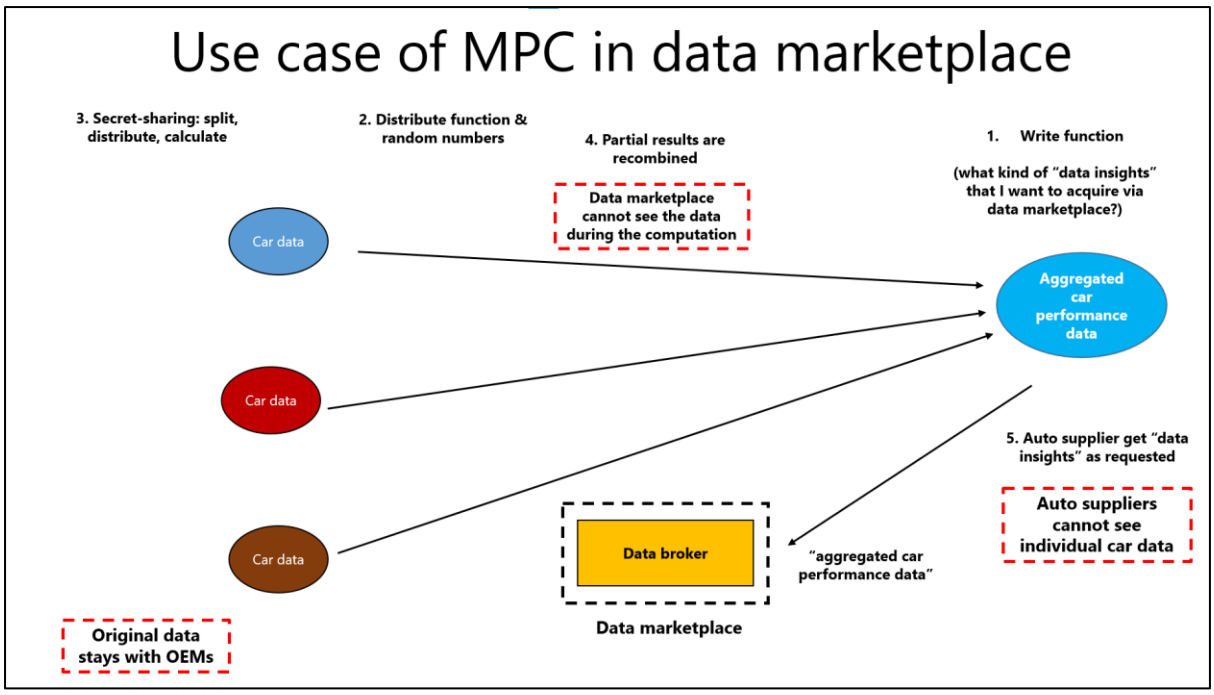


Figure 6 An excerpt of the presentation used for interviews with business actors

To guide the interview, we developed a protocol based on the conceptual background we identified in Section 4.1. The questions for each concept comprise one question about the current situation of data sharing without MPC and one question about the potential impact of MPC (see Table 10 and Appendix B). In the first part, after introducing the concept of data marketplaces, we asked questions about the current data sharing situation (without MPC). We did this to get interviewees in the mood to talk about data sharing and invite them to be as close to the current situation as possible. Then, we moved to the second part of the interview, starting with introducing interviewees to MPC and its possible use case in data marketplaces. We then asked questions about the possible impact of MPC on each concept. We did this to allow interviewees to reflect critically on how MPC could change the current data sharing situation. By dividing the interview into two parts (with and without MPC), we can better understand the baseline conditions of data sharing without MPC and compare them with the expected impact of MPC in data sharing.

Theoretical concept	Guiding questions	
	Part 1: data sharing without MPC	**Part 2: data sharing with MPC**
Organizational willingness to share data	What are the reasons behind your company's decision to share/not share data in data marketplaces?	With MPC in place, would it change your opinion on sharing those data in data marketplaces? Why?
Perceived control over data	What kind of control over data would you want while sharing in data marketplaces, and why?	With MPC in place, do you expect to have more or less control over those data while sharing in data marketplaces? Why?
Trust	What about trust towards other business actors that would get access to the data? How does it play a role in your company's decision to participate in data sharing?	With MPC in place, do you expect to have more or less trust towards other business actors while sharing those data in data marketplaces? Why?
Perceived risks	What risks might emerge if your company starts to share data in data marketplaces? How do these risks play a role in your company's decision to share those data?	With MPC in place, do you expect to encounter more or fewer risks of sharing those data in data marketplaces? Why?

Table 10 Interview protocol

Each interview was conducted online via Microsoft Teams and lasted one hour on average. All interviews were recorded, transcribed, and sent back to interviewees for approval. We anonymized transcripts to prevent revealing confidential information.

After each interview, we wrote down key insights and interesting remarks as input for analysis.

4.3.3. Data analysis

Each interview was coded and analyzed individually using ATLAS.ti 9.0. We followed the three phases of coding qualitative data as suggested by Bryant and Charmaz (2007): open, axial, and selective coding. In the open coding phase, we used an initial code list based on the theoretical concepts, the potential impact of MPC, and boundary conditions to guide the analysis. We assigned these codes to each statement in the transcript based on its relevance. However, we included additional and unexpected insights beyond the theoretical concepts as additional codes. This process of keeping an open mind is important to prevent missing out on insights that might explain our findings (Miles & Huberman, 1994). Examples of coding schemes are provided in Table 11.

Quote	**Assigned codes**
"... they will have much more control. And as you have more control over the data, you do not have to have the same level of trust in the other party. The more control you have of the data, the more you control how it can be used. The less you have to trust in the partner not cheating on you." (A06)	• MPC could increase control over data, only sharing insights, not an input data • MPC could increase trust between data providers and data buyers
"... what is the data going to be used for? I can give you an example. When we ask for sample data from an OEM, they put in the contract that they want to be informed about what things we want to	• Control: strict terms & conditions/data sharing agreements • Data sharing is based on the use case • Risk: data misuse risk

develop from the data. They want to have pretty good insight into what we are doing with the data. And, of course, it is very restricted. You can only use it for those purposes. [In that case,] no misuse is happening." (A12)	• Trust: what are the intended party does with the data
"At the moment, I think when you want to develop a service, you must discuss with them and present your idea. If they see a chance that there might be a risk, then they will discuss it with you. And maybe [they will] prohibited it and say you are not allowed to do this and that with the data. We have GDPR in Europe, so, in general, I think you must say what you will do with the data before you collect it in terms and conditions. So, I think that is what they want to check." (A17)	• control: authorization • control: strict terms & conditions/data sharing agreements • control: who is using the data and for what purpose
"I think the concept is very promising since it does not require you to share the original data or, let's say, the sensitive data outside. It basically provides you a means or a way how to share this data without actually sharing it, which is great. But then, as I mentioned again, the second issue, next to the part about the broker's involvement, is the ambiguous definition." (A21)	• MPC could increase control over data, only share insights not input data • MPC enables sharing data while preserving the confidentiality • MPC: need to understand the viability of MPC

Table 11 Examples of coding schemes

In the axial coding phase, we combined codes from all transcripts, resulting in a long list of similar and overlapping codes. Then, we analyzed similarities and relations between codes and merged them into high-level concepts. For instance, we grouped codes such as *agreements*, *contracts*, and *consent* into one broader category of *contract-based control*. See Table 12 for examples of merged codes. In this phase, we also reconsider and adapt categories and sub-categories when needed. For example, the "boundary conditions" category was not considered in the first coding round. It is only added during the axial coding because it explains conditions under which MPC impacts control, trust, and perceived risks in data sharing.

Original codes	**New code after merging**
Authorization	Contract-based control
Contract	
Data sharing agreements	
Trust towards the partner	Trust in actors
OEMs consider sharing data with trusted parties	
Not too much data sharing between OEMs	
Knowledge spill over	Competitiveness risk
Risk of having direct competition with others in selling data	
The benefit of using MPC is unclear	Perceived benefits
Need to understand if MPC changes the business model and data sharing landscape	
MPC puts back some burden of providing data to OEMs	Organizational readiness

With MPC, data buyers still need to do data cleaning	
Managerial maturity	
OEMs only willing to share generic, non-sensitive data	Perceived data sensitivity
MPC would not change the willingness to share strategically relevant data	
Data sharing depends on the data type	

Table 12 Examples of code merging in the axial coding phase

Finally, in the selective coding phase, we re-examined all codes to understand how they relate to the main topic of the potential impact of MPC use in data marketplaces on perceived control, trust, and perceived risks as antecedents of business data sharing. In doing this, we first reviewed memos and notes written down during each interview to develop argument lines. Then, we identified and structured the connections between codes and high-level concepts, which resulted in a preliminary summary outlining the possible impact of MPC on antecedents of business data sharing and its boundary conditions. After that, we critically reviewed categories and sub-categories to check for their interrelations, arguments, and consistencies. Based on this review, we made further changes to codes if necessary. Lastly, from the final set of codes, we evaluated and refined our initial propositions outlined in Table 8. This resulted in five propositions indicating the potential impact of MPC use in data marketplaces on antecedents of data sharing by businesses, as shown in Table 13. The grounding table of categories and sub-categories is accessible at the 4TU research data repository (See Appendix D).

4.4. Results

This section presents our findings based on three coding rounds. In discussing perceived control, trust, perceived risks, and boundary conditions, we first elaborate

on interviewees' views concerning the current data sharing situation without MPC (as-is conditions). This serves as a baseline for the current situation, which is essential because we want to know the changes that resulted from implementing MPC. Then, we outline the possible implications of MPC on those data sharing antecedents by businesses (to-be conditions). We use an identifier from Table 9 (e.g. (A01)) to refer to the interviewees throughout this section.

4.4.1. Perceived control over data

Most interviewees generally agreed on the importance of having control while sharing via data marketplaces. As data providers, firms demand information about who the data buyers are, what kind of data they need, and the purpose of data usage (A14). This is important so that firms can avoid mistakenly giving away (access to) sensitive data that are not supposed to be shared (A21). Firms also "still want to own the data" (A09) and maintain control "in a way that data cannot be manipulated" (A03). Hence, during data sharing via data marketplaces, firms would like to know "where [the data] is going, . . . so we know where our data is at every point of time" (A21). In this regard, ensuring compliance with data sharing rules and agreements is necessary (A08). Once data buyers violate this agreement, firms should be able to refrain from data sharing (A07, A22) to maintain control over the flow of their data (A07). Nevertheless, interviewees are well aware that such requirements of control over data are challenging to realize in practice because "if you give somebody a dataset, you can hardly regulate it and cannot find out what people ultimately do with it" (A04).

Our interviewees outlined various control sources that should be present within the context of data sharing. Based on their elaboration, we clustered those control sources by comparing their similarities and differences. Our clustering resulted in three self-developed categories representing different control sources in data sharing. First, *contract-based control* refers to an arrangement to govern providers-buyers relationships concerning data access and usage. One example includes data-sharing

agreements and contracts to define the purpose of using the data (A14), which is vital as some data is highly confidential and cannot be utilized beyond the agreed use case. Another mechanism is authorization, meaning that only parties with agreement and permission can get (access to) the data (A03). This mechanism could also be adjusted to allow different data buyers to access different data types. In this way, data providers can ensure that "the data reaches only [data buyers] that is intended to have the data" (A06) and "others who have a different security level are not allowed to look at that data" (A09).

Second, *structural-based control* refers to how the network relationships between data providers and buyers are structured. One way to implement this control mechanism is by keeping the data on the premises of data providers (A07) and "only provide [data buyers with] a way to process this data [while] not giving them access to the entire data" (A20). Firms could also opt for a bilateral partnership without intermediaries. Some interviewees prefer this approach because "it gets tricky whenever you have someone who is managing, storing, and brokering your data for you" (A21). Alternatively, companies could also implement this control mechanism by sharing data with existing partners in a closed ecosystem. This setup is more restricted and typically only filled by a network of firms that have already worked together for a long time (A19). Because firms do not want to destroy long-standing business relationships, firms are more likely to be more compliant (A14).

Third, *technology-based control* refers to any technological solutions to enforce control for data providers in data sharing. Interviewees frequently mentioned technical approaches like anonymization, encryption, and aggregation as ways to prevent data misuse, comply with privacy laws, and ensure that the data cannot be traced back to individual people (A09, A12). As long as these control measures are present, interviewees argued that companies would have no issue sharing and monetizing their data through data marketplaces (A12, A20). Nevertheless, some interviewees questioned whether "the technology is the right way to solve [the problem of control]"

(A04) because “from the technical aspect, the technology is known . . . and not the issue [in data sharing]” (A21). Instead, some interviewees suggest that measures like contracts and rules are the most appropriate solution (A04).

Interestingly, interviewees pointed out a trade-off between data usability and data misuse risk concerning control in data sharing. On the one hand, stricter control might result in unusable data. On the other hand, less control might lead to data misuse, and the data may end up somewhere unwanted. As one interviewee put it:

> *“You want to introduce these anonymization measures. But, . . . you can be very restrictive. The use case of the data that can be used is decreasing, so the data becomes less valuable for the market. That’s an important balance to keep. . . . [Y]ou need to also take into account that it doesn’t hurt the other part of the equation so that the data is still usable and still both for kind of offline and real-time use cases.”* (A12)

Zooming in on MPC, our interviewees expressed positive impressions towards this technology as a *technology-based mechanism* to enhance control over data. MPC makes it possible for data providers to “restrict what [data buyers] can do with the data” (A08) because it enables them to only share the computation results without having to release the input data (A01). The way MPC facilitates “a way how to share this data without actually sharing it” (A21) allows data providers to “preserve some information that they do not want to become public.” (A02).

> *“You control what [can be] done with the data . . . [I]n this case, you do not just offer access, but you only give away the insights you want to give . . . [W]hen I think about it again, there is a relevant improvement in what you call a control. And especially in terms of what is done with the data, that you have more control over that.”* (A01)

> *"We need [MPC technology] to make sure that you get your answers, but I am sure you will not be able to do anything else with it. Then you are getting your answers. . . . I can check with MPC what can be done and what cannot. . . . [I]t might also be a good possibility to not share my original data but share a data set with you, which looks like [it] but cannot be traced back to your actual data."* (A08)

Overall, findings suggest the relevance of control while sharing through data marketplaces, even though it requires a balance between usability and data misuse risk. Moreover, control in data sharing should be specified into three control sources: contract-based, structural-based, and technology-based control. Furthermore, technology-based control becomes more relevant with MPC since it enables control using a technical solution that facilitates sharing computation results without revealing the input data. Based on these findings, we refined the first proposition into:

P1a. Technology-based control is more relevant for data providers while sharing through data marketplaces that use MPC.

4.4.2. Trust in data sharing

Interviewees expressed that firms generally have little trust in this emerging approach for data sharing through data marketplaces (A08). In this regard, interviewees stressed the importance of establishing *trust in actors*, which is trust between actors involved in the value network of data marketplaces: data providers, data buyers, data marketplace operators, and end-users/consumers. In sharing data through data marketplaces, it is important for firms that act as data providers to establish trust with data buyers. This is because it is difficult for data providers to track the usage by data buyers once the data is shared (A06). Hence, firms would be more willing to share data with other firms that already have had a good relationship for a long time and have always been loyal to each other (A15). Moreover, given the involvement of third

parties as intermediaries, data providers also need to establish trust with data marketplace operators (A08, A11). One way to do that is by having a neutral third party as an operator. In this way, firms can ensure "everybody has an equal interest, and they are not acting in the interest of one company" (A16). Furthermore, since end-users or consumers (in our case, car owners or drivers) own the data generated in the car, they should have the final say on whether they are willing to give away their data or not (A18). Interviewees indicate that end-users have little to no degree of trust towards data providers (in our case, OEMs), saying that they either "[do] not want to share because they do not trust the OEMs" (A19) or "trusting OEMs that they do not do anything with my data … or gives me some disadvantage in any kind that you can consider" (A15).

Interestingly, other interviewees offered a different angle on trust by arguing that a lack of trust does not always hinder firms' willingness to share data (A04). This is because trust is often viewed as a secondary aim and should be discussed within the benefits and use cases of data sharing (A04). One interviewee even claimed that firms "do trust each other, but they are in a competition, and the competition matters" (A01). Therefore, decisions to share data through data marketplaces are driven more by strategic considerations than trust issues to gain economic benefits (A19).

> *"[I]f you are talking about business, I think in general there is little trust. So, I would never do anything with this type of data if I have the concerns I am talking about now based on trust."* (A07)

> *"[N]o one trusts anyone because right now [because] it is all about negotiation positions [that] you try to strengthen or weaken in the digital age. . . . [E]veryone tries to keep the data for themselves in the first place. … in this case, [trust] actually does not play a role."* (A15)

Regarding MPC, it is seen as a game-changer in the dynamics of trust in data marketplaces. MPC could facilitate collaboration between data providers and data buyers without fully trusting each other. In other words, MPC could potentially reduce the relevance of trust in actors in the context of data sharing through data marketplaces. This is possible since the input data is kept secure, and only the computation insights are provided. As a result, data providers could, in theory, allow data buyers to utilize the computation results while ensuring that they could not misuse the dataset beyond the data usage purpose. MPC could also become a novel approach to "ensure that the other partner cannot cheat the system" (A06), suggesting that trust in data marketplaces becomes less important.

> *"As you have more control over the data, you do not have to have the same level of trust in the other party. The more control you have of the data, the more you control how it can be used [and] the less you have to trust in the partner not cheating on you."* (A06)

> *"[D]oing [data sharing in data marketplaces] based on an MPC algorithm where the OEMs keep control of their data and not giving away their data to other parties gives them trust to actually collaborate because they do not have that much more to lose anymore."* (A15)

Strikingly, some interviewees argued that MPC is not just about reducing the relevance of inter-organizational trust but increasing the importance of *trust in technology*. The newness of MPC creates many questions on how it works, who the operator is, and its position in the whole data sharing process in data marketplaces. The lack of clarity makes people cautious and even skeptical about the impact of MPC on trust in data sharing through data marketplaces.

> *"I think this is highly connected to the trustworthiness of the MPC provider/operator. I think the whole thing works or does not work. But the*

question about who is the MPC is the trustworthy institution. And if that is the case, then that would add a lot to trust for the overall process. I think you need an intermediary to help there. I think you are not in a position to increase the trust among the actors by introducing a marketplace, but if you have a trustworthy process, then it could work." (A01)

"If the data marketplace or this MPC provider is not involved in the negotiation, I do not think the trust will not be provoked there. I believe they need to be involved because, again, I think big players will always fear or would always have trust issues with a party that they do not negotiate the terms of condition with." (A21)

Findings suggest that, while trust towards actors involved in data sharing is relevant, it is often a secondary aim that is embedded in the benefits and use cases of data sharing. Moreover, MPC reduces the need for trust in data sharing actors in collaborative data sharing, making it less relevant. Furthermore, MPC raises trust issues in its protocol due to a lack of accountability mechanisms, resulting in the relevance of trust in technology. Therefore, we evaluate and refine our second proposition into:

P2a. Trust in actors is less relevant for data providers while sharing through data marketplaces that use MPC.

P2b. Trust in technology is more relevant for data providers while sharing through data marketplaces that use MPC.

4.4.3. Perceived risks in data sharing

Interviewees confirmed that it is very risky for firms to participate in data sharing, especially in the emerging context of data marketplaces. *Competitiveness risks* are one of the biggest risks, in way that firms could lose an advantage over competitors

if they participate in data sharing via data marketplaces. OEMs in the automotive industry are typically reluctant to share because "they want to create a monopoly in the market" (A16). Opening up (access to) data by OEMs through data marketplaces could result in knowledge spillovers, allowing competitors to compare and develop better cars. This creates a situation where data "becomes a commodity" (A21), which is disadvantageous for OEMs because "their unique selling points are being taken away" (A06). As one interviewee put it:

> *"[T]hey are . . . not so willing to share data . . . [because of] a competitive position. If company X knows how many times the trunk will open and close and how strong you should make it, then . . . you got extra insights in product development. . . . [I]n the end, it has an impact on your competitive position because you are, for example, more efficient in producing a car because you can make it lighter or cheaper."* (A14)

Interviewees also expressed concerns regarding *data misuse risks*. If firms decided to participate in data sharing through data marketplaces, data buyers would use the data for other purposes that were not originally intended. Once the data is shared, it is difficult for data providers to check the purpose of data usage by data buyers (A08). This condition creates a risk in which data buyers could "use [the data] in any way that is harmful" (A01), such as security breaches in connected cars by hackers (A23).

> *"I can give you my data, but I cannot check if I say, "you can use this data for a certain sort of goal, which I would not like you to use for your commercial advantage or something." . . . Will they exploit [the data] to other means than I want to?"* (A08)

> *"We do not know how the third party would handle our data. . . . We [as an OEM] would like to share [data] with an automotive supplier through a data broker . . .*

but we do not know how the data marketplace would handle our information or data." (A21)

Moreover, most interviewees agree on the importance of *reputation risks* for data providers (in this case, OEMs), as they must maintain their brand and image to their end-users. In this regard, OEMs consider data sharing as a high-risk activity that could result in big protests if something terrible happens, like a privacy violation (A01) or a possibility that "others might find problems in [their] data" (A17). OEMs would like to save themselves from those troubles and prevent "negative or bad publicity when it comes to the usage of data" (A20). OEMs are also unwilling to pay huge fines since it is "damaging not only from the amount of profit that we receive per year but also from our reputation as a car company" (A13).

> *"For OEMs, especially premium OEMs, the brand is very important. So they cannot damage the brand of the car. You do not want to have a news headline that OEMs sold thousands of user data, and now you can track where you went with your car or something. That cannot happen."* (A12)

Furthermore, *end-user privacy risks* are highly relevant in data sharing via data marketplaces. Interviewees pointed out that the emergence of connected cars makes it possible for OEMs to collect data about how their car users behave every day (A19). This is not taken lightly by car users, resulting in a cautious approach by OEMs in data sharing to prevent breaching car users' privacy. The strict implementation of the General Data Protection Regulation (GDPR) also makes OEMs more aware of the importance of protecting customers' interests in safeguarding their personal data. Hence, OEMs are trying to "protect the benefits of our customers" (A18) before taking part in data sharing; otherwise, they risk getting fined for violating GDPR.

> *"[OEMs] care about privacy; they care about data providership. They question certain legal issues, whether they can share or not share data, which makes this data sharing so hard because it takes a lot of time and costs a lot."* (A08)

Additionally, interviewees mentioned that data interoperability between different OEMs might create data quality risk since the data are defined in different formats. As a result, OEMs need to make much effort in aggregating and harmonizing the data before it is usable for other parties (A12). OEMs are also afraid that poor data quality might lead to misinterpretation and misunderstanding of the data (A12, A17). One interviewee stated:

> *"If the data is wrong in the first place (garbage in), then a guy gets garbage out. But, if you do not know that, then the whole chain spent a lot of money without any use."* (A09)

As for the impact of MPC in addressing those risks, we found that MPC could reduce the relevance of competitiveness risk for data providers. MPC could restrict data usage by data buyers and only allow them to get answers from the computation. This mechanism could reduce the risk of knowledge spillover to competitors, preventing them from gaining a competitive advantage. Nevertheless, the reduced risk also comes with an increased burden for data providers and buyers, as they have to prepare better data that suits MPC requirements and clean the data themselves after acquiring it.

> *"I think it will change, but there is an increasing responsibility on the side of the data providers because they know what the requesting party wants to do with it. . . . [T]he overall risk is reduced, but the potential responsibility on the data [provider] side is increased . . ."* (A01)

> *"[W]hat they do not like is to provide the data to the competitors and allowing the competitors [to show] the advantage in using their [competitors' data] and not*

the other. If the data are not publicly available but only in an aggregated format, this usage of data from the competitor is not possible. So the risk does not exist." (A02)

However, the use of MPC could also increase the possibility of data misuse risk for data buyers. With MPC, data buyers need to ask queries (i.e., what kind of answers/insights do I want to know?), meaning that they need to reveal the process of analyzing the data (in the form of questions) through MPC-enabled data marketplaces. Such questions could allow data providers to reverse engineering and understand "the know-how" of data buyers, potentially leaking valuable information that data providers could misuse.

"Now, it is the other way around, I think. Because by specifying the aggregations, if I tell my supplier how to calculate and aggregate the values and how to assess the data, then I give away my own know-how. So now, the automotive suppliers need to at least tell the data marketplace and data providers on how the data must be aggregated. And this is sometimes already good know-how and the intellectual property of the data processors, the data analytics, and so on. So maybe now the OEMs who give away the data feel more confident, but now data buyers need to release a lot of their know-how because they need to specify how the data must be aggregated." (A11)

Findings suggest that, in sharing through data marketplaces, perceived risks should be specified into competitiveness risks, data misuse risks, end-users privacy risks, and reputation risks. As for the impact of MPC, it could make the competitiveness risks less relevant in data sharing by preventing knowledge spillover by restricting data usage. However, MPC could also shift the relevance of data misuse risks by potentially revealing firms' know-how and allowing reverse engineering through queries asked by data buyers. As such, data misuse risks could become less relevant for data providers

and more relevant for data buyers. Based on these findings, we refined the third proposition into:

P3a. Competitiveness risks are less relevant for data providers while sharing through data marketplaces that use MPC.

P3b. Data misuse risks are less relevant for data providers while sharing through data marketplaces that use MPC.

4.4.4. Boundary conditions for the impact of MPC on perceived control, trust, and perceived risks in data sharing

We also investigate boundary conditions under which MPC use in data marketplaces could impact data sharing antecedents by businesses. In this regard, interviewees pointed out three conditions that should be present: perceived benefits, organizational readiness, and perceived data sensitivity.

First, firms need to be sure of the *perceived benefits* in return for sharing data using MPC in data marketplaces. This is important as all business activities are about maximizing their profit by "creating value [and solving] customer problems" (A04). Examples of benefits mentioned by interviewees are personalization, service improvement, and direct monetization through data selling (A21). In the context of MPC, the main question is whether "some statistics [are] always enough, and how far can [firms] go [with MPC]" (A04). While MPC could be valuable for firms by generating insights from aggregated statistics, it might vary depending on domains and prospective users (A07). Therefore, firms would constantly assess if using MPC for data sharing is beneficial and valuable for their business.

> *"[I]f data privacy technology [like MPC] helps them to assure that they can monetize data better under the law, or with less risks in terms of data privacy concerns, then [firms] would be happy to employ that."* (A15)

> *"If [MPC] is really . . . cost-efficient, cost-saving, and quality increasing, then [firms] will go for it. But otherwise, they will leave it alone."* (A19)

Second, embedding MPC in data marketplaces might increase complexity for firms as they need sufficient *organizational readiness*, like data pre-processing skills such as cleaning and harmonization. This is due to the possible change in the role of data marketplace operators towards pure matchmaking. In this regard, the computation is done directly between firms (i.e., data providers and buyers). Hence, without proper data pre-processing skills, firms might be unwilling to share data in MPC-enabled data marketplaces as it is too costly and burdensome for them.

> *"[I]t looks to me that this technology shifts some value from the data marketplace provider and puts back some burden on the OEMs. If it is now more costly or complex for OEMs to provide data, then the technology adoption could be more difficult. . . . [T]he focus [of MPC] is so much on data anonymization, [while] the buyer still needs to do certain pre-processing to clean the data. . . . The fact that it seems to put more burden on the OEMs could worsen the willingness to share data if it is too costly."* (A12)

Third, the impact of MPC would depend on *perceived data sensitivity*. Most interviewees agreed that firms would only be willing to share generic and non-sensitive data through data marketplaces. In the automotive sector, OEMs will refrain from sharing data that are "relevant for the development of vehicles" (A11) or "if it comes close to competitive edges" (A14) since competitors might be able in the future to "decompose the way that the vehicle's system works and copy what an OEM has built if you give away the data" (A17). Hence, OEMs are trying to protect their data as much as possible and not share it with others. Embedding MPC into data marketplaces is unlikely to change this situation. As data providers, OEMs might hesitate to give away sensitive data, even though MPC could allow only the sharing of

computation results without revealing input data. They would consider MPC useful if the shared data are non-core, non-sensitive, and non-strategic.

> *"I think for data types for which the company does not see as core values for their own strategic interest in their own service development, it would make the idea of offering and transacting these data [using MPC in data marketplaces] would help with that."* (A01)

Taken together, the impact of MPC in data sharing would be apparent for firms when three boundary conditions are present. First, the benefits of MPC and its relevant use cases must be clear (i.e., perceived benefits). Second, firms must have a data-driven mindset and possess data pre-processing skills like data cleaning and harmonization (i.e., organizational readiness). Finally, due to the early adoption phase of MPC, firms will only share data that are considered non-sensitive and generic (i.e., perceived data sensitivity).

4.5. Discussion

When used in data marketplaces, MPC could change the relevance of perceived control, trust, and perceived risks as data sharing antecedents by businesses. These antecedents should be specified in new ways to understand firms' data-sharing decisions through MPC-based data marketplaces, namely *technology-based control*, *trust in technology*, and *data misuse risks*. Moreover, we found that trust in other actors involved and competitiveness risk, which was relevant in the current data sharing situations, are less relevant with MPC in place. Furthermore, we found three boundary conditions in which the impacts of MPC on these factors are relevant: (1) firms' perception of the benefits of using MPC, (2) firms' organizational readiness in terms of data skills and data awareness, and (3) firms' perception about the sensitivity of their data. Table 13 summarizes the changes from the initial propositions to the refined propositions derived from the results.

Theoretical concept	Initial proposition	Specified concept	Refined proposition
Perceived control over data	P1. Perceived control over data is more relevant for data providers while sharing through data marketplaces that use MPC	Technology-based control	P1a. Technology-based control is more relevant for data providers while sharing through data marketplaces that use MPC
Trust	P2. Trust is less relevant for data providers while sharing through data marketplaces that use MPC	Trust in actors	P2a. Trust in actors is less relevant for data providers while sharing through data marketplaces that use MPC
		Trust in technology	P2b. Trust in technology is more relevant for data providers while sharing through data marketplaces that use MPC
Perceived risks	P3. Perceived risks are less relevant for data providers while sharing through data marketplaces that use MPC	Competitiveness risks	P3a. Competitiveness risks are less relevant for data providers while sharing through data marketplaces that use MPC

		Data misuse risks	P3b. Data misuse risks are less relevant for data providers while sharing through data marketplaces that use MPC

Table 13 Refined propositions based on the findings

Regarding control over data, we identified contract-based control, structural based-control, and technology-based control as relevant control mechanisms in the context of data sharing. This finding extends existing knowledge of control theory in the IS literature, which typically emphasizes the object of control (input, process, output, and relations) (Tiwana, 2014; Wiener et al., 2016). We also find that MPC is seen as a technology-based mechanism that enables data providers to have more control over how their data is used (see P1a in Table 13). This finding is consistent with Garrido et al. (2022), who found that the MPC can enhance control over input data and the computation process while maintaining data utility. As a result, MPC guarantees that the data buyers will only receive the computation results and not the input data. This is important as firms mainly own sensitive and confidential data, making it necessary to ensure no leakage that might result in a competitive disadvantage or breach of end-user privacy. Furthermore, we provide support to the work of S. Spiekermann (2005), who argued that PETs could enhance control in the context of ubiquitous computing. We demonstrated that MPC, as one class of PETs, could also create a similar effect in a different setting, particularly in sharing data in automotive data marketplaces.

We also found that collaborative data sharing with MPC requires less trust between the actors involved since MPC enhances control over data during the computation. This finding means that, at least in theory, firms do not need to worry that their counterpart will get access to the input data since only the computation results will be

revealed to the requesting party. This finding also challenges the current understanding of trust, commitment, and reciprocal relationships with other firms as preconditions for data sharing (Zaheer & Trkman, 2017). However, an interesting observation is that while MPC could reduce the need for trust in actors involved in data sharing, it could raise trust issues concerning the underlying algorithms that execute the protocol. In this regard, we argue that MPC could change the way we conceptualize trust in data sharing. Traditionally, scholars see the trust between actors (i.e., inter-organizational trust) as a key aspect influencing data-sharing decisions (Müller et al., 2020). Now, in line with Lumineau et al. (2023), trust in a system based on digital technologies is increasingly relevant, especially for emerging technologies that run in the background, like MPC and blockchain. In the context of MPC, trust in technology becomes relevant because it takes over the process of enforcing control for data providers. Hence, the new conceptualization of trust should be considered when studying MPC and its implications for data sharing and collaboration (see P2b in Table 13).

A surprising finding is that while MPC reduces risks perceived by data providers, specific data sharing risks remain. For instance, the risks of revealing sensitive information might shift to data buyers. With MPC, data buyers become vulnerable because their queries could reveal insights they want to obtain, allowing data providers to guess the strategic interests of data users. This implies that MPC features are like a double-edged sword that eliminates the risk for data providers while creating risks for data buyers. We argued that this new risk might be due to the context in which MPC is implemented. The currently known MPC use cases mainly aim to address societal problems, such as financial fraud detection (Sangers et al., 2019) and healthcare predictions (van Egmond et al., 2021). In those use cases, all parties agree on data usage purposes to perform computational analysis and generate insights using MPC. Hence, the risk of revealing sensitive information by data buyers is eliminated. However, our research context of data marketplaces differs because it

involves buyer-seller relationships with unknown participants, unclear data usage purposes, and business motives. In this regard, while MPC lowers the risks for data providers, it creates new risks for data buyers in revealing sensitive information while sharing through data marketplaces. In other words, the risks of revealing sensitive information are more significant in this use case compared to existing MPC use cases. Therefore, we argue that the shift in data sharing risks should also be taken into account when investigating the implications of MPC in data sharing and collaboration (see P3b in Table 13).

We find that the impact of MPC on perceived control, trust, and perceived risks in data sharing depends on three conditions. The first condition is perceived benefits, which refers to how firms understand and appreciate the benefits of using MPC. In this regard, firms must be sufficiently informed on how MPC use cases are relevant and in line with their business activities (Kanger & Pruulmann-Vengerfeldt, 2015). The importance of perceived benefits as preconditions for data sharing is consistent with existing literature (Fu et al., 2014; Sun et al., 2018). Nevertheless, our findings show that in the early stage of MPC development and adoption, the benefits of MPC in enabling data sharing while keeping the input data private seem not highly compelling for firms. The second condition is organizational readiness. Firms must be willing to shift towards data-driven mindsets and develop data analytics skills (Svensson & Taghavianfar, 2020). Otherwise, firms will face challenges in realizing the business value of MPC in data sharing. This is important as MPC is a complex technology and might only create value for firms that are knowledgeable and aware of its potential (Zöll et al., 2021). The final condition is perceived data sensitivity. Firms must deal with highly sensitive data before considering using MPC in data sharing. However, as we found, firms will only share generic and non-sensitive data even with MPC in place, implying that data sensitivity is a necessary but insufficient condition to use MPC.

An explanation for firms' low perception of benefits and reluctance to share highly sensitive data might be due to the nature of the automotive industry, which was

chosen as our research context. This industry is known to be (1) conventional when dealing with sensitive data, (2) have low trust between actors, and (3) strongly afraid of losing a competitive edge (Svahn et al., 2017). Therefore, despite a trend toward data-driven organizations, actors in the automotive industry still perceive data sharing as a high-risk business activity. Moreover, MPC is still a relatively new technology with a lack of proven use cases in the automotive industry. As a result, actors in the automotive industry are not yet convinced of the business value of MPC. Furthermore, the added setting of data marketplaces also increases risk since it involves selling (access to) car data to other parties without knowing the purpose of data usage.

4.6. Limitations

Our work presented in this chapter has four main limitations. First, real-life and large-scale implementations of MPC are currently limited and even more lacking in the context of data marketplaces. As a result, we rely on a thought experiment on a possible scenario of MPC-enabled automotive data marketplaces rather than actual real-life implementation. In this regard, the generalizability of our findings is limited and might only be relevant in the early stage of MPC adoption. Second, we used a short presentation to explain what MPC is and how it can be used for data sharing in data marketplaces. This would lead to a potential bias in the interviews since interviewees might base their understanding of MPC mostly on our explanations, which should be considered when interpreting our findings. Third, as described in Chapter 1 and 2, we refer to data marketplaces as platforms for buying and selling datasets between firms, which was also explained in the short presentation. However, we observed that interviewees sometimes based their answers on (1) data-sharing platforms that purely focus on facilitating data exchange between partners or (2) a general view of data sharing without considering intermediaries like data marketplaces. This limitation is expected since interviewees are not very familiar with data marketplaces due to the diversity of data marketplaces' business models (Bergman et al., 2022; Fruhwirth, Rachinger, et al., 2020; van de Ven et al., 2021).

Therefore, we kept an eye on this issue during the interviews and clarified the concepts to the interviewees when this issue arose. Fourth, the findings in our study were derived in the context of automotive data marketplaces, which is a setting with high data sharing hurdles (i.e., high fear of losing data control, low trust between actors, and high risks). Given that the magnitude of data sharing hurdles is important in assessing the impact of MPC, different findings might emerge in settings with a lower magnitude of hurdles.

4.7. Conclusions

In this chapter, we investigate the impact that MPC could have on antecedents of data sharing by businesses in data marketplaces. Our findings showed that implementing MPC in automotive data marketplaces could increase the relevance of **technology-based control**, **reduce the need for trust in other organizations,** and **reduce the relevance of competitiveness risks** by preventing the leakage of datasets. However, MPC changes the relevance of trust as now companies **need to trust the technology** instead of other companies participating in the computation. At the same time, MPC also introduces **a new risk in the form of possible misuse of data insights**. Nevertheless, MPC could only impact antecedents of business data sharing if companies (1) perceive clear benefits from sharing data and using MPC for data sharing, (2) have sufficient organizational readiness regarding data-related capabilities, and (3) are aware of the sensitivity of their business data.

From this chapter, we provide answers to the second research question and obtain an essential understanding of the business perspective on MPC and data sharing. However, these findings only tell parts of the story since MPC could also impact consumer data sharing decisions, in which privacy concerns become an important antecedent. This is also in line with our observation in Chapter 2 that MPC could also impact privacy concerns. Hence, in the next chapter, we emphasize the consumer

perspective to investigate the potential impact of MPC implementation in data marketplaces on these antecedents.

5 Understanding consumer perspective on MPC and data sharing: a quantitative study[3]

In this chapter, we answer the third research question: *what could be the impact of MPC use in data marketplaces on antecedents of data sharing by consumers?* We found in Chapter 3 that control, privacy, trust, and risks are data sharing antecedents that could be impacted by MPC use in data marketplaces. Then, focusing on a business perspective, we found in Chapter 4 that MPC use in data marketplaces could impact antecedents of business data sharing by increasing the relevance of technology-based control, trust in technology, and risks of data insight misuse. Nevertheless, as described in Section 1.1, sharing in data marketplaces not only benefits business actors in creating value by combining multiple data sources but is also beneficial to consumers in allowing them to monetize their personal data. In this regard, our findings in Chapter 4 did not capture the view of individual consumers, especially on privacy concerns as one important data sharing antecedent. Thus, we need to complement insights from businesses with consumers' perspectives regarding the potential impact of MPC use in data marketplaces on antecedents of consumer data sharing. By integrating multiple perspectives from MPC developers,

---

[3] This chapter is based on: (1) Agahari, W., & de Reuver, M. (2022). Rethinking consumers' data sharing decisions with the emergence of multi-party computation: an experimental design for evaluation. In *Proceedings of the 30th European Conference on Information Systems (ECIS 2022)*, Timisoara, Romania.; and (2) Agahari, W., Fiebig, T., & de Reuver, M. *Does multi-party computation impact consumers' willingness to share data? An experimental study.* [Manuscript in preparation].

businesses, and consumers, we can comprehensively understand how MPC could change the dynamics of data sharing in data marketplaces.

Because we need a theoretical understanding of why privacy concerns can influence consumers' data sharing decisions, we first elaborate on information privacy literature, which also covers other data sharing antecedents as identified in Chapter 3. Based on the literature analysis, we develop a conceptual foundation on privacy, control, trust, and risks as antecedents of consumer data sharing in data marketplaces. Then, drawing from MPC characteristics identified in Chapter 2, we specify the conceptual foundation of consumer data sharing antecedents to the MPC domain, resulting in the proposed hypotheses for this chapter. Subsequently, we describe our approach to conducting an online experiment, including the design, survey questions, procedures, and selection of participants. After that, we present and discuss the results of our analyses and hypotheses testing. We conclude this chapter by outlining the limitations of this chapter and providing an answer to the third research question.

5.1. Consumer perspectives on data sharing: a review on information privacy literature

Information privacy refers to the ability of individuals to control when, how, and to what extent information about them is shared with others (Malhotra et al., 2004; Popovič et al., 2017; Westin, 1968). Scholars working in this literature stream have been attempting to explain consumers' information disclosure decisions in various contexts, such as e-commerce (Dinev & Hart, 2006), social media (Hajli & Lin, 2016), e-health (Dinev et al., 2016), IoT services (Bélanger et al., 2021), and mobility (Kehr et al., 2015). This resulted in various antecedents of consumers' data sharing decisions, which we summarize in Table 14.

One of the key factors in consumers' decision to share data is the notion of control over data (Dinev et al., 2013; S. Spiekermann, 2005), as it is closely related to the definition of information privacy. *Perceived control* refers to the extent to which an individual believes that he/she can manage the release and dissemination of personal information (Xu et al., 2011). Prior studies have suggested that perceived (lack of) control might cause consumers to refrain from sharing data (Cichy et al., 2021; Kehr et al., 2015; Xu et al., 2011). In the context of social media, consumers might even try to falsify the information they provide on social media as a means to retain control (Alashoor et al., 2017). In this regard, it is important for consumers to have more ability to control their data via various means like anonymization (Harborth et al., 2020) and control functionality (Kato et al., 2016) before deciding to share their data (Farrelly & Chew, 2016; Markos et al., 2017).

Privacy concerns are another central antecedent of individual willingness to share data. It is defined as the degree of an individual's concern about who has access to the data that is being shared and how other parties use it (Smith et al., 1996, 2011). Privacy concerns in the digital sphere might emerge from collecting personal information, unauthorized secondary use, improper access, and errors (Hsu & Lin, 2016; Smith et al., 1996). In this regard, individuals who are deeply concerned about their information privacy will likely refrain from data sharing and demand more privacy protection. This is evident in prior studies within the context of location-based services (Keith et al., 2013; L. Zhao et al., 2012), connected cars (Cichy et al., 2021), and privacy-preserving data markets (Schomakers et al., 2020).

Information privacy literature also highlights the vital role of trust as a prerequisite for data sharing (Pavlou, 2003; Richter & Slowinski, 2019; M. Spiekermann, 2019). Following Kehr et al. (2015), we define trust as an individual's belief that another party will act as expected and not do harmful things, such as misusing personal data. In the context of data marketplaces, there are two important aspects of trust from the perspective of consumers as data providers: *trust in data buyers* and *trust in data*

marketplace operators. In this regard, greater trust would make consumers more willing to share their data (Kehr et al., 2015; Pal et al., 2021).

Further, trust is often associated with perceived risks (Hart & Saunders, 1997), in which a higher degree of trust reduces perceived risk in data sharing (Dinev & Hart, 2006; Pavlou, 2003). Building on Xu et al. (2011), *perceived risks* refer to the expectation of losses if someone engages in data sharing. Hence, if people think that sharing their data is a risky thing to do and could cause harm to them, they will refrain from data sharing (Wang et al., 2016). The importance of trust and perceived risks in consumers' data sharing decisions have been outlined in prior research, not only in the mobility-related context, such as connected cars (Buck & Reith, 2020; Cichy et al., 2021) and location-based services (Keith et al., 2013; Zhou, 2011) but also in other domains like e-commerce (Dinev & Hart, 2006; Malhotra et al., 2004; Pavlou, 2003; Xu et al., 2011), social media (Alashoor et al., 2017; Krasnova et al., 2010), and IoT (Kim & Choi, 2022; Pal et al., 2021).

Antecedents	**Description**	**Relevant studies**
Perceived control	the extent to which an individual believes that he/she can manage the release and dissemination of personal information (Xu et al., 2011)	Dinev et al. (2013), Farrelly and Chew (2016), Kim and Choi (2022), Krasnova et al. (2010), Markos et al. (2017), S. Spiekermann (2005), Schomakers et al. (2020), Xu et al. (2011)
Privacy concerns	the degree of an individual's concern about who has access to the data that is being shared and how other	Cichy et al. (2021), Kato et al. (2016), Kehr et al. (2015), Keith et al. (2013), Malhotra et al. (2004), Mangiò et al. (2020), Naous et al. (2019), Pal et al.

	parties use it (Smith et al., 1996)	(2021), Smith et al. (2011), L. Zhao et al. (2012)
Trust (in data buyers and in data marketplace operators)	an individual's belief that another party (data buyers and/or data marketplace operators) will act as expected and not do harmful things, such as misusing personal data (Kehr et al., 2015)	Buck and Reith (2020), Dinev and Hart (2006), Kehr et al. (2015), Krasnova et al. (2010), Liu et al. (2005), Malhotra et al. (2004), Pavlou (2003), Wessels et al. (2019)
Perceived risks	the expectation of losses if someone engages in data sharing (Xu et al., 2011)	Alashoor et al. (2017), Cichy et al. (2021), Dinev and Hart (2006), Kehr et al. (2015), Kim and Choi (2022), Krasnova et al. (2010), Malhotra et al. (2004), Pal et al. (2021), Pavlou (2003), Zhou (2011)

Table 14 Antecedents of consumers' willingness to share data derived from the information privacy literature

5.2. Research hypotheses

Next, we develop hypotheses by contextualizing perceived control, privacy concerns, trust, and perceived risks as antecedents of consumer data sharing to the MPC domain. Prior studies suggest that PETs could address privacy concerns and allow consumers to have greater control in protecting themselves from unauthorized access and use of their data (Cichy et al., 2021; Ram et al., 2018; S. Spiekermann, 2005). Hence, consumers would be more willing to share their data with companies implementing PETs (Kato et al., 2016). Metzger (2007) found that users of e-commerce platforms that incorporate privacy protections are more willing to disclose

their data to such platforms. Meanwhile, a study by Zhou (2017) outlined the importance of privacy control in reducing privacy concerns in location-based services. Similarly, Li and Slee (2014) also found that having the ability to control personal information could reduce privacy concerns and increase patients' willingness to digitize their personal health records. In line with this, Cichy et al. (2021) found that, with the presence of privacy protection, connected car users are less concerned about their privacy in sharing their driving data. Furthermore, Schomakers et al. (2020) revealed that users' willingness to share data is significantly impacted by the presence of anonymization in the privacy-preserving personal data market.

In line with these studies, we expect that MPC could have a similar effect in increasing consumers' perception of control over data and lowering privacy concerns compared to conventional solutions like TTP. With MPC, data buyers will only receive insights from the computational analysis between multiple data providers. In this way, consumers have more control over how data buyers utilize the data. Therefore, we hypothesize that:

H1. Consumers' perceived control while sharing through data marketplaces that use MPC is higher than TTP.

H2. Consumers' privacy concerns while sharing through data marketplaces that use MPC are lower than TTP.

PETs could also increase trust toward actors involved in data sharing (Smith et al., 2011). A study by Krasnova et al. (2010) found that social media users have more trust in social media providers and other users when control measures are present. Similarly, Naous et al. (2019) revealed that with privacy control in place, users of location-based services would have higher trust in service providers and other users. Some use cases of MPC have also been shown to have the potential to alleviate issues of trust in other actors. For instance, by using MPC-based analytics solutions,

companies in Boston became more cooperative and trusted third parties that analyzed the gender wage gap data (Lapets et al., 2018). In another example, Bogetoft et al. (2009) outlined that, thanks to the MPC solution, farmers have more trust in the buyer company during an auction in determining the market clearing price of sugar beets in Denmark.

Based on the above elaboration, we also expect that implementing MPC in data marketplaces could lead to higher trust toward data marketplace operators and data buyers. The data marketplace operator will not be able to see the data during the computation, as it was done directly between consumers (as data providers) and data buyers. Meanwhile, data buyers can only access computation results, not input data. Therefore, in theory, it is not possible to misuse individual consumers' data. Thus, we propose the following hypotheses:

H3. Consumers' trust in operator while sharing through data marketplaces that use MPC is higher than TTP.

H4. Consumers' trust in data buyers while sharing through data marketplaces that use MPC is higher than TTP.

Related to its ability to enhance the feeling of control over data, PETs could also lower risk perceptions in data sharing (Ziefle et al., 2016). In the context of social media, the presence of privacy control features plays an essential role in reducing risk perceptions of their users (Hajli & Lin, 2016). Moreover, a study by Kim and Choi (2022) found that various privacy control options in smart home services also decrease users' perceived risks and the sensitivity of shared information. Furthermore, implementing technology privacy mechanisms for sharing electronic health record data positively affects perceived control, ultimately reducing data sharing risks (Dinev et al., 2016).

Taken together, we expect consumers to perceive lower risks while sharing through data marketplaces that use MPC than TTP. Since data buyers will not receive individual consumers' data, the risks for consumers to take part in data sharing might be lower. Hence, we hypothesize that:

H5. Consumers' perceived risks while sharing through data marketplaces that use MPC is lower than TTP.

Finally, existing research has shown that the presence of PETs can increase consumers' willingness to share data. A study by Kato et al. (2016) has shown that by giving consent and controlling what kind of data can be accessed by third parties, consumers are more willing to share their data. In line with this, Schomakers et al. (2020) also found that users will demand more privacy protection, such as anonymization, before deciding to share data. Hence, with all the features of MPC, we expect that consumers are more willing to share data through data marketplaces that use MPC than TTP. Thus, we propose the following hypotheses:

H6. Consumers' willingness to share through data marketplaces that use MPC is higher than TTP.

5.3. Methodology

We describe the approach used in the quantitative study, including the experimental design, measures, procedures, and participant selection.

5.3.1. Experimental design

We conducted a controlled, survey-based online experiment, which is suitable for our study as we aim to explore the effect of a new instrument (i.e., MPC) in changing the current situations (i.e., consumers' data sharing decisions). Achieving this goal

requires a high internal validity, which is the main advantage of experimental research (Verschuren & Doorewaard, 2010). Specifically, we used a between-subject post-test-only design, meaning that each participant was randomly assigned to one treatment only and got an identical post-test after the treatment (Charness et al., 2012). We opted for this design to minimize order effect bias while participants compared different treatments in a within-subject design. By assigning participants to one treatment only, other treatments will not influence their answers, and they cannot do hypothesis guessing (Campbell & Stanley, 2015).

We conducted an experiment with three treatments representing three data sharing scenarios in data marketplaces (see Figure 7). Treatment 1 examined existing data marketplaces that used a Trusted Third Party (TTP), an intermediary that facilitates data exchange between buyer and seller. Meanwhile, Treatment 2 examined data marketplaces described as using MPC protocol to compute data from different data providers without giving away the underlying data. However, since we used a description of MPC and not a working demonstrator or prototype, participants could attribute value to the term MPC rather than to the underlying ideas in the technology. Therefore, we also added Treatment 3, which examined data marketplaces that are described as using a fictitious technology called Data-Computation-Protection (DCP). This way, we can see if different technologies would make any differences in perception or do not matter to participants, even if the technology does not exist.

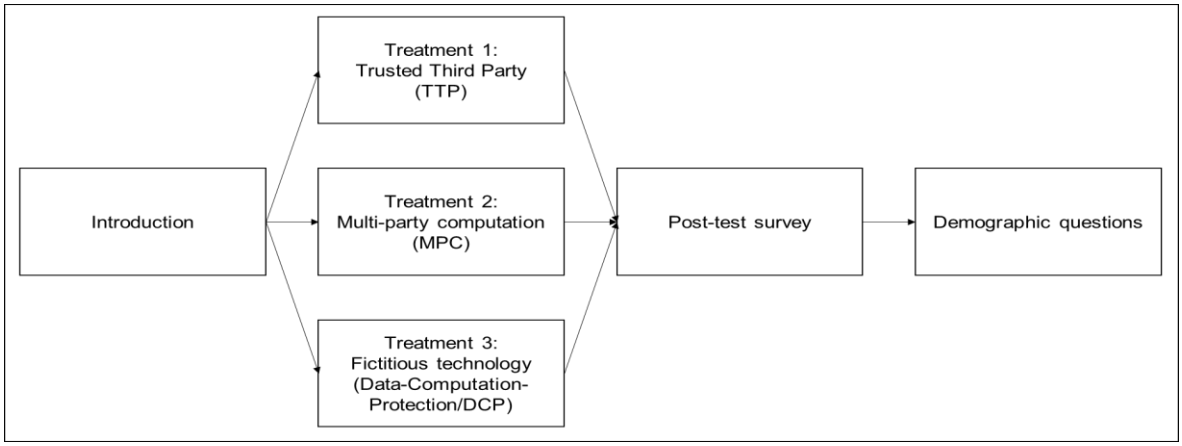


Figure 7 Experimental design

As a treatment, we visualized three mock-ups of data marketplaces based on the work of Faujdar (2019), Petronia (2020), and van der Wel (2021), who lay a basis for the components of data marketplaces that we need to visualize in the mock-up. Given that those works focus on revenue management, supply chain, and healthcare domains, we adapted the mock-ups to fit our context of sharing driving data through personal data marketplaces. We do this by taking inspiration from Kehr et al. (2015) and adjusting data types that consumers could sell through data marketplaces, namely “trip date and time”, “origin and destination”, “routes”, and “speed”. These data types and desired compensations were already pre-filled with a deliberately low amount of money. We do this so that participants do not focus on the amount of compensation they can receive. Instead, we emphasized that participants can receive compensation for sharing their driving data. Figure 8 shows the three mock-ups of data marketplaces that we used in the experiment.

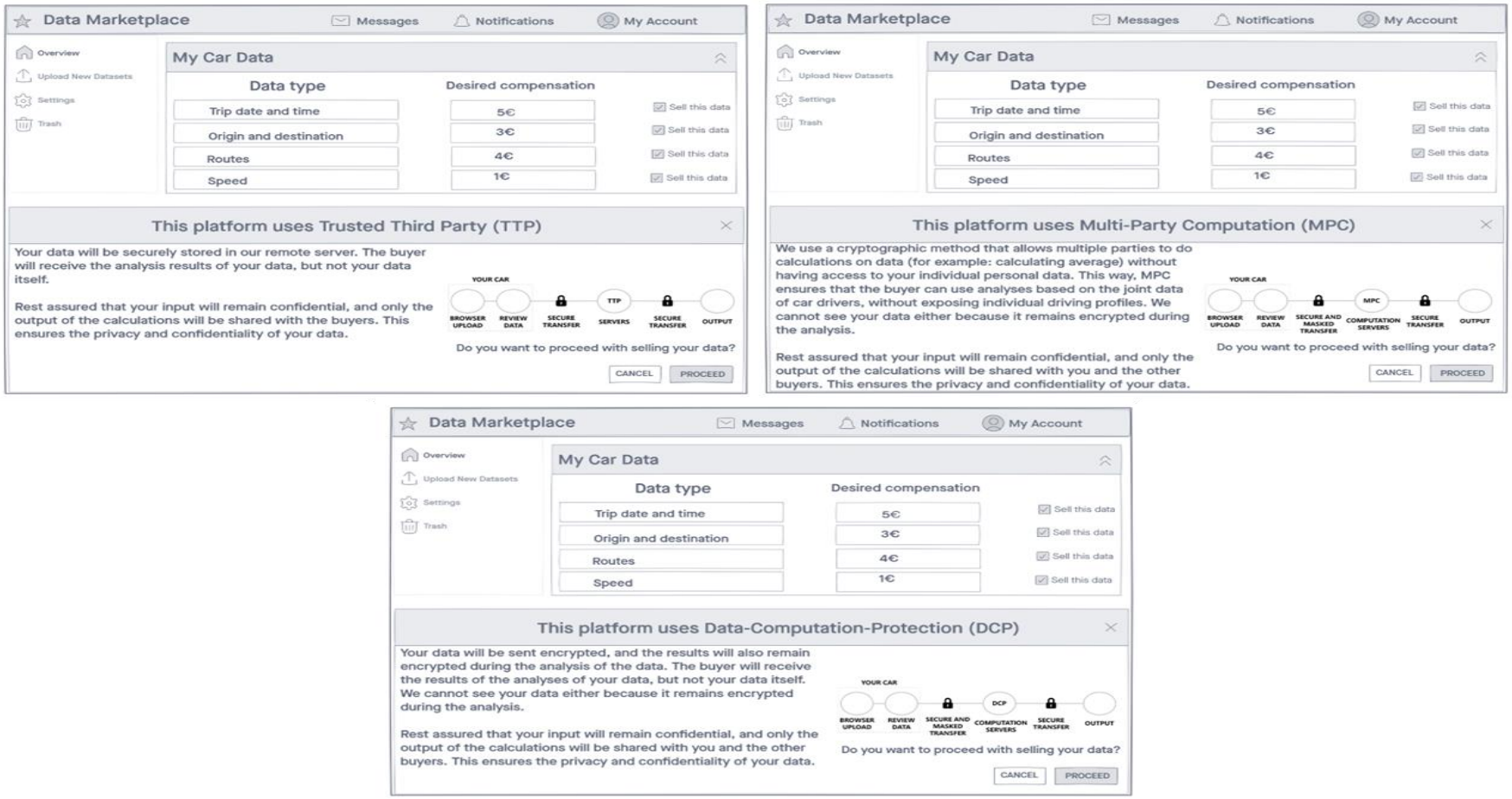


Figure 8 The mock-ups of data marketplaces for TTP (top-left), MPC (top-right), and DCP (bottom) treatments

The three mock-ups differ in the bottom part, in which we illustrated the underlying technology as our treatment. In doing so, we synthesized the work of Faujdar (2019), Petronia (2020), and van der Wel (2021) and combined them with insights from our literature study (Chapter 2) as well as interviews with MPC experts (Chapter 3) and business actors (Chapter 4). Integrating these insights resulted in the visualization consisting of the title, the description, and a simplified process illustration, as shown in the bottom part of the mock-ups in Figure 8. For TTP-based data marketplaces (Treatment 1), we described that consumers send data to a central system where the data will be analyzed and stored. For MPC-based data marketplaces (Treatment 2), we described how MPC works, in which the data is encrypted in the car, and only the analysis results are revealed to the prospective buyers. Meanwhile, for DCP-based data marketplaces (Treatment 3), the description is identical to Treatment 2, with the name of the technology being the only difference. This way, we can see whether the different terminology or the description of the technology matters for participants, even if it is a fictitious technology.

5.3.2. Measures

We adapted the scales from previous studies in the information privacy literature and specified them to the novel context of sharing driving data via data marketplaces. We modified existing measures from Xu et al. (2011) to measure perceived control and perceived risks, while measures by Dinev and Hart (2006) were modified to measure participants' privacy concerns. To measure trust, we modified measures by Kehr et al. (2015) and made distinctions between trust in data buyers and trust in data marketplace operators (see Section 5.1). Finally, we adapted measures by Pavlou (2003) to measure participants' willingness to share data given the scenario of privacy-enhancing data marketplaces presented to them.

As control variables, we used Westin's Privacy Segmentation Index (Kumaraguru & Cranor, 2005) to classify participants' privacy attitudes into one of the three groups:

fundamentalists (i.e., most protective of their privacy), unconcerned (i.e., least protective of their privacy), and pragmatists (i.e., weighing the pros and cons of sharing information). We also used other demographic characteristics as control variables, like industry type, car ownership, awareness of data marketplaces, and awareness of PETs.

We validated our measures and the three mock-ups in two rounds of pre-tests. In the first pre-test, we recruited six researchers as participants (5 male and 1 female) to check the content validity of the constructs. We also checked whether the description of how the technology works is understandable. Based on their feedback, we refined the questions, case descriptions, and experiment flow. In the second pre-test, we recruited 300 participants (165 male, 126 female, and 9 others/prefer not to say) from an online crowdsourcing platform Prolific. We used Prolific to recruit pre-test participants since the same platform is used to recruit participants for the main study. Thus, we would like to validate further our measures and the description of our treatments with participants with similar characteristics to our target participants for the main study. For this second pre-test, our population comprises consumers with a driving license. We restricted the sample to participants 18 years old and older, as this is the minimum age to have a driving license in most countries. Also, since our sample for the main study is participants from the United Kingdom (e.g., participants who have a UK nationality or currently live in the UK), we excluded this group of participants to ensure they did not participate in the same study twice, increasing the reliability of the answers for the main study. We offered financial compensation (3.75 GBP) to participants based on the recommendation provided by Prolific.

We collected the data for the second pre-test on 9 September 2021. The average age of participants was 30.1 years old (SD = 8.87), and about 70.7% are part of the younger generation (18-34 years old). Most of them reside in the United States (53.3%), France (20.7%), and South Africa (6.7%). The majority had already finished a graduate degree (35%), followed by an undergraduate degree (30.3%) and high school diploma/A-level

education (18%). More than half of the participants currently work full-time (56%) or part-time (13.7%) and primarily work in the IT (19.7%) or finance industry (7.3%). About one-third of our participants hold a managerial position, either at a junior (5.7%), middle (18%), or upper management level (9.3%). Regarding access to and ownership of cars, only 10% of participants did not have access at all. The rest either own a car (63.7%), have access via family members (22.3%), or have access via leasing or rental (4%). Further, 53.4% of participants claimed they were familiar with data marketplaces, while only 23% of participants had prior knowledge about privacy-enhancing technologies before participating in the survey. The underlying datasets for the second pre-test are accessible at the 4TU research data repository (See Appendix D).

Based on the data collected in this second pre-test, we conducted a Confirmatory Factor Analysis (CFA) using JASP 0.16.1 to validate our constructs and measurement model (Brown & Moore, 2012). Using three criteria by Hu and Bentler (1999) (Comparative Fit Index/CFI ≥ 0.95, Tucker-Lewis Index/TLI ≥ 0.95, Root Mean Square Error of Approximation/RMSEA ≤ 0.06), we found a good fit of the model (CFI = 0.99, TLI = 0.98, RMSEA = 0.04). We then removed two survey items with factor loadings lower than the cut-off values of 0.7 (Fornell & Larcker, 1981). To check for cross-loadings, we removed another two items with modification indices higher than ten and one item that cross-loaded on all constructs. Moreover, we established internal reliability (CR > 0.7), convergent validity (AVE > 0.5), and discriminant validity (inter-construct correlation < √AVE) (Fornell & Larcker, 1981). Furthermore, we conducted Multi-Group Confirmatory Factor Analysis (MGCFA) and found that the three treatment conditions also satisfy all criteria. Our final measures for the main study comprise six constructs and 16 items (see Table 15).

Construct	**Items**	**Item wording***	**Adapted from**
Perceived control	CTRL_1	I believe I have control over who can access the	Xu et al. (2011)

		sensitive data I provided to this data marketplace.	
	CTRL_2	I think I have control over what kind of sensitive data is shared by this data marketplace to other companies.	
	CTRL_3	I believe I have control over how other companies use the sensitive data I provided to this data marketplace.	
Privacy concerns	PRIV_1	I am concerned that other parties could find sensitive information about me on this data marketplace.	Dinev and Hart (2006)
	PRIV_2	I am concerned about providing my sensitive data to this data marketplace because of what other parties might do with it.	
Trust in data marketplace operators	TRSD_1	I expect this data marketplace would be trustworthy regarding my sensitive data.	Kehr et al. (2015)

	TRSD_2	This data marketplace would tell the truth and fulfil promises related to my sensitive data.	
	TRSD_3	I expect this data marketplace would be honest with me regarding the sensitive data I would provide.	
Trust in data buyers	TRSB_1	I expect that data buyers would be trustworthy in handling the data they got from this data marketplace.	Kehr et al. (2015)
	TRSB_2	I expect that data buyers would tell the truth and fulfill promises in handling the data they got from this data marketplace.	
	TRSB_3	I expect that data buyers would be honest when handling the data they got from this data marketplace.	
Perceived risk	RISK_1	I find it risky to provide my sensitive data via this data marketplace.	Xu et al. (2011)

	RISK_2	There would be too much uncertainty associated with providing my sensitive data to this data marketplace.	
Willingness to share data via data marketplaces	WTSD_1	Given the chance, I would share my data via this data marketplace.	Pavlou (2003)
	WTSD_2	Given the chance, I predict that I should share my data via this data marketplace in the future.	
	WTSD_3	It is likely that I will share my data via this data marketplace in the near future.	

*5-point Likert scale, 1 (strongly disagree) to 5 (strongly agree)

Table 15 Survey items

5.3.3. Procedures

At the start of the experiment, we introduced participants to the purpose of the study and its approximate completion duration. Then, we provided an informed consent form approved by the university's Human Research Ethics Committee (HREC). In this informed consent form, we explained that the datasets would be stored in protected storage and only accessible to the research team. We also informed participants that anonymized datasets would be publicly available in an open research data repository.

Next, we presented participants with a case description about sharing driving data from connected cars via data marketplaces. We asked participants to imagine owning a connected car that generates driving data and could sell it via data marketplaces. Then, mobility service providers are interested in buying participants' driving data (e.g., trip date and time, destination and route history, and driving speed) via data marketplaces. In return, these mobility service providers can offer innovative products and services based on driving data they bought via data marketplaces. Furthermore, despite the importance of privacy calculus in consumers' data sharing decisions, we expected that MPC only impacts perceived risks and not the perceived benefits. Therefore, in the case description, we controlled the perceived benefits by asking participants to assume they would always get benefits through better driving advice and financial compensation. We provided a simplified illustration to help participants understand better the process of sharing driving data through personal data marketplaces, as seen in Figure 9.

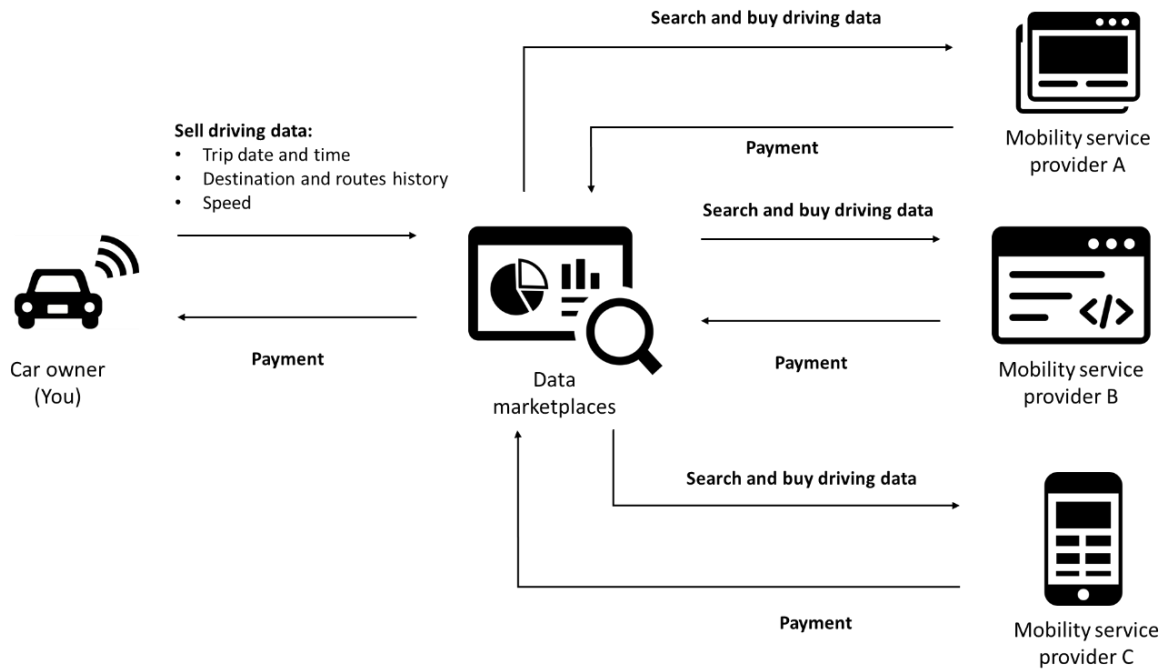


Figure 9 Illustration of sharing data via data marketplaces

After the case description, we randomly assigned participants to one of the three experimental conditions (TTP, MPC, or DCP). In each condition, we explained how the technology works with an illustrative example of sharing salary information among four colleagues to determine who should pay for dinner. Then, we further explained how the technology works in data marketplaces and showed the mock-up. Finally, we asked participants to enter the scenario code used to double-check the equal distribution of the three scenarios. At the end of the experiment, we asked participants in all treatments to answer identical survey questions about their perceptions of data marketplaces presented to them (see Table 15). It is important to note that we did not assign all three treatments to participants and asked them to compare the three treatments before answering the survey. Instead, we only asked them to answer the survey based on one treatment they received. Then, we asked three questions about the privacy attitude based on Westin's Privacy Segmentation Index. We conclude the experiment with demographic questions.

It is essential to stress that, given that we employ a between-subject experiment design, we did not assign all three treatments to participants and asked them to compare the three treatments before answering the survey. Instead, what we did was

For the privacy attitude questions based on Westin's Privacy Segmentation Index, participants had to give opinions about (1) consumers' control over data collected by companies, (2) how companies handle consumers' data, and (3) the sufficiency of regulations. We grouped participants into three privacy attitudes based on their answers (1 = strongly disagree, 4 = strongly agree). Participants who agreed (strongly or somewhat) with the first statement and disagreed (strongly or somewhat) with the second and third statements were classified as fundamentalists, while those who answered with the opposite were considered unconcerned. The remaining participants were classified as pragmatists.

We employed two instructional manipulation checks (Oppenheimer et al., 2009) to ensure that our participants answered thoughtfully and paid attention to our instruction. Specifically, we randomly placed the following two questions in the survey, in which we instructed participants to select particular answers regardless of their opinions.

(1) There is nothing wrong with companies that collect personal information without consent. Regardless of what you think, please select "somewhat agree."

(2) Do you agree that data is the new oil? Regardless of what you think, please select "strongly disagree."

The complete procedure of the experiment is provided in Appendix C. On average, participants spent 9.2 (SD = 5.4) minutes on the experiment, which is faster than the average time spent during the pre-study (average 15.6 minutes, SD = 9.3). We offered financial compensation to participants by following the minimum hourly rate recommended by Prolific (7.5 GBP/hour). Then, we used the average time spent by participants in the pre-study to estimate the completion time, which we rounded up to 20 minutes. Hence, each participant was paid 2.5 GBP for participating in the experiment, regardless of the actual completion time.

5.3.4. Participants

We recruited participants using an online crowdsourcing platform Prolific, which is commonly used in academic research nowadays (Palan & Schitter, 2018; Peer et al., 2017). Participants recruited via Prolific are more diverse, naïve, and honest than similar crowdsourcing platforms like Amazon Mechanical Turk (MTurk) (Adams et al., 2020; Peer et al., 2017). Also, Prolific claims to be able to offer representative samples based on age, gender, and ethnicity (Prolific, 2022). Nevertheless, using Prolific for academic research also has limitations like participant selection bias and monetary

incentives, which should be considered when interpreting the results (Kaufmann et al., 2011).

Our population of interest comprises citizens of the United Kingdom (i.e., those with UK nationality or currently living in the UK) aged 18 years and older. We selected the UK population so that we can leverage the representative samples feature offered by Prolific and ultimately can generalize our findings for the UK population based on age, gender, and ethnicity (Prolific, 2022). We collected the data on 15 November 2021, and we recruited 1500 participants who are representative of the UK population, according to Prolific. We excluded 43 of them because they failed to answer two instructional manipulation checks correctly, suggesting that these participants did not participate in the experiment seriously (Oppenheimer et al., 2009). The final sample of 1457 participants is represented in terms of gender (47.9% male compared to 49.4% in the target population), ethnicity (85.5% white compared to 84.8% in the target population), and car ownership (64.9% own/have access to the car compared to 76% in the target population). However, the sample is biased toward the younger generation (58% between 18 and 37 compared to 32.6% in the target population) and highly educated people (60.7% higher education compared to 47% in the target population), although the median age was representative (35 years compared to 39 years in the target population).

Looking at other demographics, more than half of the participants currently work full-time (55.7%) or part-time (18.7%) and primarily work in the education (12.4%), IT (8.9%), or retail industry (7.9%). About a quarter of our participants hold a managerial position, either at a junior (8.9%), middle (14.1%), or upper management level (3.4%). Moreover, 45.4% of participants claimed they were aware of data marketplaces, for which they provided examples like Snowflake, Facebook, Prolific, Compare the Market, and YouGov. Meanwhile, 20.8% of participants were aware of PETs before participating in the survey, with many different encryption protocols named, such as end-to-end encryption, homomorphic encryption, zero-knowledge proofs, and Virtual

Private Networks (VPN). Further, the majority of our participants are privacy pragmatists (53.6%), followed by privacy fundamentalists (26.5%) and privacy unconcerned (19.9%), which is broadly similar to the distribution of privacy perspectives in comparison to other studies (Hughes-Roberts, 2013; Jensen et al., 2005; Kumaraguru & Cranor, 2005). See Table 16 for the demographic characteristics of the conducted sample. The complete datasets of the experiment are accessible at the 4TU research data repository (See Appendix D).

Variable	**Characteristics**	**N**	**%**
Age	18-27	395	27.1 %
	28-37	450	30.9 %
	38-47	257	17.6 %
	48-57	202	13.9 %
	58+	153	10.5 %
Gender	Male	698	47.9 %
	Female	745	51.1 %
	None of the above	11	0.8 %
	Prefer not to say	3	0.2 %
Education level	Doctorate degree (Ph.D./other)	39	2.7%
	Graduate degree (MA/MSc/MPhil/other)	276	18.9%
	Undergraduate degree (BA/BSc/other)	569	39.1%
	Technical/community college	135	9.3%
	High school diploma/A-levels	265	18.2%
	Secondary education (e.g., GED/GCSE)	158	10.8%
	No formal qualifications	12	0.8%
	I do not know/not applicable	1	0.1%

	Prefer not to say	2	0.1%
Employment status	Full-time	812	55.7 %
	Part-time	273	18.7 %
	Self-employed/freelance	18	1.2 %
	Not in paid work (e.g., homemaker, retired, or disabled)	184	12.6 %
	Not employed (students)	65	4.5 %
	Due to starting a new job within the next month	12	0.8 %
	Unemployed (and job-seeking)	74	5.1 %
	Prefer not to say	19	1.3 %
Industry type	Education & Training	180	12.4 %
	Information Technology	129	8.9 %
	Retail	115	7.9 %
	Medicine	104	7.1 %
	Finance	96	6.6 %
	Others	833	57.1%
Role at work	Upper Management	50	3.4 %
	Middle Management	205	14.1 %
	Junior Management	129	8.9 %
	Others	1000	68.6%
	Prefer not to say	73	5%
Car ownership	Yes	946	64.9 %
	Have access via parents/family	214	14.7 %
	Have access via leasing/rental	50	3.4 %

	No	247	17.0 %
Awareness of data marketplaces	Shared data through data marketplaces multiple times	77	5.3 %
	Shared data through data marketplaces once	67	4.6 %
	Know but never shared data through data marketplaces	518	35.6 %
	Never heard of data marketplaces	795	54.6 %
Awareness of PETs	Already know before the survey	303	20.8 %
	Have some idea because of the survey	876	60.1 %
	Still have no idea after the survey	278	19.1 %
Westin's Privacy Segmentation Index	Privacy fundamentalists	386	26.5 %
	Privacy unconcerned	290	19.9 %
	Privacy pragmatists	781	53.6 %

Table 16 Demographic characteristics (N=1457)

5.3.5. Data preparation

We conduct a Confirmatory Factor Analysis (CFA) using JASP version 0.16.1 to validate our constructs and measurement model (Brown & Moore, 2012). Similar to the pre-test (see Section 5.3.2), we first assess the model fit using the cut-off value suggested by Hu and Bentler (1999) for a good fit of those three measures: CFI ≥ 0.95, TLI ≥ 0.95, and RMSEA ≤ 0.06. The results show a good level of the fit index of the model, with CFI = 0.988, TLI = 0.984, and RMSEA = 0.043.

Next, we assess the validity of our constructs by looking into the factor loadings of each survey item using a threshold of 0.70 (Fornell & Larcker, 1981). Based on this threshold, we remove one item (CTRL_3) that does not meet this criterion. As shown

in Table 17, the rest of the items have factor loadings greater than 0.79, higher than the recommended threshold. We then assess the internal reliability of our model by looking at the Composite Reliability (CR) and Cronbach's alpha of each construct, which should have a value of 0.7 or higher (Hair et al., 2011, 2014). Table 17 shows that we establish convergent validity as all constructs have CR and Cronbach's alpha values greater than 0.8 and 0.74, respectively. Subsequently, we examine convergent validity through the Average Variance Extracted (AVE), which should be greater than 0.5 (Fornell & Larcker, 1981). Table 17 suggests all constructs satisfied the recommended value, with the lowest AVE being 0.67 for perceived control and the highest value of 0.93 for trust in data buyers.

Construct	**Items**	**Factor loadings**	**Mean**	**SD**	**R^2**	**α**	**CR**	**AVE**
Perceived control (CTRL)	CTRL_1	0.79	3.25	1.14	0.63	0.74	0.80	0.67
	CTRL_2	0.84	3.45	1.12	0.71			
Privacy concerns (PRIV)	PRIV_1	0.91	3.20	1.17	0.83	0.89	0.93	0.87
	PRIV_2	0.96	3.36	1.13	0.92			
Trust in data marketplace operators (TRSD)	TRSD_1	0.90	3.43	0.89	0.81	0.90	0.93	0.82
	TRSD_2	0.90	3.35	0.87	0.80			
	TRSD_3	0.92	3.50	0.90	0.84			
Trust in data buyers (TRSB)	TRSB_1	0.94	3.06	1.01	0.89	0.95	0.97	0.93
	TRSB_2	0.98	3.09	1.02	0.95			
	TRSB_3	0.97	3.09	1.03	0.94			
Perceived risk (RISK)	RISK_1	0.93	3.07	1.07	0.86	0.90	0.94	0.88
	RISK_2	0.95	3.07	1.09	0.89			
	WTSD_1	0.97	3.06	1.15	0.94	0.94	0.96	0.90
	WTSD_2	0.95	2.99	1.12	0.91			

Willingness to share data via data marketplaces (WTSD)	WTSD_3	0.92	2.88	1.17	0.85			

Note: SD = Standard Deviation; α = Cronbach's Alpha; CR = composite reliability; AVE = average variance extracted

Table 17 Descriptive statistics, convergent validity, internal consistency, and reliability

We also examine the discriminant validity of the constructs by checking whether the correlation among constructs is lower than the square root of AVE (Fornell & Larcker, 1981). All inter-construct correlation coefficients are well below the square root of AVE, suggesting that we also establish discriminant validity (see Table 18). Our final model comprises six constructs and 15 items (see Table 17).

	CTRL	**PRIV**	**TRSD**	**TRSB**	**RISK**	**WTSD**
CTRL	**0.82**					
PRIV	-0.34	**0.93**				
TRSD	0.45	-0.52	**0.90**			
TRSB	0.37	-0.45	0.68	**0.96**		
RISK	-0.43	0.75	-0.58	-0.48	**0.94**	
WTSD	0.43	-0.61	0.61	0.57	-0.70	**0.95**

Note: diagonals represent the square root of the average variance extracted, and other values represent the correlations

Table 18 Discriminant validity: correlation among constructs and the Square Root of the AVE

In the last step, we conduct Multi-Group Confirmatory Factor Analysis (MGCFA) to check whether all the criteria are also met in the three treatment conditions. We estimate the model using configural invariance testing and find a good level of the fit index, with CFI = 0.980, TLI = 0.975, and RMSEA = 0.053. All treatment conditions also show convergent and discriminant validity, with all factor loadings, CR, and Cronbach's

alpha higher than 0.7 and AVE higher than 0.5. Furthermore, comparing the square root of AVE and all inter-construct correlation coefficients in all treatment conditions suggests discriminant validity.

5.4. Results

5.4.1. Comparison of three treatments

We first conduct a MANOVA test to check whether antecedents of consumers' willingness to share data differ across our treatments (TTP, MPC, and DCP). We do this since we used a between-subject design with several antecedents of consumers' willingness to share data as multiple dependent variables. Subsequently, we conduct one-way ANOVAs to compare the effect of our treatments on each of the antecedents of consumers' data sharing decisions (see Table 19). For the analysis, we use composite scores for each construct, derived from aggregating the score of items that belong to each construct divided by the number of items. For instance, we used two survey items to measure "perceived control" (see Table 17). Therefore, we calculate the average of these two items to form a composite score for the "perceived control" construct. Further, we conduct Levene's test to test for equal variances and find that the variances for each construct are equal across our treatment.

The MANOVA shows that there is a significant effect of our treatments on antecedents of consumers' data sharing decisions (Pillai's trace = 0.08, $F(12, 2900) = 9.96$, $p < 0.001$). The subsequent one-way ANOVAs (see Table 19) reveal a significant effect of our treatments on perceived control [$F(2,1454) = 50.35$, $p < 0.001$, $\omega^2 = 0.06$], privacy concerns [$F(2,1454) = 9.49$, $p < 0.001$, $\omega^2 = 0.01$], trust in data buyers [$F(2,1454) = 5.65$, $p = 0.004$, $\omega^2 = 0.01$], perceived risk [$F(2,1454) = 15.38$, $p < 0.001$, $\omega^2 = 0.02$], and willingness to share data [$F(2,1454) = 7.17$, $p < 0.001$, $\omega^2 = 0.01$] at the $p < .05$ level. However, we find no significant differences across our treatments concerning trust in data marketplace operators [$F(2, 1454) = 2.21$, $p = 0.11$]. Therefore, H3 is not supported.

Construct		TTP (N=486)	MPC (N=484)	DCP (N=487)	df	F-value	ω²	p
CTRL	M	2.99	3.51	3.55	2	F(2, 1454) = 50.35	0.063	< .001***
	SD	1.02	0.94	0.96				
PRIV	M	3.45	3.17	3.22	2	F(2, 1454) = 9.49	0.012	< .001***
	SD	1.06	1.09	1.10				
TRSD	M	3.37	3.46	3.46	2	F(2, 1454) = 2.21	0.002	0.11
	SD	0.83	0.80	0.79				
TRSB	M	2.96	3.16	3.12	2	F(2, 1454) = 5.65	0.006	0.004**
	SD	0.99	0.93	0.99				
RISK	M	3.28	2.95	2.98	2	F(2, 1454) = 15.38	0.019	< .001***
	SD	1	1.05	1.01				
WTSD	M	2.83	3.08	3.02	2	F(2, 1454) = 7.17	0.008	< .001***
	SD	1.08	1.09	1.07				

Note: ** p < .01, *** p < .001

Table 19 The results of one-way ANOVA

To further detail the difference between the three treatments and test the remaining hypotheses, we conduct a series of post hoc tests using Tukey's correction (see Table 20 and Figure 10). Participants who receive MPC treatment perceive higher control over data (mean difference = 0.52, p < 0.001), have more trust in data buyers (mean difference = 0.2, p < 0.05), and are more willing to share data (mean difference = 0.25, p < 0.001) than those who receive TTP treatment. Thus, we provide support for H1, H4, and H6. Meanwhile, privacy concerns (mean difference = 0.29, p < 0.001) and perceived risks (mean difference = 0.38, p < 0.001) are lower for participants who

receive MPC treatment than those who receive TTP treatment. Therefore, H2 and H5 are also supported.

Construct	Comparison	Mean difference	SE	t	P_{tukey}
CTRL	TTP - MPC	0.52	0.06	-8.31	< .001***
	TTP - DCP	-0.56	0.06	-9.03	< .001***
	MPC - DCP	-0.04	0.06	-0.71	0.76
PRIV	TTP - MPC	0.29	0.07	4.11	< .001***
	TTP - DCP	0.23	0.07	3.32	0.003**
	MPC - DCP	-0.06	0.07	-0.80	0.706
TRSB	TTP - MPC	-0.2	0.06	-3.21	0.004**
	TTP - DCP	-0.15	0.06	-2.47	0.036*
	MPC - DCP	0.05	0.06	0.74	0.737
RISK	TTP - MPC	0.33	0.07	5	< .001***
	TTP - DCP	0.30	0.07	4.58	< .001***
	MPC - DCP	-0.03	0.07	-0.43	0.904
WTSD	TTP - MPC	-0.25	0.07	-3.61	< .001***
	TTP - DCP	-0.19	0.07	-2.80	0.014*
	MPC - DCP	0.06	0.07	0.81	0.696

Note: * p < .05, ** p < .01, *** p < .001

Table 20 Post-hoc comparisons

Meanwhile, participants who receive DCP treatment also perceive higher control over data (mean difference = 0.56, p < 0.001), have fewer privacy concerns (mean difference = 0.23, p < 0.003), have more trust in data buyers (mean difference = 0.15, p < 0.036), perceive lower risks (mean difference = 0.3, p < 0.001), and more willing to share data (mean difference = 0.19, p < 0.014) than those who receive TTP treatment.

However, there are no significant differences between participants who receive MPC treatment and DCP treatment in terms of perceived control (mean difference = 0.04, p = 0.76), perceived risks (mean difference = 0.03, p = 0.904), privacy concerns (mean difference = 0.06, p = 0.706), trust in data buyers (mean difference = 0.05, p = 0.737), and willingness to share data (mean difference = 0.06, p = 0.696).

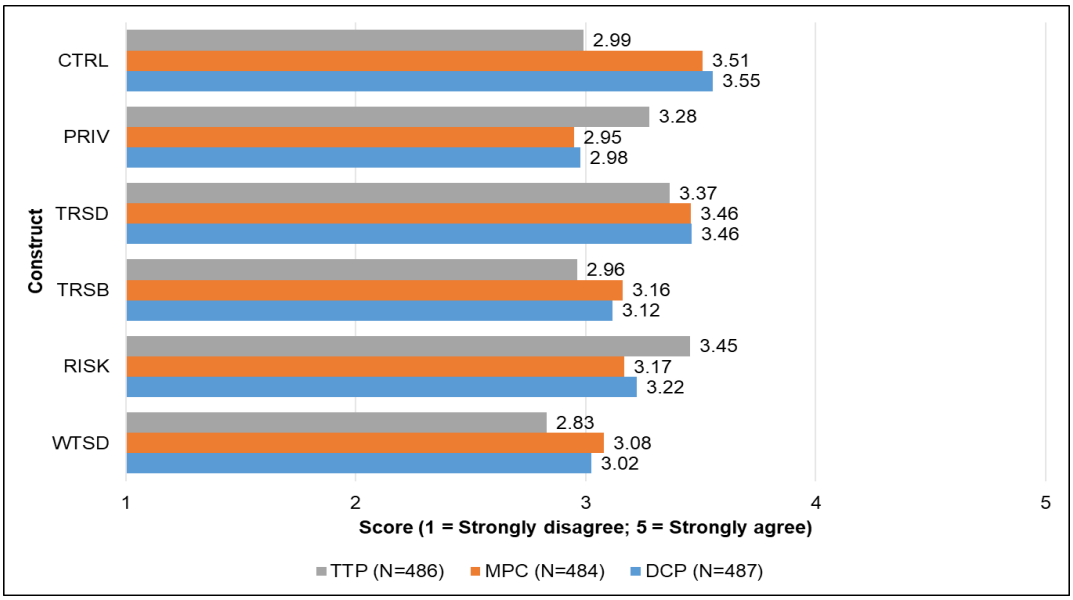


Figure 10 Mean differences between the three treatments

We summarize the results of the hypotheses testing in Table 21.

Hypotheses		Results
H1	Consumers' perceived control while sharing through data marketplaces that use MPC is higher than TTP	Supported
H2	Consumers' privacy concerns while sharing through data marketplaces that use MPC are lower than TTP	Supported
H3	Consumers' trust in operator while sharing through data marketplaces that use MPC is higher than TTP	Not supported

H4	Consumers' trust in data buyers while sharing through data marketplaces that use MPC is higher than TTP	Supported
H5	Consumers' perceived risks while sharing through data marketplaces that use MPC is lower than TTP	Supported
H6	Consumers' willingness to share through data marketplaces that use MPC is higher than TTP	Supported

Table 21 Summary of hypotheses testing

5.4.2. Main and interaction effects of control variables

Next, we perform two-way ANOVAs to test whether our control variables (i.e., Westin's Privacy Segmentation Index, industry type, car ownership, awareness of data marketplaces, and awareness of PETs) result in significant differences concerning the effect of our treatments on all constructs. As shown in Table 22, we find significant main effects of Westin's Privacy Segmentation Index on all constructs. Participants categorized as 'privacy unconcerned' have the greatest data sharing attitude, followed by the 'privacy fundamentalists' and 'privacy pragmatists' groups. Awareness of PETs also has significant main effects on all constructs (see Table 22). Participants aware of PETs have a more positive data sharing attitude than those who are not. Moreover, industry type has a significant main effect on trust in data buyers ($F(1, 1451) = 7.69$, $p = 0.006$, $\omega^2 = 0.01$). Tukey's post hoc correction shows that participants who work in non-IT sectors have more trust in data buyers than those who work in the IT sector ($p = 0.006$). However, we find no significant main effects of car ownership and awareness of data marketplaces on willingness to share data and its antecedents.

Finally, we test the interaction effect between control variables and our treatments (see Table 23). Initially, we found a significant interaction effect between our treatments and awareness of data marketplaces on one construct, namely perceived control ($F(1, 1451) = 5.58$, $p = 0.004$, $\omega^2 = 0.006$). This effect indicates that our

treatments affect different familiarity levels of data marketplaces differently. However, Tukey's post hoc correction showed that these differences are not significant. In other words, there are no significant differences in perceived control between participants who are aware and not aware of data marketplaces in TTP ($p = 0.265$), MPC ($p = 0.962$), and DCP treatments ($p = 0.118$).

Construct	Main effect									
	WESTIN		INDUSTRY		CAR		DMP		PETS	
	F-value	p	F-value	p	F-value	p	F-value	p	F-value	p
CTRL	F(2, 1448) = 49.03	< .001***	F(1, 1451) = 0.07	0.786	F(1, 1451) = 0.73	0.393	F(1, 1451) = 0.5	0.481	F(1, 1451) = 19.76	< .001***
PRIV	F(2, 1448) = 108.29	< .001***	F(1, 1451) = 0.34	0.559	F(1, 1451) = 0.01	0.935	F(1, 1451) = 0.13	0.716	F(1, 1451) = 18.3	< .001***
TRSD	F(2, 1448) = 162.75	< .001***	F(1, 1451) = 2.39	0.122	F(1, 1451) = 0.54	0.463	F(1, 1451) = 3.18	0.075	F(1, 1451) = 15.12	< .001***
TRSB	F(2, 1448) = 163.23	< .001***	F(1, 1451) = 7.69	0.006*	F(1, 1451) = 0.09	0.77	F(1, 1451) = 0.43	0.513	F(1, 1451) = 6.77	0.009**
RISK	F(2, 1448) = 116.52	< .001***	F(1, 1451) = 0.31	0.58	F(1, 1451) = 0.9	0.343	F(1, 1451) = 0.21	0.644	F(1, 1451) = 18.69	< .001***
WTSD	F(2, 1448) = 128.15	< .001***	F(1, 1451) = 0.01	0.906	F(1, 1451) = 0.01	0.907	F(1, 1451) = 3.34	0.068	F(1, 1451) = 16.04	< .001***

Note: WESTIN = Westin's Privacy Segmentation Index; INDUSTRY = Industry type; CAR = Car ownership; DMP = awareness of data marketplaces; PETS = awareness of Privacy-Enhancing Technologies (PETs); * p < .05, ** p < .01, *** p < .001

Table 22 The results of two-way ANOVA for the main effects

Construct	Interaction effect									
	SCENARIO x WESTIN		SCENARIO x INDUSTRY		SCENARIO x CAR		SCENARIO x DMP		SCENARIO x PETS	
	F-value	p	F-value	p	F-value	p	F-value	p	F-value	p
CTRL	F(4, 1448) = 0.41	0.805	F(2, 1451) = 2.32	0.099	F(2, 1451) = 0.44	0.643	F(2, 1451) = 5.58	0.004 **	F(2, 1451) = 1.21	0.299
PRIV	F(4, 1448) = 0.28	0.89	F(2, 1451) = 0.23	0.793	F(2, 1451) = 0.85	0.743	F(2, 1451) = 0.85	0.43	F(2, 1451) = 1.13	0.323
TRSD	F(4, 1448) = 0.3	0.88	F(2, 1451) = 1.04	0.354	F(2, 1451) = 2.49	0.519	F(2, 1451) = 2.49	0.083	F(2, 1451) = 1.24	0.29
TRSB	F(4, 1448) = 1.77	0.132	F(2, 1451) = 0.21	0.814	F(2, 1451) = 1.32	0.342	F(2, 1451) = 1.32	0.266	F(2, 1451) = 0.37	0.694
RISK	F(4, 1448) = 0.15	0.962	F(2, 1451) = 0.51	0.603	F(2, 1451) = 2.61	0.432	F(2, 1451) = 2.61	0.074	F(2, 1451) = 0.65	0.523
WTSD	F(4, 1448) = 1.05	0.38	F(2, 1451) = 0.02	0.983	F(2, 1451) = 2.1	0.199	F(2, 1451) = 2.1	0.122	F(2, 1451) = 0.49	0.612

Note: SCENARIO = Data sharing approaches(TTP/MPC/DCP); WESTIN = Westin's Privacy Segmentation Index; INDUSTRY = Industry type; CAR = Car ownership; DMP = awareness of data marketplaces; PETS = awareness of Privacy-Enhancing Technologies (PETs); * p < .05, ** p < .01, *** p < .001

Table 23 Two-way ANOVA for the interaction effect between treatments and control variables

5.5. Discussion

Our study conducted in this chapter shows the profound effect of MPC on data sharing decisions of consumers. We find that consumers are more willing to share data in a scenario that has MPC compared to one where a trusted third party handles the data exchange. In addition, the MPC scenario exhibits significantly higher scores on nearly all typical antecedents of data sharing decisions. MPC provides a higher sense of control over their data, lowers data sharing risks and privacy concerns, and increases trust toward data buyers. The only exception is trust in the data marketplace operator, which is not affected by MPC.

Our findings align with studies that suggest that properly communicating privacy-enhancing approaches positively affects users' perceptions and attitudes around data sharing. For instance, Kainda et al. (2010) and Yee (2002) emphasize the importance of adequately communicating MPC (and PETs in general), including their inner workings and usefulness for consumers, as consumers need the information to assess how those technologies can protect their data. Furthermore, our findings are also consistent with prior studies that suggest that communicating privacy protection and data control measures to consumers can reduce privacy concerns, increase trust in companies, and positively influence individuals' data disclosure decisions (Brandimarte et al., 2013; Cavusoglu et al., 2016; Miyazaki & Krishnamurthy, 2002).

We find no differences regarding trust in data marketplace operators. Apparently, in the context of data sharing, technical solutions like MPC themselves do not enhance trust in intermediaries (i.e., data marketplace operators). This finding goes against other studies, such as Ratnasingam et al. (2002), who found that technology does play a role in establishing a trustworthy environment for electronic transactions. Possibly, the use of MPC in itself does not signal that the operator is trustworthy. Another explanation is that MPC reduces the role of the data marketplace operators in facilitating data exchange, making the operator's role less relevant.

We find no significant differences between MPC and the fictitious technology DCP for any of the constructs. We included such a fictitious technology to control for the fact that MPC is a currently hyped technology, which could bias the participants. Our findings suggest that this plays no role. Further, this finding suggests that individual consumers might not scrutinize each detail of the underlying technology used in data sharing. Instead, consumers apparently focus on the usefulness and the (privacy) value of the technology being explained. In other words, consumers would perceive new technologies similarly in real-life as long as they are explained convincingly with terms like “encrypted” and “secure”.

We find that attitudes about privacy (i.e., Westin’s Privacy Segmentation Index) affect consumers’ data sharing decisions, whereas our other controls do not. Participants who are ‘unconcerned’ are more positive about data sharing than ‘fundamentalists’, while the ‘pragmatists’ group is somewhere in between. However, privacy attitude has no significant interaction effect, which implies that the impact of MPC on data sharing decisions is similar for each sub-group. This finding has two implications. First, the impact of MPC on data sharing preferences does not depend on a person’s privacy attitude. Hence, although ‘fundamentalists’ or ‘pragmatists’ consumers are more willing to share data in general, MPC impacts each group similarly. Nevertheless, implementing MPC is, to some extent, in line with what the ‘fundamentalists’ and ‘pragmatists’ groups want, namely, more robust mechanisms to protect consumers’ privacy (Kumaraguru & Cranor, 2005). Second, consumers already have a baseline intention to share data depending on their privacy attitude. The ‘unconcerned’ are more open to sharing data because they perceive that the benefits far outweigh the risks, while the ‘fundamentalists’ are very protective of their data (Kumaraguru & Cranor, 2005).

5.6. Limitations

Our work presented in this chapter has three main limitations. First, we excluded perceived benefits in our model despite being a dominant factor in explaining individual data sharing decisions (Culnan & Armstrong, 1999). We did this since we expect MPC will only change the “costs” in the privacy calculus model and will not impact perceived benefits. In our study, we keep the benefit constant by informing participants that being paid for the data is a part of each treatment. In this regard, the results are only valid in settings where people get monetary compensation. Nevertheless, if data buyers understand that MPC reduces the risks and concerns of consumers, data buyers can pay less for consumers’ data, implying that perceived benefits could also be affected by MPC. In this way, the privacy calculus would remain the same.

Second, at the start of the experiment, we explicitly disclosed the title and purposes of the experiment to participants (see Appendix C), which is a mandatory requirement for the Human Research Ethics Committee. A concern is that this could create undesirable effects, like participants guessing the hypothesis and answering questions based on the study purpose. However, we argue that this is not a big issue as we design our experiment so that participants only receive one treatment, making it not possible to compare different treatments. Nevertheless, future research might opt for an alternative way of presenting the title and the study purpose by disclosing them at the end as part of the debriefing for participants.

Third, our study only used hypothetical scenarios and mock-ups that are not working prototypes. This is due to the context of privacy-enhancing data marketplaces based on MPC, which is still limited in implementation. To counter this, we extensively informed participants about the setting at the beginning of the survey. In this regard, we essentially evaluate visualizations and explanations of MPC and not MPC itself.

We also rely on self-reporting answers from our participants rather than how participants actually use MPC in data sharing.

Fourth, while our study can be generalizable to the UK population, we only focused on the specific context of privacy-enhancing personal data marketplaces in the automotive domain. In this regard, the transferability of our findings to other domains can be questioned. Nevertheless, in other contexts dealing with similarly sensitive data, such as healthcare and energy, our findings could also be transferrable, especially when there is a common goal for the public good, like sharing data to improve healthcare treatment or promote energy transition.

5.7. Conclusions

In this chapter, we investigate the potential impact of MPC use in data marketplaces on antecedents of consumer data sharing. We found that, if visualized and appropriately communicated, **implementing MPC in data marketplaces could enhance consumers' control over data, increase trust in data buyers, reduce perceptions of risks, and lower privacy concerns, ultimately increasing willingness to share.** At the same time, we also found that **MPC did not increase trust in data marketplace operators**, suggesting that other mechanisms beyond technical solutions are needed to enhance trust in intermediaries. These effects are robust across different privacy attitudes.

From this chapter, we provide answers to the third research question and derive an important insight into MPC and data sharing from the consumer perspective. This way, we can establish a comprehensive understanding of the potential impact of MPC on antecedents of data sharing decisions from the perspective of MPC developers, businesses, and consumers. In the next chapter, we synthesize and integrate the key findings of the three studies to elaborate on how MPC use in data marketplaces could

change antecedents of data sharing. We also discuss the implication of our findings to theory and practice and suggest possible directions for future research.

6 Conclusions, limitations, and future research

In this research, we have investigated the potential impact of MPC on businesses' and consumers' data-sharing decisions through data marketplaces. In doing so, we conducted mixed-method research through business model analysis, semi-structured interviews with business actors in the automotive industry, and an online experiment through a crowdsourcing platform. In this chapter, we elaborate on our main findings and provide answers to our three research questions. Then, we discuss how findings from each study are similar, different, and reinforcing each other. Subsequently, we elaborate on our theoretical and practical implications. We conclude by outlining the limitations of the study and possible next steps for future research.

6.1. Revisiting research questions: main findings

This research aims to theorize the socio-technical implications of MPC on sharing through data marketplaces, by investigating data sharing antecedents that are potentially impacted by MPC and its resulting impact on data sharing decisions by businesses and individuals. To fulfill our research objective, we formulated three research questions, which we answered through mixed-method research based on three empirical research processes. In the remainder of this section, we outline answers for each research question and explain its importance in relation to our research objective.

Research question 1: what types of data sharing antecedents could be impacted by MPC use in data marketplaces?

This question aimed to identify antecedents of data sharing decisions that are potentially impacted by implementing MPC in data marketplaces. Answering this

question is essential to understand data sharing antecedents that we should focus on within the context of MPC and data marketplaces.

Findings from our interviews with MPC experts, which we derived using the unified business model framework, show that MPC could change the architecture of data marketplaces into peer-to-peer or (either single or multiple) cloud-based models. MPC could also change the role of data marketplaces towards data brokers that mediates data providers and buyers without storing and maintaining data flow between two parties. As a result, MPC could enable new value propositions in data marketplaces by facilitating **distributed, trustless, and privacy-enhancing data sharing that maintains control and reduces risks**. This is possible due to the capabilities of MPC in facilitating distributed data sharing in which data stays with data providers, and only the computation results are revealed to data buyers. **In this regard, we elicit four antecedents of data sharing decisions that could be impacted by MPC use in data marketplaces: control, privacy, trust, and risks.** As such, the answer to the first research question provides a foundation for subsequent empirical studies in which we investigate business and consumer perspectives on MPC and data sharing. We summarize our answer to the first research question in Figure 11.

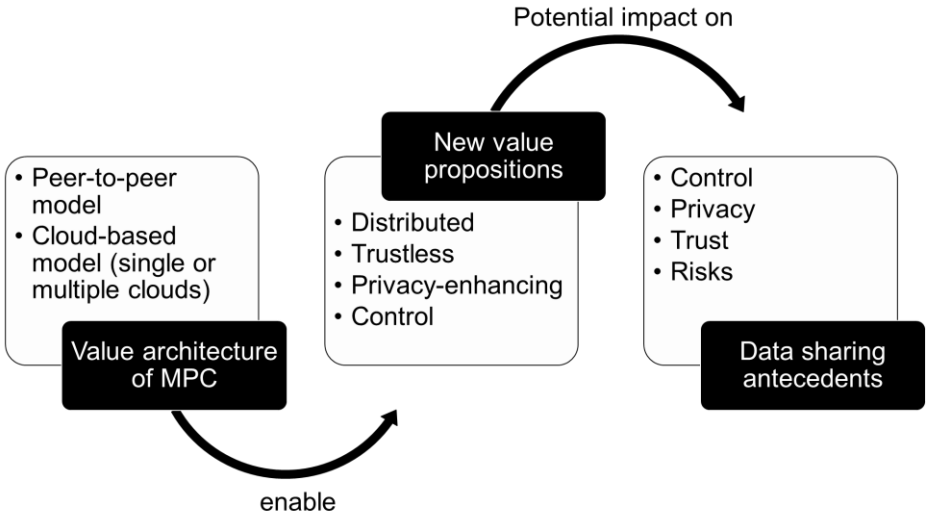


Figure 11 Graphical summary of answers to research question 1

Research question 2: what could be the impact of MPC use in data marketplaces on antecedents of data sharing by businesses?

The purpose of this question was to understand the business perspective on MPC and data sharing by investigating the extent to which MPC use in data marketplaces could impact data sharing antecedents identified in the first question. Answering this question is instrumental in substantiating whether there are changes regarding the relevance of antecedents of business data sharing and what kind of changes, if any. Our findings from interviews with business actors in the automotive industry indicate that MPC use in data marketplaces could result in three main changes regarding the relevance of data sharing antecedents by businesses.

First, we found that **MPC use in data marketplaces could enhance technology-based control by data providers.** From our domain exploration in Chapter 2, we learned that businesses could use MPC to collaborate in joint computation to produce meaningful insights without revealing input data. Then, as our findings in Chapter 4 show, businesses demand various control sources while sharing through data marketplaces: contract-based, structural-based, and technology-based control. Thus, MPC could play a crucial role in fulfilling the needs of businesses regarding control over data by acting as a technology-based control mechanism for data providers.

Second, our findings suggest that **MPC use in data marketplaces could shift the importance of trust for data providers, from trust in other business actors to trust in the technology.** With MPC in place, we found that businesses can participate in a distributed data sharing process without fully trusting each other and relying on a trusted third party. This is because the input data stays with data providers, while the computation is performed directly with data buyers to generate results instead of sharing the original data source. Therefore, data providers can be assured that data buyers can only utilize data insights without being able to misuse the original datasets. At the same time, our findings also show that MPC emphasizes the relevance of trust

in the underlying algorithm. As such, our findings indicate that it is important for data providers to understand how MPC works, whether it produces an accurate calculation, and whether the result indeed cannot reveal anything about the input data. Furthermore, our findings underline the multiple impacts of MPC on the need for trust in the context of business data sharing.

Third, **MPC use in data marketplaces could reduce the relevance of competitiveness risks and data misuse risks for data providers**. Our findings indicate that, given the ability of MPC to only allow the sharing of computational results and not the original datasets, it could prevent knowledge spillover to competitors. Also, data providers can be assured that data buyers can only utilize data insights without being able to misuse the original datasets. However, an interesting finding is a possible shift in who would suffer from data misuse risks, from data providers to data buyers. This is because queries asked by data buyers could reveal their know-how, which increases the possibility of reverse engineering by data providers. In this regard, our finding underlines the dual role of MPC as a double-edged sword that could eliminate the risks for data providers while creating one for data buyers.

Furthermore, our findings indicate three main boundary conditions under which MPC use in data marketplaces could impact antecedents of business data sharing. First, the **benefits** of using MPC in data sharing should be clear for businesses. Given the novelty of MPC, the use cases and business relevance might not be apparent yet and are highly dependent on different contexts and domains. As a result, companies are playing a waiting game to see the true value of MPC, ultimately hindering its potential impact on data sharing. Second, businesses should have sufficient **organizational readiness**, such as data pre-processing skills like cleaning and harmonization. MPC will change the role of data marketplaces towards full matchmaking, leading to data providers' increasing responsibility in data preparation before executing joint computation. Thus, businesses should have a data-driven mindset and the willingness to enhance organizational readiness to get the best out of MPC. Third, given the low

maturity of MPC, **data sensitivity** is regarded as crucial consideration by businesses. Much of what MPC can and cannot do is still relatively unknown, which leads to a cautious approach by businesses to share data via MPC-enabled data marketplaces. As such, businesses will start small by initially sharing generic and non-sensitive data. Once proven use cases of MPC exist, companies will be more willing to share more sensitive data. We summarize our answer to the second research question in Figure 12.

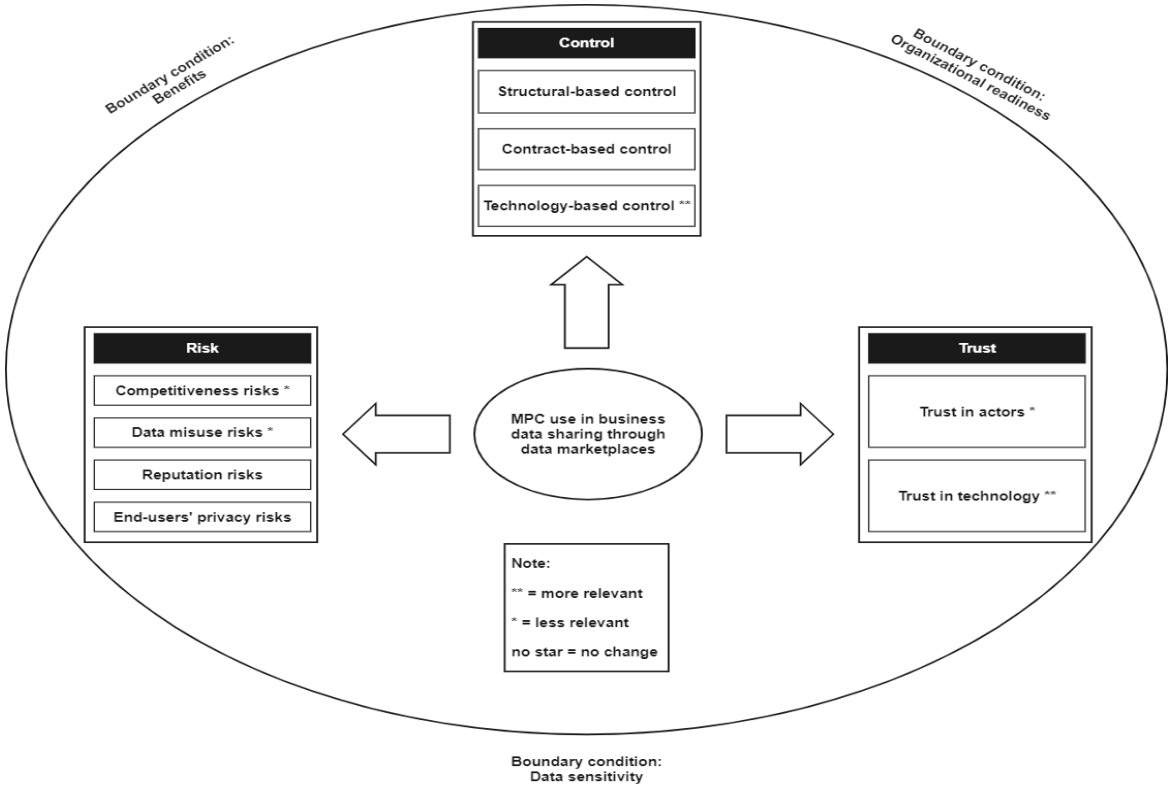


Figure 12 Graphical summary of answers to research question 2

<u>Research question 3: what could be the impact of MPC use in data marketplaces on antecedents of data sharing by consumers?</u>

This question aimed to understand the consumer perspective regarding the impact of MPC use in data marketplaces on antecedents of consumer data sharing. Given that MPC can also be implemented in consumer-facing technologies, answers to this

question lay a foundation for understanding changes resulting from MPC use in data marketplaces. Further, answering this question provided an avenue to substantiate whether privacy concerns, as one of the data sharing antecedents, could be impacted by MPC use in data marketplaces.

Based on an online experiment with consumers, in which we compared MPC with a baseline scenario of a Trusted Third Party (TTP), we found that MPC use in data marketplaces could impact all antecedents of consumer data sharing. Specifically, our findings indicate that **consumers who share through data marketplaces that use MPC perceive higher control over data, lower privacy concerns, and lower risks** than those who share through TTP. One explanation might be due to the capabilities of MPC (see Chapter 2), which facilitate distributed computation to generate insights for data buyers without revealing anything about consumers' data. With MPC, data buyers only receive computation results and not the original input data, which restricts what data buyers can do with the data. As such, consumers might feel in control over their data because it stays with them and is not transferred to an intermediary or data buyers, limiting the possibility of personal data misuse. This differs from TTP because the data is transferred and stored by an intermediary that processes the data exchange further before sending the data to data buyers. Thus, consumers might perceive less control over data given that there is a movement of consumers' data to an intermediary and data buyers. The possibility of data misuse is also higher since data buyers receive a complete dataset rather than only the computation results, threatening consumers' privacy concerns in the process.

Moreover, our findings also indicate different effects of MPC on trust. On the one hand, findings show that **consumers perceive higher trust in data buyers while sharing through data marketplaces that use MPC** than TTP. MPC already restricts what data buyers can do with the data by means of computation results instead of original datasets. In this regard, MPC could serve as a safeguard that allows individual consumers to trust data buyers. On the other hand, there is **no significant difference**

between MPC and TTP regarding trust in data marketplace operators. This finding might mean that trust in data marketplace operators is not relevant anymore since MPC could change its role toward data brokers (see Chapter 3). Alternatively, this finding might also indicate that trust in data marketplace operators is still important due to their role as data brokers. However, MPC alone cannot enhance trust in data marketplace operators. Instead, MPC should be complemented with other mechanisms like certification to enhance trust in data marketplace operators.

Furthermore, comparing MPC with fictitious technology DCP, our findings show no differences between the two conditions for all antecedents that we investigate. This finding suggests that consumers attribute value to the underlying ideas of the technology rather than the term itself. Thus, consumers were not biased by the fact that MPC is a currently hyped technology. Put differently, how the technology is named does not matter to consumers as long as it is communicated clearly, concisely, and convincingly since it would lead to higher control, higher trust, lower privacy concerns, lower risks, and a greater willingness to share. We summarize our answer to the third research question in Figure 13.

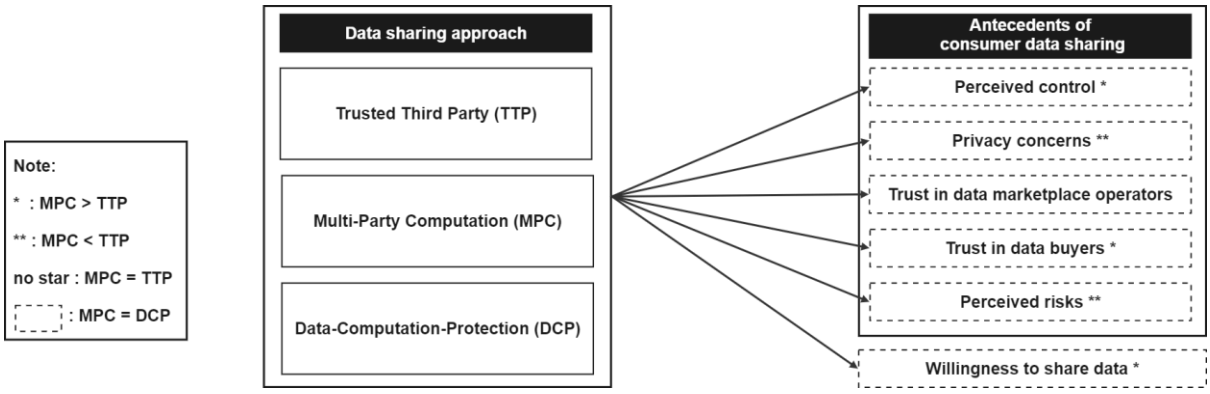


Figure 13 Graphical summary of answers to research question 3

Overall, our findings are useful for academics studying data sharing and privacy-enhancing technologies (PETs), especially MPC. This study is one of the first to investigate the impact of MPC, as one class of PETs, on antecedents of data sharing

decisions by businesses and consumers in the context of data marketplaces. Our study shows that control, privacy concerns, trust, and perceived risks are four data sharing antecedents that could be impacted by MPC use in data marketplaces. Thus, future researchers should focus on these four antecedents when investigating the impact of MPC on data sharing decisions. In particular, future researchers should focus on technology-based control, trust in technology, and data misuse risks (from data buyers' perspectives) while considering perceived benefits, organizational readiness, and perceived data sensitivity as boundary conditions. Additionally, researchers and practitioners can use our study as a foundation to effectively communicate and visualize how MPC works to non-experts to enhance control over data, lower privacy concerns, reduce risks, increase trust in data buyers, and ultimately increase willingness to share data. Moreover, academics and practitioners should start positioning MPC as a data collaboration tool rather than a data privacy tool to ensure businesses can realize the benefits of using MPC. Furthermore, academics and practitioners should not view MPC as the only solution to address data sharing challenges and instead complement MPC with other data governance mechanisms to ensure a trustworthy and privacy-enhancing data sharing environment.

6.2. Discussion of findings

In this section, we compare the findings between business and consumer perspectives on MPC and data sharing decisions in data marketplaces. We base our discussion on the findings from answering the first research question (i.e., antecedents of data sharing decisions impacted by MPC), given its instrumental role in providing a foundation for empirical studies. Specifically, we discuss how findings from the business perspective study (Chapter 4) and the consumer perspective study (Chapter 5) are different, similar, or reinforce each other. This diversity of findings may partly be explained by differences in methods used in studies on business and consumer perspectives. For instance, a qualitative study enables us to capture a rich array of perspectives from business actors regarding how MPC could change the

relevance of data sharing antecedents. This differs from a quantitative study based on an online experiment with consumers as it allows us to test the possible relation between MPC and data sharing antecedents.

Perceived control

We found a similar effect of MPC on businesses and consumers in enhancing control over data while participating in data sharing through data marketplaces. An explanation for this similar effect is the way MPC works, which is the same for businesses and consumers by enabling joint computation to generate insights without revealing input data. The difference is mainly about the positioning of MPC towards other control sources. From investigating the business perspective on data sharing and MPC (Chapter 4), we found three important control sources for businesses: contract-based, structural-based, and technology-based control, in which MPC can be categorized as technology-based control. Meanwhile, in understanding the consumer perspective on data sharing and MPC (Chapter 5), we compared MPC with TTP, which can be broadly positioned as a structural-based control mechanism due to its reliance on trusted intermediaries. In this regard, although our quantitative findings indicate that MPC enhances more control over data than TTP, which can be interpreted as the superiority of technology-based control over structural-based control, it does not necessarily mean that technology-based control is the most important, as shown in the qualitative study. Instead, combining three control sources should provide optimal control over data while sharing in data marketplaces.

Privacy concerns

We found differences regarding how businesses and consumers view privacy concerns. The findings of the business study (Chapter 4) suggest that businesses see privacy concerns as a secondary objective rather than the primary goal. Put differently, businesses mainly look at how sharing data in a privacy-enhancing way using MPC-

based data marketplaces can enable them to develop novel products and services for their end users. Nevertheless, despite being driven by business motives, businesses still need to take action to protect end-users' privacy. Otherwise, if something bad happens, the implications are more complicated, like possible misuse of end-users' data, reputational harm, and fines due to regulatory violations, which are all related to perceived risks. In this regard, the implications would extend beyond the company itself to other entities like the end-users.

Meanwhile, in the consumer study (Chapter 5), consumers are referring to their privacy concerns, which are about managing who can access their data and the purpose of use. In this regard, consumers are more concerned about the possibility that their personal data is exposed and can be accessed by other parties that should not have access to it. Nevertheless, although there is a possible misuse of personal data if something wrong happens in the data sharing process, other issues like reputational harm and regulatory violations are irrelevant to consumers. Because of this, consumers have more relative freedom to engage in data sharing than businesses, as the dynamics faced by consumers are less complicated than businesses. Hence, our findings indicate that once MPC is used in data marketplaces and explained convincingly, consumers would likely be more willing to share data because their privacy concerns would be reduced.

Trust

We found the differing impact of MPC on trust between the two studies. In the business study (Chapter 4), trust in other actors involved in data sharing was initially seen as a key aspect for businesses. However, MPC reduces the importance of trust in other actors, given its ability to facilitate joint computation to generate insights without fully trusting other actors. At the same time, MPC increases the importance of trust in the technology (i.e., MPC algorithm) that it will execute the correct computation and protect input data as promised. Meanwhile, in the consumer study

(Chapter 5), consumers view trust in other actors involved as an important factor in data sharing decisions, although MPC only enhances trust in data buyers and does not have an impact on enhancing trust in intermediaries like data marketplace operators. However, trust in technology does not seem to be important for consumers. While this difference is partly by design (i.e., the literature review does not point to the importance of consumers' trust in the technology), it might also be due to the different need to understand the technology's underlying details. Because businesses deal with sensitive and confidential data, they need to ensure that the whole MPC process is trustworthy and can really protect their data during the computation. This need for trust is not evident from the consumers' perspective, as they mainly focus on how the technology is communicated.

Perceived risks

Findings from both studies indicate the similar impact of MPC in reducing data sharing risks by jointly computing input data to generate meaningful insights without giving away the input data to data buyers. This way, the data access and usage are restricted to only the computation results and not the original datasets as the input data. This mechanism creates positive impressions for businesses and consumers, ultimately leading to lower perceptions of risks.

However, the type of risks that are impacted by MPC differs due to the contrasting views and importance of risks between them. Risks faced by businesses are more complex, like losing competitive advantage, violating data protection regulations, and reputational harm due to data leakage and misuse. These risks are not only about confidential business-sensitive data but also end-users' personal data, which links to the privacy concerns discussed before. In this regard, companies will greatly value MPC if it can address business-related risks. Meanwhile, consumers' view of risk is more about the possible misuse and exploitation of data by third parties beyond the agreed purpose and not about competitiveness and reputational harm. Hence, MPC

will be seen as a valuable tool for consumers if it can reduce the risk of exposing their personal data that is supposed to be private.

Another difference is regarding the shift of risk. Findings from the business study (Chapter 4) indicate that using MPC might shift the data misuse risk to data buyers by revealing know-how and the possibility of reverse engineering based on the query asked by data buyers. However, this risk shift is not apparent in consumer data sharing (Chapter 5) because the query asked by data buyers is, in theory, not valuable to consumers. Hence, the shift of data misuse risk is not relevant for consumers. The same cannot be said for businesses because data providers could use the query asked by data buyers to perform reverse engineering and steal the intellectual ideas of data buyers.

We summarize the comparison of findings between business and consumer perspectives findings in Table 24.

Data sharing antecedents	**Business perspective**	**Consumer perspective**
Perceived control	MPC enhances companies' control over data through technology-based control	MPC significantly enhances consumers' control over data more than TTP
Privacy concerns	MPC protects the privacy of companies' end-user data, but it is not a sufficient condition to justify business adoption of MPC	MPC significantly reduces consumers' privacy concerns more than TTP
Perceived risks	MPC reduces competitiveness risks due to knowledge spillover but shifts the data misuse risk to	MPC significantly reduces consumers' data sharing risks more than TTP

	data buyers by revealing know-how and reverse engineering	
Trust	MPC reduces the relevance of trust in other actors but increases the relevance of trust in the technology	MPC significantly enhances trust toward data buyers more than TTP, but has no impact on trust in data marketplace operators

Table 24 Comparison between business and consumer perspectives on MPC and data sharing antecedents

6.3. Theoretical contributions

This research is among the first to investigate how novel privacy-enhancing technologies like MPC challenge the current understanding regarding data sharing in data marketplaces. We discuss each of our theoretical contributions below.

Theorizing the potential impact of MPC on data sharing decisions

Our main contribution from this study is theorizing how MPC could potentially impact antecedents of data sharing decisions in the context of data marketplaces. Specifically, we contribute to the literature on business-to-business data sharing by specifying the concepts of perceived control, trust, and perceived risks into a set of propositions, as well as three boundary conditions (i.e., benefits, readiness, and data sensitivity) that scholars should consider when studying the impact of MPC on data sharing decisions. We also contribute to the information privacy literature by introducing MPC as a novel context/boundary condition that affects almost all of the key antecedents of consumers' data sharing decisions in data marketplaces. These contributions are crucial because, so far, we lack knowledge of the meaning of MPC for data sharing decisions by businesses and consumers, especially in the early stage

of MPC adoption. Furthermore, MPC differs from existing data sharing approaches that primarily rely on a Trusted Third Party (TTP) (Bruun et al., 2020; Helminger & Rechberger, 2022). Hence, we cannot simply transfer existing knowledge to this new phenomenon (cf., Alvesson & Sandberg, 2011; Gkeredakis & Constantinides, 2019).

By understanding the potential impact of MPC on data sharing decisions, researchers working on the emerging topics of data sharing could leverage our findings as a foundation for future studies in these increasingly important areas. As Xu and Dinev (2022) argued, the emergence of AI and machine learning has prompted a shift in privacy governance from data-centered privacy concerns to knowledge-centered privacy concerns. At the same time, privacy assurance techniques are growing rapidly and gaining more attention from academics and practitioners. In this regard, we respond to the call from Xu and Dinev (2022) by laying a basis for future research on the societal implications of MPC as one class of privacy-enhancing technologies. Researchers can build upon our findings by, for instance, developing and testing the theoretical relations among the specified concepts based on the boundary conditions. Alternatively, researchers could use MPC or other PETs as a moderator that strengthens or weakens existing explanations of data sharing decisions. Furthermore, qualitative work could be done as a follow-up to examine the reasons why businesses and consumers perceive, for instance, fewer privacy risks and more control over their data when faced with MPC.

Socio-technical understanding of MPC

We also contribute to the socio-technical studies of MPC by providing insights on the potential value of MPC for businesses and consumers within the context of data sharing in data marketplaces. Specifically, the value of MPC for businesses and consumers includes functional elements inherent in the technology, such as enhancing privacy perceptions. However, the value goes beyond that, affecting interrelated aspects of perceived control and risks and more symbolic aspects such

as trust in data buyers. This way, as our findings show, businesses can create value from MPC by sharing data to generate new insights while maintaining control.

At the same time, our findings indicate that consumers would be more willing to share data via privacy-enhancing approaches like MPC than a conventional solution like Trusted Third Party (TTP). This finding resonates with the general tendency in the data sharing field towards privacy preservation and decentralization. Both aspects are core to MPC, which enhances privacy by moving the algorithm to the data. Our findings suggest that decentralization and privacy preservation are not only societally desirable but also benefit the business interests of involved parties by making consumers more inclined to share their data.

Furthermore, this research supports expectations that MPC and related privacy-enhancing technologies will facilitate a data economy where businesses and consumers feel in control of their data, trust data buyers, and ultimately share their data in return for monetary compensation. Our research thus sketches a fertile ground for socio-technical studies on the affordances and values of MPC, which is overlooked in the MPC literature (Agrawal et al., 2021; Bruun et al., 2020; Kanger & Pruulmann-Vengerfeldt, 2015).

Addressing data marketplaces challenges

To data marketplaces literature, our research contributes to understanding adoption challenges. Data marketplaces adoption is generally low, potentially due to a lack of trust and confidence in data buyers as well as privacy concerns of data providers. Businesses and consumers also demand stronger data control and privacy protection. In this regard, our findings indicate that data marketplace operators could implement MPC to proactively protect privacy, give control, and enable privacy-friendly business models in data marketplaces (Bonazzi et al., 2010; Conger et al., 2013). This way, we show that implementing MPC in data marketplaces likely favors businesses' and

consumers' adoption of data marketplaces. At the same time, data marketplace operators can implement MPC to fulfill their moral obligation to maintain data control by businesses and protect individual privacy (Bonazzi et al., 2010; Conger, 2009).

Meanwhile, from the consumers' perspective, MPC impacts the risks part of the privacy calculus. Hence, without substantially increasing the monetary compensation, MPC could foster the willingness of consumers to use data marketplaces. Further, our findings suggest that giving consumers more control over their data through technology favors willingness to share data, which aligns with recent calls to move away from passive data monetization strategies (e.g., Van Alstyne et al., 2021).

6.4. Implications for practitioners

Our research also offers recommendations that practitioners can consider when considering using MPC in their data sharing practice.

Positioning of MPC as data collaboration tools

Our research is relevant to MPC developers and service providers to rethink the positioning of MPC for businesses. Currently, MPC is promoted as a privacy tool (i.e., privacy-enhancing technology) that offers the business value of generating aggregated output while retaining the input data from multiple data providers (Ofe et al., 2022). However, as our finding in the study on business perspective (Chapter 4) shows, businesses do not see privacy as a compelling value proposition despite repeated calls from scholars to implement privacy-friendly business models (e.g., Bonazzi et al., 2010; Conger et al., 2013; Zöll et al., 2021). Instead, businesses often view privacy as a secondary aim, making MPC less appealing for businesses due to its framing as a privacy tool. In other words, although MPC might benefit consumers by lowering their privacy concerns, it does not create sufficient benefits for businesses to justify the adoption of MPC.

To address this challenge, our findings indicate that MPC could be promoted differently as a data collaboration tool (Lundy-Bryan, 2021) that goes beyond privacy protection (Agrawal et al., 2021). We show the value of MPC in governing collaboration in data sharing by improving companies' perception of control over data, establishing trust, and reducing data sharing risks (Lundy-Bryan, 2021). This is important because control, trust, and risk reduction are a basis for inter-organizational data sharing and collaboration. Thus, we offer an alternative way for MPC developers and service providers to frame and promote the benefits of MPC beyond privacy.

MPC could transform the business models of data marketplaces

Our research could benefit data marketplace operators and other intermediary platforms facilitating data sharing. As pointed out by Abbas et al. (2021), MPC could affect the value proposition of those platforms by (1) enabling sharing and computation of data insights without disclosing the input data; and (2) affording control over data without a trusted third party. Our study provides empirical evidence that MPC could address control, privacy, trust, and risk issues in data sharing, which are challenges data sharing platforms struggle to deal with (M. Spiekermann, 2019). MPC could even create other values for data sharing platforms by changing how these platforms perform matchmaking based on data collaboration potential. For instance, instead of matching data buyers with data providers that want to sell their data, data sharing platforms could perform matchmaking between multiple parties that have the potential to collaborate in addressing collective problems such as financial fraud, traffic congestion, and energy transition. Then, platform owners could offer an end-to-end solution by implementing MPC-based privacy-enhancing analytics as a way to address those collective problems. Alternatively, data marketplaces could also completely move from the matchmaking function and entirely focus on offering privacy-enhancing data analytics platforms to potential customers. Therefore, in line with our findings in Chapter 3, those platforms could transform their business models

by implementing MPC to offer unique services for their customers and gain a competitive advantage.

MPC could enable companies to implement privacy-friendly business models and level the playing field in the data economy

Broadly, our findings are important for proponents of data economy and data sharing. Current challenges, such as consumers' lack of trust in data sharing, may be overcome by solutions such as MPC. By enhancing control, reducing risks, lowering privacy concerns, and increasing trust in companies, MPC offers important business values for companies to support their data-driven businesses. Therefore, we argue that companies should start using PETs like MPC to balance consumers' privacy and value creation from data sharing. In this regard, our suggestion aligns with repeated calls from scholars that businesses should shift towards implementing data-driven and privacy-friendly business models (Bonazzi et al., 2010; Conger et al., 2013; Zöll et al., 2021). By increasing the chances that consumers will share data, MPC may contribute to unlocking the economic and societal value of data sharing.

Our findings also show that sharing data through MPC is more favorable than with a Trusted Third Party (TTP). Thus, our findings could challenge the relevance of trusted parties in the data economy that typically holds massive amount of data. By using MPC for data sharing, companies (and even individual consumers) could directly share data without relying on big technology companies, which would level the playing field in the data economy. In this regard, policymakers have started to propose various policy recommendations that encourage the use of privacy-enhancing technologies (European Commission, 2023; European Data Protection Board, 2021). While this is a positive first step, policymakers should take further action by incorporating MPC and other privacy-enhancing technologies in the future regulatory framework to stimulate free flow of data while maintaining fairness in data access (European Commission, 2020).

MPC must be communicated effectively to improve knowledgeability

Findings from our empirical works indicate the importance of knowledgeability for businesses and consumers to maximize the impact of MPC on data sharing decisions. Businesses need to enhance their data-related competencies as MPC usage would shift the responsibility towards data providers in, among others, data preparation and cleaning. Meanwhile, consumers are more willing to share data if they are aware of privacy-enhancing technologies like MPC. However, MPC is generally deemed too complex for non-experts to understand (Agrawal et al., 2021; Bruun et al., 2020), which hinders the uptake of MPC. In this regard, our research developed mock-ups to explain and visualize how MPC works, which were preceded by exploration with MPC experts (Chapter 3) and business actors (Chapter 4). As such, we lay a blueprint to effectively and concisely explain and visualize MPC to a general audience, which practitioners can build upon. By educating businesses and consumers, they can better see and appreciate the value of using MPC in data sharing, ultimately increasing their willingness to share.

Companies can implement unethical practices in communicating MPC

Our findings also demonstrate that corporations can realize the benefits in terms of consumers' trust without actually implementing specific PETs. Since consumers' perceptions do not differ between MPC and a fictitious technology, companies might also take our findings to justify strategies where they claim to use PETs by handling buzzwords without implementing the actual PETs. However, we argue that this is not necessarily the case because consumers are becoming increasingly aware and critical over time, especially in light of scandals like Facebook and Cambridge Analytica. In this regard, to make these informed decisions, consumers need transparency on how corporations use PETs to interact with their data, including (1) what kind of data is collected, (2) for what purposes, and (3) how their data is protected. Furthermore, policymakers could also play a role in this dynamic through

(1) capacity building to further raise consumers' awareness and understanding of how companies should be transparent in protecting consumers' data and (2) enforcing transparency through regulations and verifications to check whether companies are actually implementing those protection mechanisms. By pressuring even further from multiple fronts, companies should be more incentivized to protect consumers' data, including implementing PETs like MPC and being transparent about its usage.

MPC should be complemented with other governance mechanisms

Our findings in the quantitative study show that MPC does not significantly enhance consumers' trust in data marketplace operators. These findings underline the importance of acknowledging the limitations of MPC, which practitioners should be aware of. Specifically, our findings indicate that, given the complex nature of trust in data sharing, increasing trust in intermediaries cannot simply rely on technical solutions like MPC alone. Instead, intermediaries like data marketplace operators should complement MPC with other governance mechanisms to ensure a trustworthy data sharing environment. Examples include gatekeeping for prospective participants through certification (Otto & Jarke, 2019) and screening (Richter & Slowinski, 2019; Son et al., 2006) before they can perform transactions in data marketplaces. Data marketplace operators could also establish data usage policies (Noorian et al., 2014; Otto & Jarke, 2019) and traditional contractual arrangements (Bergman et al., 2022) to control the purpose and data usage access. Furthermore, measures like rating (Zavolokina et al., 2018), review systems (Fruhwirth et al., 2020; Subramanian, 2017), and smart contracts (Fruhwirth et al., 2020; van de Ven et al., 2021) could also be adopted by data marketplace operators.

6.5. Limitations

The research conducted in this dissertation has three main limitations. First, we used a short presentation and textual explanation as mock-ups to illustrate how MPC works

and hypothetical scenarios of MPC-enabled data marketplaces in the automotive sector. Hence, we essentially evaluate visualizations and explanations of MPC and not evaluate MPC itself. Thus, interviewees and survey participants might not have a clear idea of MPC and data marketplaces as they only read the explanation for a short time and not actually using MPC for sharing in data marketplaces. This is evident in our observation during the qualitative study, in which interviewees often answer their questions based on data sharing platforms without monetary compensation or the generic view of data sharing. As such, our findings should be interpreted under specific constraints of the early stage of MPC adoption. This way, we provide the foundation for comparison and analysis of subsequent studies, such as comparing the impact of MPC on data sharing decisions when it is widely adopted by businesses and information about MPC is easily accessible to the general audience.

Second, this research was focused on a specific context of data marketplaces in the automotive domain, which represent a setting that deals with highly sensitive data and has high data sharing hurdles. This approach allows us to use this context as a starting point for future research on other facets. In this regard, our findings should be interpreted under these specific constraints, especially within a context that deals with sharing similarly sensitive data to realize a common goal for the public good. Furthermore, we provide a blueprint for studying other domains, like sharing non-sensitive data between business partners, which could provide additional insights into the context under which MPC is relevant and can improve data sharing decisions.

Lastly, in this research, we have considered the perspective of MPC experts, business actors in the automotive domain, and consumers in different studies. Nevertheless, we did not incorporate the view of data marketplace operators in any of our studies. A rationale is that we focused on how MPC can address data sharing barriers faced by businesses and consumers as data providers, although those barriers might overlap with those faced by data marketplace operators. Thus, it is likely that involving

data marketplace operators could also offer new insights regarding businesses' and consumers' decisions to share data and the role of MPC.

6.6. Recommendations for future research

In this research, we conducted each empirical work based on a single perspective. That is, the business model analysis to determine data sharing antecedents (Chapter 3) was based on the perspective of MPC developers. In contrast, we conducted a qualitative study to investigate the business perspective (Chapter 4), while the study from the consumer perspective was conducted using a quantitative study (Chapter 5). Future research could perform a follow-up study to expand further our understanding of the impact of MPC on data sharing decisions. For instance, from investigating business perspectives on MPC and data sharing (Chapter 4), future research could test the resulting propositions and the nomological net between data sharing antecedents and willingness to share using quantitative approaches such as Structural Equation Modelling (SEM) or online experiments. The same approaches can also be useful to compare the impact of MPC with the currently known data sharing solutions such as a trusted third party (TTP) within the context of business-to-business data sharing. The three conditions of benefits, readiness, and data sensitivity are also relevant for future research. One way is to consider them as moderating effects, which strengthen or weaken the relationship between MPC and the antecedents of data sharing decisions. For instance, in experiments, researchers should keep benefits and readiness at constant and high levels, while sensitivity should be maintained at constant and low levels. Alternatively, researchers could treat these three conditions as the boundary conditions under which the relationship between MPC and the three antecedents holds, or, rather, that MPC certainly has no effect in settings with low readiness and benefits and high sensitivity (c.f., Busse et al., 2017).

Similarly, drawing upon findings of our quantitative study (Chapter 5), future studies can treat MPC as either a boundary condition or a moderator to test whether (1) the currently known theoretical relationships of data sharing decisions still hold and (2) it strengthens or weakens existing explanations of data sharing decisions. Also, since we kept the benefit constant in the quantitative study, future studies can compare different benefits received for sharing data (either monetary or other intangible benefits) based on different scenarios (trusted third party, MPC, or fictitious technology). Moreover, future research might incorporate a certification process for data buyers and data marketplace operators, as our findings indicate the need to complement MPC with non-technical measures to increase trust and willingness to share. Other than that, further research might benefit from defining a trusted third party explicitly, as it can range from big tech corporations (e.g., Google, Facebook, Amazon) to public authorities (e.g., government institutions) and non-governmental organizations, which might lead to different perceptions when comparing trusted third party with MPC. Furthermore, qualitative studies like interviews or case studies could also be done as a follow-up to examine the reason behind changes due to MPC, like why consumers perceive lower privacy concerns and greater control over data.

Our research also relies on a mock-up in the form of a short presentation that explains what MPC is. In other words, our research is based on a thought experiment in which interviewees (in the qualitative study) and participants (in the quantitative study) were asked to imagine a possible scenario of sharing data through MPC-based data marketplaces. Future research could employ design science research (DSR) to develop an artifact or working prototype of MPC-based data marketplaces. The resulting artifact can then be used to evaluate the impact of MPC on decisions to share data through data marketplaces by businesses and consumers. This way, we can also reduce bias from researchers' influence and dependency of participants on the presentation that explains what MPC is. At the same time, we can account that people often use technology in ways not envisioned by the developer (c.f., Kallinikos

et al., 2013). Evaluating the impact of MPC based on a working prototype is also important because it can consider that people use technologies in ways not intended or envisioned.

In this research, we focus on a specific context of data marketplaces, meaning that data marketplaces facilitate data exchange between unknown parties with no prior relationships for monetary compensation. Future research can expand our study by designing an artifact of MPC and evaluating its impact on data sharing decisions on other types of intermediary platforms that facilitate data exchange between known partners. This represents a closed relationship between multiple partners with a typically well-defined purpose of data usage, which means greater control, a high degree of trust, and lower risks even without MPC in place. Hence, it might be interesting to see whether the impact of MPC differs between the two setups of intermediary platforms.

Our emphasis in this research is on data sharing through data marketplaces in the automotive domain, which is relatively conventional regarding data-related practice. Future research can also evaluate the impact of MPC on data sharing decisions in different domains, such as digital health, finance, economics, and energy. For instance, researchers could build upon the work of Bampoulidis et al. (2022), who developed a privacy-preserving analytics solution to connect mobile phone data with individuals' health records. Specifically, scholars could use the developed MPC-based solutions as a scenario to investigate the perception of mobile phone users if their telecom operators decide to share data with healthcare institutions using MPC-based solutions. Evaluating the impact of MPC in various domains is important given the novelty of the technology and its potential to address societal challenges that were not possible to address before.

Our study theorizes the link between MPC and data sharing decisions by drawing on governance-related concepts like perceived control, privacy concerns, trust, and

perceived risks. Future studies can also link MPC with the literature on platform governance, particularly the openness of data platforms. As argued by de Reuver et al. (2022), the openness of data platforms might entail different objects like data-driven insights, which can be realized through new mechanisms like MPC. In this regard, scholars can explore how MPC breaks the tension between openness and control and whether new tensions arise when implementing MPC. Alternatively, data sharing decisions through MPC-based solutions could also be explained by a unifying theory of polycentric information commons (Mindel et al., 2018). This theory conceptualizes decentralized online information systems like MPC as information commons that require polycentric information practices like regulations, incremental adaptation, shared accountability, and recognition. Thus, scholars can also explore the impact of polycentric information practices in ensuring the sustainability of MPC-based solutions in deriving values for relevant actors.

Further research might benefit from linking MPC with other theories to understand the socio-technical implications of MPC beyond data sharing decisions. For instance, drawing from the dynamic capabilities (e.g., Teece, 2018; Teece & Pisano, 1994) and data-driven business model literature (e.g., Günther et al., 2017; Hartmann et al., 2016; Sorescu, 2017), we can view MPC as a privacy-preserving analytics tool and investigate how it can enable firms' capabilities to implement data-driven business models (DDBM). Scholars could investigate the resources and capabilities needed by firms to incorporate MPC or privacy-preserving analytics into their business processes. Moreover, the emergence of MPC service providers requires a further understanding of their business models, such as data sources, new services enabled by MPC, added value for customers, revenue models, and possible beneficiaries. In this regard, we can leverage emerging data-driven business models tooling, like data service cards (Breitfuss et al., 2020, 2023) and data-driven business canvas (Fruhwirth, Breitfuss, et al., 2020), to deepen our understanding of the business model of MPC-based services.

Finally, from an ecosystem perspective, more actors are becoming interested in the technology by actively developing MPC solutions, establishing consortiums to develop MPC jointly, or becoming active users of MPC solutions. Given this enormous growth, further research could dig deep into the current state of MPC ecosystems, including key stakeholder groups and their role within the ecosystem. Scholars could also draw upon collective action theory in investigating collaboration issues in projects aimed at developing MPC solutions, such as actors' roles, resources, and interests in the technology and the alignment of stakeholders' interests. This way, we can understand whether the novelty of MPC creates unique challenges and opportunities while executing the project.

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.

Adams, T. L., Li, Y., & Liu, H. (2020). A replication of beyond the turk: Alternative platforms for crowdsourcing behavioral research–sometimes preferable to student groups. *AIS Transactions on Replication Research*, *6*(1), 15.

Agrawal, N., Binns, R., Van Kleek, M., Laine, K., & Shadbolt, N. (2021). Exploring Design and Governance Challenges in the Development of Privacy-Preserving Computation. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–13. https://doi.org/10.1145/3411764.3445677

Alashoor, T., Han, S., & Joseph, R. C. (2017). Familiarity with big data, privacy concerns, and self-disclosure accuracy in social networking websites: An APCO model. *Communications of the Association for Information Systems*, *41*(1), 4.

Al-Debei, M. M., & Avison, D. (2010). Developing a unified framework of the business model concept. *European Journal of Information Systems*, *19*(3), 359–376. https://doi.org/10.1057/ejis.2010.21

Alt, R., & Zimmermann, H.-D. (2001). Preface: Introduction to special section–business models. *Electronic Markets*, *11*(1), 3–9.

Alter, G., Falk, B. H., Lu, S., & Ostrovsky, R. (2018). Computing Statistics from Private Data. *Data Science Journal*, *17*(0), Article 0. https://doi.org/10.5334/dsj-2018-031

Alvesson, M., & Sandberg, J. (2011). Generating research questions through problematization. *Academy of Management Review*, *36*(2), 247–271.

Alvsvåg, R., Bokolo, A., & Petersen, S. A. (2022). The Role of a Data Marketplace for Innovation and Value-Added Services in Smart and Sustainable Cities. In F. Phillipson, G. Eichler, C. Erfurth, & G. Fahrnberger (Eds.), *Innovations for*

Community Services (pp. 215–230). Springer International Publishing. https://doi.org/10.1007/978-3-031-06668-9_16

Archer, D. W., Bogdanov, D., Lindell, Y., Kamm, L., Nielsen, K., Pagter, J. I., Smart, N. P., & Wright, R. N. (2018). From Keys to Databases—Real-World Applications of Secure Multi-Party Computation. *The Computer Journal*, *61*(12), 1749–1771. https://doi.org/10.1093/comjnl/bxy090

Arnaut, C., Pont, M., Scaria, E., Berghmans, A., & Leconte, S. (2018). Study on data sharing between companies in Europe. *A Study Prepared for the European Commission Directorate-General for Communications Networks, Content and Technology by Everis Benelux*, *24*.

Arrow, K. J. (1972). Economic Welfare and the Allocation of Resources for Invention. In C. K. Rowley (Ed.), *Readings in Industrial Economics: Volume Two: Private Enterprise and State Intervention* (pp. 219–236). Macmillan Education UK. https://doi.org/10.1007/978-1-349-15486-9_13

Asare, A. K., Brashear-Alejandro, T. G., & Kang, J. (2016). B2B technology adoption in customer driven supply chains. *Journal of Business & Industrial Marketing*, *31*(1), 1–12. https://doi.org/10.1108/JBIM-02-2015-0022

Athanasopoulou, A., de Reuver, M., Nikou, S., & Bouwman, H. (2019). What technology enabled services impact business models in the automotive industry? An exploratory study. *Futures*, *109*, 73–83.

Azarm-Daigle, M., Kuziemsky, C., & Peyton, L. (2015). A Review of Cross Organizational Healthcare Data Sharing. *Procedia Computer Science*, *63*, 425–432. https://doi.org/10.1016/j.procs.2015.08.363

Ažderska, T. (2012). Co-evolving trust mechanisms for catering user behavior. *IFIP International Conference on Trust Management*, 1–16.

Baden-Fuller, C., & Haefliger, S. (2013). Business models and technological innovation. *Long Range Planning*, *46*(6), 419–426.

Balazinska, M., Howe, B., Koutris, P., Suciu, D., & Upadhyaya, P. (2013). A Discussion on Pricing Relational Data. In V. Tannen, L. Wong, L. Libkin, W. Fan, W.-C. Tan,

& M. Fourman (Eds.), *In Search of Elegance in the Theory and Practice of Computation: Essays Dedicated to Peter Buneman* (pp. 167–173). Springer. https://doi.org/10.1007/978-3-642-41660-6_7

Balson, D., & Dixon, W. (2020). *Cyber Information Sharing: Building Collective Security*. World Economic Forum. https://www3.weforum.org/docs/WEF_Cyber_Information_Sharing_2020.pdf

Bampoulidis, A., Bruni, A., Helminger, L., Kales, D., Rechberger, C., & Walch, R. (2022). Privately Connecting Mobility to Infectious Diseases via Applied Cryptography. *Proceedings on Privacy Enhancing Technologies*, *4*, 768–788.

Bataineh, A. S., Mizouni, R., Bentahar, J., & El Barachi, M. (2020). Toward monetizing personal data: A two-sided market analysis. *Future Generation Computer Systems*, *111*, 435–459.

Beaver, D., Micali, S., & Rogaway, P. (1990). The round complexity of secure protocols. *Proceedings of the Twenty-Second Annual ACM Symposium on Theory of Computing*, 503–513.

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

Bélanger, F., Crossler, R. E., & Correia, J. (2021). Privacy Maintenance in Self-Digitization: The Effect of Information Disclosure on Continuance Intentions. *ACM SIGMIS Database: The DATABASE for Advances in Information Systems*, *52*(2), 7–24.

Bergman, R., Abbas, A. E., Jung, S., Werker, C., & de Reuver, M. (2022). Business model archetypes for data marketplaces in the automotive industry. *Electronic Markets*. https://doi.org/10.1007/s12525-022-00547-x

Bestavros, A., Lapets, A., & Varia, M. (2017). User-centric distributed solutions for privacy-preserving analytics. *Communications of the ACM*, *60*(2), 37–39.

Beverungen, D., Hess, T., Köster, A., & Lehrer, C. (2022). From private digital platforms to public data spaces: Implications for the digital transformation. *Electronic Markets*, *32*(2), 493–501. https://doi.org/10.1007/s12525-022-00553-z

Bogdanov, D., Jõemets, M., Siim, S., & Vaht, M. (2015). How the Estonian Tax and Customs Board Evaluated a Tax Fraud Detection System Based on Secure Multi-party Computation. In R. Böhme & T. Okamoto (Eds.), *Financial Cryptography and Data Security* (pp. 227–234). Springer. https://doi.org/10.1007/978-3-662-47854-7_14

Bogdanov, D., Talviste, R., & Willemson, J. (2012). Deploying Secure Multi-Party Computation for Financial Data Analysis. In A. D. Keromytis (Ed.), *Financial Cryptography and Data Security* (pp. 57–64). Springer. https://doi.org/10.1007/978-3-642-32946-3_5

Bogetoft, P., Christensen, D. L., Damgård, I., Geisler, M., Jakobsen, T., Krøigaard, M., Nielsen, J. D., Nielsen, J. B., Nielsen, K., Pagter, J., Schwartzbach, M., & Toft, T. (2009). Secure Multiparty Computation Goes Live. In R. Dingledine & P. Golle (Eds.), *Financial Cryptography and Data Security* (pp. 325–343). Springer. https://doi.org/10.1007/978-3-642-03549-4_20

Bonazzi, R., Fritscher, B., & Pigneur, Y. (2010). Business model considerations for privacy protection in a mobile location based context. *2010 14th International Conference on Intelligence in Next Generation Networks*, 1–8. https://doi.org/10.1109/ICIN.2010.5640885

Bons, R., 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*, *10*, 198–225.

Bons, R., Lee, R. M., & Wagenaar, R. W. (1998). Designing Trustworthy Interorganizational Trade Procedures for Open Electronic Commerce. *International Journal of Electronic Commerce*, *2*(3), 61–83. https://doi.org/10.1080/10864415.1998.11518316

Borking, J. J. (2011). Why adopting privacy enhancing technologies (pets) takes so much time. In *Computers, privacy and data protection: An element of choice* (pp. 309–341). Springer.

Borking, J. J., & Raab, C. (2001). Laws, PETs and other technologies for privacy protection. *Journal of Information, Law and Technology*, *1*, 1–14.

Bouwman, H., de Vos, H., & Haaker, T. (2008). *Mobile service innovation and business models*. Springer Science & Business Media.

Bouwman, H., Nikou, S., & de Reuver, M. (2019). Digitalization, business models, and SMEs: How do business model innovation practices improve performance of digitalizing SMEs? *Telecommunications Policy*, *43*(9), 101828.

Brandão, A., Mamede, H. S., & Gonçalves, R. (2019). Trusted Data's Marketplace. In Á. Rocha, H. Adeli, L. P. Reis, & S. Costanzo (Eds.), *New Knowledge in Information Systems and Technologies* (pp. 515–527). Springer International Publishing. https://doi.org/10.1007/978-3-030-16181-1_49

Brandimarte, L., Acquisti, A., & Loewenstein, G. (2013). Misplaced confidences: Privacy and the control paradox. *Social Psychological and Personality Science*, *4*(3), 340–347.

Breitfuss, G., Fruhwirth, M., Wolf-Brenner, C., Riedl, A., de Reuver, G. A., Ginthoer, R., & Pimas, O. (2020). Data Service Cards-A Supporting Tool For Data-Driven Business. *33rd Bled EConference: Enabling Technology for a Sustainable Society*, 599–614.

Breitfuss, G., Santa-Maria, T., Fruhwirth, M., & Disch, L. (2023, January 3). Use Your Data: Design and Evaluation of a Card-Based Ideation Tool for Data-Driven Services. *Proceedings of the 56th Hawaii International Conference on System Sciences*. https://hdl.handle.net/10125/103267

Brink, H. I. L. (1993). Validity and reliability in qualitative research. *Curationis*, *16*(2), Article 2. https://doi.org/10.4102/curationis.v16i2.1396

Brown, T. A., & Moore, M. T. (2012). Confirmatory factor analysis. *Handbook of Structural Equation Modeling*, 361–379.

Bruun, M. H., Andersen, A. O., & Mannov, A. (2020). Infrastructures of trust and distrust: The politics and ethics of emerging cryptographic technologies. *Anthropology Today*, *36*(2), 13–17. https://doi.org/10.1111/1467-8322.12562

Bryant, A., & Charmaz, K. (2007). *The SAGE Handbook of Grounded Theory*. SAGE Publications Ltd. https://doi.org/10.4135/9781848607941

Buck, C., & Reith, R. (2020). Privacy on the road? Evaluating German consumers' intention to use connected cars. *International Journal of Automotive Technology and Management*, *20*(3), 297–318.

Burkert, H. (1997). Privacy-enhancing technologies: Typology, critique, vision. *Technology and Privacy: The New Landscape*, *125*.

Busse, C., Kach, A. P., & Wagner, S. M. (2017). Boundary Conditions: What They Are, How to Explore Them, Why We Need Them, and When to Consider Them. *Organizational Research Methods*, *20*(4), 574–609. https://doi.org/10.1177/1094428116641191

Campbell, D. T., & Stanley, J. C. (2015). *Experimental and quasi-experimental designs for research*. Ravenio books.

Cavusoglu, H., Phan, T. Q., Cavusoglu, H., & Airoldi, E. M. (2016). Assessing the impact of granular privacy controls on content sharing and disclosure on Facebook. *Information Systems Research*, *27*(4), 848–879.

Charness, G., Gneezy, U., & Kuhn, M. A. (2012). Experimental methods: Between-subject and within-subject design. *Journal of Economic Behavior & Organization*, *81*(1), 1–8. https://doi.org/10.1016/j.jebo.2011.08.009

Chen, Y.-H., Lin, T.-P., & Yen, D. C. (2014). How to facilitate inter-organizational knowledge sharing: The impact of trust. *Information & Management*, *51*(5), 568–578. https://doi.org/10.1016/j.im.2014.03.007

Chesbrough, H., & Rosenbloom, R. S. (2002). The role of the business model in capturing value from innovation: Evidence from Xerox Corporation's technology spin-off companies. *Industrial and Corporate Change*, *11*(3), 529–555.

Cho, H., Wu, D. J., & Berger, B. (2018). Secure genome-wide association analysis using multiparty computation. *Nature Biotechnology*, *36*(6), Article 6. https://doi.org/10.1038/nbt.4108

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

Cichy, P., Salge, T.-O., & Kohli, R. (2021). Privacy Concerns and Data Sharing in the Internet of Things: Mixed Methods Evidence from Connected Cars. *MIS Quarterly*.

Conger, S. (2009). Personal Information Privacy: A Multi-Party Endeavor. *Journal of Electronic Commerce in Organizations (JECO)*, *7*(1), 71–82. https://doi.org/10.4018/jeco.2009010106

Conger, S., Pratt, J. H., & Loch, K. D. (2013). Personal information privacy and emerging technologies. *Information Systems Journal*, *23*(5), 401–417. https://doi.org/10.1111/j.1365-2575.2012.00402.x

Constant, D., Kiesler, S., & Sproull, L. (1994). What's Mine Is Ours, or Is It? A Study of Attitudes about Information Sharing. *Information Systems Research*, *5*(4), 400–421. https://doi.org/10.1287/isre.5.4.400

Creswell, J. W., & Clark, V. L. P. (2017). *Designing and conducting mixed methods research*. Sage publications.

Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.

Cropanzano, R., & Mitchell, M. S. (2005). Social Exchange Theory: An Interdisciplinary Review. *Journal of Management*, *31*(6), 874–900. https://doi.org/10.1177/0149206305279602

Culnan, M. J., & Armstrong, P. K. (1999). Information privacy concerns, procedural fairness, and impersonal trust: An empirical investigation. *Organization Science*, *10*(1), 104–115.

Curry, E., Scerri, S., & Tuikka, T. (2022). Data Spaces: Design, Deployment, and Future Directions. In E. Curry, S. Scerri, & T. Tuikka (Eds.), *Data Spaces: Design, Deployment and Future Directions* (pp. 1–17). Springer International Publishing. https://doi.org/10.1007/978-3-030-98636-0_1

Cusumano, M. A., Gawer, A., & Yoffie, D. B. (2019). *The Business of Platforms: Strategy in the Age of Digital Competition, Innovation, and Power*. Harper Business. https://www.harpercollins.co.uk/9780062896322/the-business-of-platforms-strategy-in-the-age-of-digital-competition-innovation-and-power/

Dahlberg, T., & Nokkala, T. (2019). *Willingness to Share Supply Chain Data in an Ecosystem Governed Platform–An Interview Study*.

De Prieëlle, F., De Reuver, M., & Rezaei, J. (2020). The Role of Ecosystem Data Governance in Adoption of Data Platforms by Internet-of-Things Data Providers: Case of Dutch Horticulture Industry. *IEEE Transactions on Engineering Management*, 1–11. https://doi.org/10.1109/TEM.2020.2966024

de Reuver, M., Bouwman, H., & MacInnes, I. (2009). Business models dynamics for start-ups and innovating e-businesses. *International Journal of Electronic Business*, *7*(3), 269–286.

de Reuver, M., Ofe, H., Agahari, W., Abbas, A. E., & Zuiderwijk, A. (2022). The openness of data platforms: A research agenda. *Proceedings of the 1st International Workshop on Data Economy*, 34–41. https://doi.org/10.1145/3565011.3569056

Dinev, T., Albano, V., Xu, H., D'Atri, A., & Hart, P. (2016). Individuals' Attitudes Towards Electronic Health Records: A Privacy Calculus Perspective. In A. Gupta, V. L. Patel, & R. A. Greenes (Eds.), *Advances in Healthcare Informatics and Analytics* (pp. 19–50). Springer International Publishing. https://doi.org/10.1007/978-3-319-23294-2_2

Dinev, T., & Hart, P. (2006). An extended privacy calculus model for e-commerce transactions. *Information Systems Research*, *17*(1), 61–80.

Dinev, T., Xu, H., Smith, J. H., & Hart, P. (2013). Information privacy and correlates: An empirical attempt to bridge and distinguish privacy-related concepts. *European Journal of Information Systems*, *22*(3), 295–316.

Docherty, I., Marsden, G., & Anable, J. (2018). The governance of smart mobility. *Transportation Research Part A: Policy and Practice*, *115*, 114–125.

Dolci, R. (2020). *Realising platform control in data marketplaces through Secure Multi-Party Computation: A qualitative study exploring the use of Secure Multi-Party Computation (MPC) as an instrument for realising platform control in data marketplaces* [Master's thesis, Delft University of Technology]. https://repository.tudelft.nl/islandora/object/uuid%3A1d568346-86d5-402b-babe-26d2ba46809b

Drees, H., Kubitza, D. O., Lipp, J., Pretzsch, S., & Langdon, C. S. (2021). Mobility Data Space-First Implementation and Business Opportunities. *ITS World Congress*.

Du, T. C., Lai, V. S., Cheung, W., & Cui, X. (2012). Willingness to share information in a supply chain: A partnership-data-process perspective. *Information & Management*, *49*(2), 89–98. https://doi.org/10.1016/j.im.2011.10.003

Dwork, C. (2006). Differential Privacy. In M. Bugliesi, B. Preneel, V. Sassone, & I. Wegener (Eds.), *Automata, Languages and Programming* (pp. 1–12). Springer. https://doi.org/10.1007/11787006_1

Dwork, C., & Roth, A. (2014). The Algorithmic Foundations of Differential Privacy. *Foundations and Trends® in Theoretical Computer Science*, *9*(3–4), 211–407. https://doi.org/10.1561/0400000042

Elgarah, W., Falaleeva, N., Saunders, C. C., Ilie, V., Shim, J. T., & Courtney, James. F. (2005). Data exchange in interorganizational relationships: Review through multiple conceptual lenses. *ACM SIGMIS Database: The DATABASE for Advances in Information Systems*, *36*(1), 8–29. https://doi.org/10.1145/1047070.1047073

Elliot, D., & Quest, L. (2020, January 14). *It's time to redefine how data is governed, controlled and shared. Here's how.* World Economic Forum.

https://www.weforum.org/agenda/2020/01/future-of-data-protect-and-regulation/

Elsaify, M., & Hasan, S. (2021). Data exchanges among firms. *Digital Business*, *1*(2), 100010.

Emerson, R. M. (1976). Social Exchange Theory. *Annual Review of Sociology*, *2*(1), 335–362. https://doi.org/10.1146/annurev.so.02.080176.002003

Emsley, D., & Kidon, F. (2007). The Relationship between Trust and Control in International Joint Ventures: Evidence from the Airline Industry*. *Contemporary Accounting Research*, *24*(3), 829–858. https://doi.org/10.1506/car.24.3.7

Etikan, I., Musa, S. A., & Alkassim, R. S. (2015). Comparison of Convenience Sampling and Purposive Sampling. *American Journal of Theoretical and Applied Statistics*, *5*(1), Article 1. https://doi.org/10.11648/j.ajtas.20160501.11

Eurich, M., Oertel, N., & Boutellier, R. (2010). The impact of perceived privacy risks on organizations' willingness to share item-level event data across the supply chain. *Electronic Commerce Research*, *10*(3), 423–440.

European Commission. (2020). *A European strategy for data*. https://ec.europa.eu/info/sites/default/files/communication-european-strategy-data-19feb2020_en.pdf

European Commission. (2023). *Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on European statistics on population and housing, amending Regulation (EC) No 862/2007 and repealing Regulations (EC) No 763/2008 and (EU) No 1260/2013*. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52023PC0031

European Data Protection Board. (2021). *Recommendations 01/2020 on measures that supplement transfer tools to ensure compliance with the EU level of protection of personal data*. https://edpb.europa.eu/our-work-tools/our-documents/recommendations/recommendations-012020-measures-supplement-transfer_en

Evans, D., Kolesnikov, V., & Rosulek, M. (2018). A Pragmatic Introduction to Secure Multi-Party Computation. *Foundations and Trends® in Privacy and Security*, *2*(2–3), 70–246. https://doi.org/10.1561/3300000019

Farrelly, R., & Chew, E. (2016). *Who's in to win?: Participation rate in a primary personal information market*.

Faujdar, V. (2019). *Customer Acceptance of a Revenue Management Platform with Multi-Party Computation: Application of Multi-Party Computation to Revenue Management in the Semiconductor Industry* [Master's thesis, Delft University of Technology]. https://repository.tudelft.nl/islandora/object/uuid%3Ac0bd3fed-307b-4ab6-bf33-79e6f52cd991

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39–50.

Franklin, M., Halevy, A., & Maier, D. (2005). From databases to dataspaces: A new abstraction for information management. *ACM SIGMOD Record*, *34*(4), 27–33. https://doi.org/10.1145/1107499.1107502

Fricker, S. A., & Maksimov, Y. V. (2017). Pricing of data products in data marketplaces. *International Conference of Software Business*, 49–66.

Fruhwirth, M., Breitfuss, G., & Pammer-Schindler, V. (2020). The Data Product Canvas—A Visual Collaborative Tool for Designing Data-Driven Business Models. *33rd Bled EConference: Enabling Technology for a Sustainable Society*. https://aisel.aisnet.org/bled2020/8

Fruhwirth, M., Rachinger, M., & Prlja, E. (2020). Discovering Business Models of Data Marketplaces. *Proceedings of the 53rd Hawaii International Conference on System Sciences*.

Fu, H.-P., Chang, T.-H., Ku, C.-Y., Chang, T.-S., & Huang, C.-H. (2014). The critical success factors affecting the adoption of inter-organization systems by SMEs. *Journal of Business & Industrial Marketing*, *29*(5), 400–416. https://doi.org/10.1108/JBIM-04-2012-0070

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*, 103465. https://doi.org/10.1016/j.jnca.2022.103465

Gartner. (2021). *Gartner Says Digital Ethics is at the Peak of Inflated Expectations in the 2021 Gartner Hype Cycle for Privacy*. https://www.gartner.com/en/newsroom/press-releases/2021-09-30-gartner-says-digital-ethics-is-at-the-peak-of-inflate

Gast, J., Gundolf, K., Harms, R., & Matos Collado, E. (2019). Knowledge management and coopetition: How do cooperating competitors balance the needs to share and protect their knowledge? *Industrial Marketing Management*, *77*, 65–74. https://doi.org/10.1016/j.indmarman.2018.12.007

Gentry, C. (2009). Fully homomorphic encryption using ideal lattices. *Proceedings of the Forty-First Annual ACM Symposium on Theory of Computing*, 169–178. https://doi.org/10.1145/1536414.1536440

Gerber, N., Zimmermann, V., Henhapl, B., Emeröz, S., & Volkamer, M. (2018). Finally johnny can encrypt: But does this make him feel more secure? *Proceedings of the 13th International Conference on Availability, Reliability and Security*, 1–10.

Ghazawneh, A., & Henfridsson, O. (2013). Balancing platform control and external contribution in third-party development: The boundary resources model. *Information Systems Journal*, *23*(2), 173–192. https://doi.org/10.1111/j.1365-2575.2012.00406.x

Gkeredakis, M., & Constantinides, P. (2019). Phenomenon-based problematization: Coordinating in the digital era. *Information and Organization*, *29*(3), 100254.

Goldbach, T., Benlian, A., & Buxmann, P. (2018). Differential effects of formal and self-control in mobile platform ecosystems: Multi-method findings on third-party developers' continuance intentions and application quality. *Information & Management*, *55*(3), 271–284. https://doi.org/10.1016/j.im.2017.07.003

Goldreich, O., Goldwasser, S., & Micali, S. (1986). How to construct random functions. *Journal of the ACM*, *33*(4), 792–807. https://doi.org/10.1145/6490.6503

Günther, W. A., Mehrizi, M. H. R., Huysman, M., & Feldberg, F. (2017). Debating big data: A literature review on realizing value from big data. *The Journal of Strategic Information Systems*, *26*(3), 191–209.

Guo, C., Katz, J., Wang, X., & Yu, Y. (2020). Efficient and Secure Multiparty Computation from Fixed-Key Block Ciphers. *2020 IEEE Symposium on Security and Privacy (SP)*, 825–841. https://doi.org/10.1109/SP40000.2020.00016

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a silver bullet. *Journal of Marketing Theory and Practice*, *19*(2), 139–152.

Hair, J. F., Sarstedt, M., Hopkins, L., & Kuppelwieser, V. G. (2014). Partial least squares structural equation modeling (PLS-SEM): An emerging tool in business research. *European Business Review*.

Hajli, N., & Lin, X. (2016). Exploring the security of information sharing on social networking sites: The role of perceived control of information. *Journal of Business Ethics*, *133*(1), 111–123.

Hall, H., & Widén-Wulff, G. (2008). Social Exchange, Social Capital and Information Sharing in Online Environments: Lessons from Three Case Studies. In M.-L. Huotari & D. Elisabeth (Eds.), *From Information Provision to Knowledge Production* (p. 21). University of Oulu.

Harborth, D., Pape, S., & Rannenberg, K. (2020). Explaining the Technology Use Behavior of Privacy-Enhancing Technologies: The Case of Tor and JonDonym. *Proceedings on Privacy Enhancing Technologies*, *2020*(2), 111–128. https://doi.org/10.2478/popets-2020-0020

Harris, D., Khan, L., Paul, R., & Thuraisingham, B. (2007). Standards for secure data sharing across organizations. *Computer Standards & Interfaces*, *29*(1), 86–96. https://doi.org/10.1016/j.csi.2006.01.004

Hart, P., & Saunders, C. (1997). Power and trust: Critical factors in the adoption and use of electronic data interchange. *Organization Science*, *8*(1), 23–42.

Hartmann, P. M., Zaki, M., Feldmann, N., & Neely, A. (2016). Capturing value from big data – a taxonomy of data-driven business models used by start-up firms. *International Journal of Operations & Production Management*, *36*(10), 1382–1406. https://doi.org/10.1108/IJOPM-02-2014-0098

Helminger, L., & Rechberger, C. (2022). Multi-Party Computation in the GDPR. *Privacy Symposium 2022-Data Protection Law International Convergence and Compliance with Innovative Technologies (DPLICIT)*.

Hemenway, B., Lu, S., Ostrovsky, R., & Welser IV, W. (2016). High-Precision Secure Computation of Satellite Collision Probabilities. In V. Zikas & R. De Prisco (Eds.), *Security and Cryptography for Networks* (pp. 169–187). Springer International Publishing. https://doi.org/10.1007/978-3-319-44618-9_9

Heurix, J., Zimmermann, P., Neubauer, T., & Fenz, S. (2015). A taxonomy for privacy enhancing technologies. *Computers & Security*, *53*, 1–17. https://doi.org/10.1016/j.cose.2015.05.002

Hsu, C.-L., & Lin, J. C.-C. (2016). An empirical examination of consumer adoption of Internet of Things services: Network externalities and concern for information privacy perspectives. *Computers in Human Behavior*, *62*, 516–527.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55.

Hughes-Roberts, T. (2013). Privacy and Social Networks: Is Concern a Valid Indicator of Intention and Behaviour? *2013 International Conference on Social Computing*, 909–912. https://doi.org/10.1109/SocialCom.2013.140

Jagadeesh, K. A., Wu, D. J., Birgmeier, J. A., Boneh, D., & Bejerano, G. (2017). Deriving genomic diagnoses without revealing patient genomes. *Science*, *357*(6352), 692–695. https://doi.org/10.1126/science.aam9710

Janssen, M., & Zuiderwijk, A. (2014). Infomediary Business Models for Connecting Open Data Providers and Users. *Social Science Computer Review*, *32*(5), 694–711. https://doi.org/10.1177/0894439314525902

Jarman, H., Luna-Reyes, L. F., & Zhang, J. (2016). Public Value and Private Organizations. In H. Jarman & L. F. Luna-Reyes (Eds.), *Private Data and Public Value: Governance, Green Consumption, and Sustainable Supply Chains* (pp. 1–23). Springer International Publishing. https://doi.org/10.1007/978-3-319-27823-0_1

Jensen, C., Potts, C., & Jensen, C. (2005). Privacy practices of Internet users: Self-reports versus observed behavior. *International Journal of Human-Computer Studies*, *63*(1), 203–227. https://doi.org/10.1016/j.ijhcs.2005.04.019

Jernigan, S., Kiron, D., & Ransbotham, S. (2016). Data sharing and analytics are driving success with IoT. *MIT Sloan Management Review*, *58*(1).

Johnson, M. E. (2009). Managing information risk and the economics of security. In *Managing Information Risk and the Economics of Security* (pp. 1–16). Springer.

Kagal, L., Finin, T., & Joshi, A. (2001). Trust-based security in pervasive computing environments. *Computer*, *34*(12), 154–157. https://doi.org/10.1109/2.970591

Kainda, R., Flechais, I., & Roscoe, A. W. (2010). Security and usability: Analysis and evaluation. *2010 International Conference on Availability, Reliability and Security*, 275–282.

Kaiser, C., Stocker, A., Viscusi, G., Fellmann, M., & Richter, A. (2021). Conceptualising value creation in data-driven services: The case of vehicle data. *International Journal of Information Management*, *59*, 102335.

Kallinikos, J., Aaltonen, A., & Marton, A. (2013). The Ambivalent Ontology of Digital Artifacts. *MIS Quarterly*, *37*(2), 357–370.

Kallio, H., Pietilä, A.-M., Johnson, M., & Kangasniemi, M. (2016). Systematic methodological review: Developing a framework for a qualitative semi-structured interview guide. *Journal of Advanced Nursing*, *72*(12), 2954–2965. https://doi.org/10.1111/jan.13031

Kamm, L., Bogdanov, D., Laur, S., & Vilo, J. (2013). A new way to protect privacy in large-scale genome-wide association studies. *Bioinformatics*, *29*(7), 886–893. https://doi.org/10.1093/bioinformatics/btt066

Kamm, L., & Willemson, J. (2015). Secure floating point arithmetic and private satellite collision analysis. *International Journal of Information Security*, *14*(6), 531–548. https://doi.org/10.1007/s10207-014-0271-8

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

Kato, N., Murakami, Y., Endo, T., & Nawa, K. (2016). Study on privacy setting acceptance of the drivers for the data utilization on the car. *2016 14th Annual Conference on Privacy, Security and Trust (PST)*, 287–293.

Kaufmann, N., Schulze, T., & Veit, D. (2011). More than fun and money: Worker motivation in crowdsourcing-a study on Mechanical Turk. *AMCIS 2011 Proceedings*. https://aisel.aisnet.org/amcis2011_submissions/340

Kehr, F., Kowatsch, T., Wentzel, D., & Fleisch, E. (2015). Blissfully ignorant: The effects of general privacy concerns, general institutional trust, and affect in the privacy calculus. *Information Systems Journal*, *25*(6), 607–635.

Keith, M. J., Thompson, S. C., Hale, J., Lowry, P. B., & Greer, C. (2013). Information disclosure on mobile devices: Re-examining privacy calculus with actual user behavior. *International Journal of Human-Computer Studies*, *71*(12), 1163–1173.

Kembro, J., Näslund, D., & Olhager, J. (2017). Information sharing across multiple supply chain tiers: A Delphi study on antecedents. *International Journal of Production Economics*, *193*, 77–86.

Kerber, W. (2018). Data Governance in Connected Cars: The Problem of Access to In-Vehicle Data. *Journal of Intellectual Property, Information Technology and Electronic Commerce Law*, *9*, 310.

Khurana, M., Mishra, P., & Singh, A. R. (2011). Barriers to Information Sharing in Supply Chain of Manufacturing Industries. *International Journal of Manufacturing Systems*, *1*, 9–29. https://doi.org/10.3923/ijmsaj.2011.9.29

Kim, M., & Choi, B. R. (2022). The Impact of Privacy Control on Users' Intention to Use Smart Home Internet of Things (IoT) Services. *Asia Marketing Journal*, *24*(1), 4.

Kirkpatrick, K. (2021). Monetizing your personal data. *Communications of the ACM*, *65*(1), 17–19. https://doi.org/10.1145/3495563

Klein, T., & Verhulst, S. (2017). *Access to New Data Sources for Statistics: Business Models and Incentives for the Corporate Sector* (SSRN Scholarly Paper ID 3141446). Social Science Research Network. https://doi.org/10.2139/ssrn.3141446

Koch, K., Krenn, S., Pellegrino, D., & Ramacher, S. (2021). Privacy-Preserving Analytics for Data Markets Using MPC. In M. Friedewald, S. Schiffner, & S. Krenn (Eds.), *Privacy and Identity Management* (pp. 226–246). Springer International Publishing. https://doi.org/10.1007/978-3-030-72465-8_13

Kolekofski, K. E., & Heminger, A. R. (2003). Beliefs and attitudes affecting intentions to share information in an organizational setting. *Information & Management*, *40*(6), 521–532. https://doi.org/10.1016/S0378-7206(02)00068-X

Koutroumpis, P., Leiponen, A., & Thomas, L. D. (2020). Markets for data. *Industrial and Corporate Change*, *29*(3), 645–660.

Krasnova, H., Spiekermann, S., Koroleva, K., & Hildebrand, T. (2010). Online social networks: Why we disclose. *Journal of Information Technology*, *25*(2), 109–125.

Kumaraguru, P., & Cranor, L. F. (2005). *Privacy indexes: A survey of Westin's studies*. Carnegie Mellon University, School of Computer Science, Institute for ….

Lapets, A., Dak Albab, K., Issa, R., Qin, L., Varia, M., Bestavros, A., & Jansen, F. (2019). Role-Based Ecosystem for the Design, Development, and Deployment of Secure Multi-Party Data Analytics Applications. *2019 IEEE Cybersecurity Development (SecDev)*, 129–140. https://doi.org/10.1109/SecDev.2019.00023

Lapets, A., Jansen, F., Albab, K. D., Issa, R., Qin, L., Varia, M., & Bestavros, A. (2018). Accessible Privacy-Preserving Web-Based Data Analysis for Assessing and Addressing Economic Inequalities. *Proceedings of the 1st ACM SIGCAS*

Conference on Computing and Sustainable Societies, 1–5. https://doi.org/10.1145/3209811.3212701

Li, H., Chen, Q., Zhu, H., Ma, D., Wen, H., & Shen, X. S. (2020). Privacy Leakage via De-Anonymization and Aggregation in Heterogeneous Social Networks. *IEEE Transactions on Dependable and Secure Computing*, *17*(2), 350–362. https://doi.org/10.1109/TDSC.2017.2754249

Li, J., Sikora, R., Shaw, M. J., & Tan, G. W. (2006). A strategic analysis of inter organizational information sharing. *Decision Support Systems*, *42*(1), 251–266.

Li, T., Lin, L., & Gong, S. (2019). AutoMPC: Efficient Multi-Party Computation for Secure and Privacy-Preserving Cooperative Control of Connected Autonomous Vehicles. *SafeAI@ AAAI*.

Li, T., & Slee, T. (2014). The effects of information privacy concerns on digitizing personal health records. *Journal of the Association for Information Science and Technology*, *65*(8), 1541–1554. https://doi.org/10.1002/asi.23068

Lindell, Y. (2020). Secure multiparty computation. *Communications of the ACM*, *64*(1), 86–96. https://doi.org/10.1145/3387108

Liu, C., Marchewka, J. T., Lu, J., & Yu, C.-S. (2005). Beyond concern—A privacy-trust-behavioral intention model of electronic commerce. *Information & Management*, *42*(2), 289–304. https://doi.org/10.1016/j.im.2004.01.003

Lumineau, F., Schilke, O., & Wang, W. (2023). Organizational Trust in the Age of the Fourth Industrial Revolution: Shifts in the Form, Production, and Targets of Trust. *Journal of Management Inquiry*, *32*(1), 21–34. https://doi.org/10.1177/10564926221127852

Lundy-Bryan, L. (2021). *Privacy Enhancing Technologies: Part 2- The Coming Age of Collaborative Computing* (Lunar Insight Series). Lunar Ventures.

Malhotra, N. K., Kim, S. S., & Agarwal, J. (2004). Internet Users' Information Privacy Concerns (IUIPC): The Construct, the Scale, and a Causal Model. *Information Systems Research*, *15*(4), 336–355. https://doi.org/10.1287/isre.1040.0032

Mangiò, F., Andreini, D., & Pedeliento, G. (2020). Hands off my data: Users' security concerns and intention to adopt privacy enhancing technologies. *Italian Journal of Marketing*, *2020*(4), 309–342.

Markos, E., Milne, G. R., & Peltier, J. W. (2017). Information sensitivity and willingness to provide continua: A comparative privacy study of the United States and Brazil. *Journal of Public Policy & Marketing*, *36*(1), 79–96.

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

Metzger, M. J. (2007). Communication privacy management in electronic commerce. *Journal of Computer-Mediated Communication*, *12*(2), 335–361.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. sage.

Mindel, V., Mathiassen, L., & Rai, A. (2018). The sustainability of polycentric information commons. *MIS Quarterly*, *42*(2), 607–632. https://doi.org/10.25300/MISQ/2018/14015

Mingers, J. (2001). Combining IS Research Methods: Towards a Pluralist Methodology. *Information Systems Research*, *12*(3), 240–259. https://doi.org/10.1287/isre.12.3.240.9709

Mingers, J. (2003). The paucity of multimethod research: A review of the information systems literature. *Information Systems Journal*, *13*(3), 233–249. https://doi.org/10.1046/j.1365-2575.2003.00143.x

Mišura, K., & Žagar, M. (2016). Data marketplace for Internet of Things. *2016 International Conference on Smart Systems and Technologies (SST)*, 255–260. https://doi.org/10.1109/SST.2016.7765669

Miyazaki, A. D., & Krishnamurthy, S. (2002). Internet seals of approval: Effects on online privacy policies and consumer perceptions. *Journal of Consumer Affairs*, *36*(1), 28–49.

Morgan, R. M., & Hunt, S. D. (1994). The Commitment-Trust Theory of Relationship Marketing. *Journal of Marketing*, *58*(3), 20–38. https://doi.org/10.1177/002224299405800302

Morrell, M., & Ezingeard, J.-N. (2002). Revisiting adoption factors of inter-organisational information systems in SMEs. *Logistics Information Management*.

Mosterd, L., Sobota, V. C. M., van de Kaa, G., Ding, A. Y., & de Reuver, M. (2021). Context dependent trade-offs around platform-to-platform openness: The case of the Internet of Things. *Technovation*, *108*, 102331. https://doi.org/10.1016/j.technovation.2021.102331

Mucha, T., & Seppala, T. (2020). *Artificial Intelligence Platforms–A New Research Agenda for Digital Platform Economy*.

Mukhopadhyay, S., de Reuver, M., & Bouwman, H. (2016). Effectiveness of control mechanisms in mobile platform ecosystem. *Telematics and Informatics*, *33*(3), 848–859. https://doi.org/10.1016/j.tele.2015.12.008

Müller, J. M., Veile, J. W., & Voigt, K.-I. (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

Muschalle, A., Stahl, F., Löser, A., & Vossen, G. (2013). Pricing Approaches for Data Markets. In M. Castellanos, U. Dayal, & E. A. Rundensteiner (Eds.), *Enabling Real-Time Business Intelligence* (pp. 129–144). Springer. https://doi.org/10.1007/978-3-642-39872-8_10

Naehrig, M., Lauter, K., & Vaikuntanathan, V. (2011). Can homomorphic encryption be practical? *Proceedings of the 3rd ACM Workshop on Cloud Computing Security Workshop*, 113–124. https://doi.org/10.1145/2046660.2046682

Naous, D., Kulkarni, V., Legner, C., & Garbinato, B. (2019). Information Disclosure in Location-based Services: An Extended Privacy Calculus Model. *ICIS*.

Narayanan, S., Marucheck, A. S., & Handfield, R. B. (2009). Electronic Data Interchange: Research Review and Future Directions*. *Decision Sciences*, *40*(1), 121–163. https://doi.org/10.1111/j.1540-5915.2008.00218.x

Nicolaou, A. I., & McKnight, D. H. (2006). Perceived Information Quality in Data Exchanges: Effects on Risk, Trust, and Intention to Use. *Information Systems Research*, *17*(4), 332–351. https://doi.org/10.1287/isre.1060.0103

Noble, A., Cohen, G., Crowcroft, J., Gascón, A., Oswald, M., & Sasse, A. (2019). *Protecting Privacy in Practice: The current use, development and limits of Privacy Enhancing Technologies in data analysis* [Technical report]. The Royal Society.

Nokkala, T., Salmela, H., & Toivonen, J. (2019, July 12). Data Governance in Digital Platforms. *AMCIS 2019 Proceedings*. https://aisel.aisnet.org/amcis2019/ebusiness/ebusiness/12

Noorian, Z., Iyilade, J., Mohkami, M., & Vassileva, J. (2014). Trust Mechanism for Enforcing Compliance to Secondary Data Use Contracts. *2014 IEEE 13th International Conference on Trust, Security and Privacy in Computing and Communications*, 519–526. https://doi.org/10.1109/TrustCom.2014.66

Ofe, H., Minnema, H., & de Reuver, M. (2022). The business value of privacy-preserving technologies: The case of multiparty computation in the telecom industry. *Digital Policy, Regulation and Governance*, *24*(6), 541–557. https://doi.org/10.1108/DPRG-10-2021-0132

Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology*, *45*(4), 867–872.

Opriel, S., Fraunhofer, I., Skubowius, G. E., Fraunhofer, I. M. L., & Lamberjohann, M. (2021). How Usage Control Fosters Willingness To Share Sensitive Data In Inter-Organizational Processes of Supply Chains. *International Scientific Symposium on Logistics*, 91.

Osterwalder, A., & Pigneur, Y. (2010). *Business model generation: A handbook for visionaries, game changers, and challengers*. John Wiley & Sons.

Otto, B. (2022). The Evolution of Data Spaces. In B. Otto, M. ten Hompel, & S. Wrobel (Eds.), *Designing Data Spaces: The Ecosystem Approach to Competitive Advantage* (pp. 3–15). Springer International Publishing. https://doi.org/10.1007/978-3-030-93975-5_1

Otto, B., & Jarke, M. (2019). Designing a multi-sided data platform: Findings from the International Data Spaces case. *Electronic Markets*, *29*(4), 561–580. https://doi.org/10.1007/s12525-019-00362-x

Otto, B., Steinbuß, S., Teuscher, A., & Lohmann, S. (2019). *Reference architecture model—International data spaces (Version 3.0)*. International Data Spaces Association. https://internationaldataspaces.org/wp-content/uploads/IDS-Reference-Architecture-Model-3.0-2019.pdf

Özyilmaz, K. R., Doğan, M., & Yurdakul, A. (2018). IDMoB: IoT Data Marketplace on Blockchain. *2018 Crypto Valley Conference on Blockchain Technology (CVCBT)*, 11–19. https://doi.org/10.1109/CVCBT.2018.00007

Pal, D., Funilkul, S., & Zhang, X. (2021). Should I Disclose My Personal Data? Perspectives From Internet of Things Services. *IEEE Access*, *9*, 4141–4157. https://doi.org/10.1109/ACCESS.2020.3048163

Palan, S., & Schitter, C. (2018). Prolific. Ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, *17*, 22–27.

Pavlou, P. A. (2003). Consumer acceptance of electronic commerce: Integrating trust and risk with the technology acceptance model. *International Journal of Electronic Commerce*, *7*(3), 101–134.

Pavlou, P. A., & Gefen, D. (2004). Building Effective Online Marketplaces with Institution-Based Trust. *Information Systems Research*, *15*(1), 37–59. https://doi.org/10.1287/isre.1040.0015

Pedersen, T. B., Saygın, Y., & Savas, E. (2007). *Secret Sharing vs. Encryption-based Techniques For Privacy Preserving Data Mining*.

Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, *70*, 153–163.

Petronia, M. (2020). *Multiparty Computation: The effect of multiparty computation on firms' willingness to contribute protected data* [Master's thesis, Delft University of Technology]. https://repository.tudelft.nl/islandora/object/uuid%3Ab0de4a4b-f5a3-44b8-baa4-a6416cebe26f

Popovič, A., Smith, H. J., Thong, J. Y. L., & Wattal, S. (2017). Information Privacy. In Ashley Bush & Arun Rai (Eds.), *MIS Quarterly Research Curations*. http://misq.org/research-curations

Praditya, D., Janssen, M., & Sulastri, R. (2017). Determinants of Business-to-Government Information Sharing Arrangements. *Electronic Journal of E-Government*, *15*(1), Article 1.

Priego, L. P., Osimo, D., & Wareham, J. D. (2019). Data Sharing Practice in Big Data Ecosystems. *SSRN Electronic Journal*.

Prolific. (2022). *Representative samples*. https://researcher-help.prolific.co/hc/en-gb/articles/360019236753-Representative-samples

Qin, L., Lapets, A., Jansen, F., Flockhart, P., Albab, K. D., Globus-Harris, I., Roberts, S., & Varia, M. (2019). From usability to secure computing and back again. *Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019)*, 191–210.

Rabin, M. O. (1981). *How to exchange secrets by oblivious transfer*. Technical report, Aiken Computation Laboratory. Harvard University.

Rajan, A., Qin, L., Archer, D. W., Boneh, D., Lepoint, T., & Varia, M. (2018). Callisto: A Cryptographic Approach to Detecting Serial Perpetrators of Sexual Misconduct. *Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies*, 1–4. https://doi.org/10.1145/3209811.3212699

Ram, P., Markkula, J., Friman, V., & Raz, A. (2018). Security and privacy concerns in connected cars: A systematic mapping study. *2018 44th Euromicro Conference on Software Engineering and Advanced Applications (SEAA)*, 124–131.

Ramachandran, G. S., Radhakrishnan, R., & Krishnamachari, B. (2018). Towards a Decentralized Data Marketplace for Smart Cities. *2018 IEEE International Smart Cities Conference (ISC2)*, 1–8. https://doi.org/10.1109/ISC2.2018.8656952

Ranerup, A., Henriksen, H. Z., & Hedman, J. (2016). An analysis of business models in Public Service Platforms. *Government Information Quarterly*, *33*(1), 6–14. https://doi.org/10.1016/j.giq.2016.01.010

Ratnasingam, P., Pavlou, P. A., & Tan, Y.-H. (2002). *The Importance of Technology Trust for B2B Electronic Commerce* (SSRN Scholarly Paper ID 2380727). Social Science Research Network. https://papers.ssrn.com/abstract=2380727

Recker, J. (2021). *Scientific Research in Information Systems: A Beginner's Guide*. Springer International Publishing. https://doi.org/10.1007/978-3-030-85436-2

Reimsbach-Kounatze, C. (2021). *Enhancing access to and sharing of data: Striking the balance between openness and control over data*. 25–68. https://doi.org/10.5771/9783748924999-25

Richter, H., & Slowinski, P. R. (2019). The data sharing economy: On the emergence of new intermediaries. *IIC-International Review of Intellectual Property and Competition Law*, *50*(1), 4–29.

Robey, D., Im, G., & Wareham, J. D. (2008). Theoretical foundations of empirical research on interorganizational systems: Assessing past contributions and guiding future directions. *Journal of the Association for Information Systems*, *9*(9), 4.

Roman, D., Nikolov, N., Elvesæter, B., Soylu, A., Prodan, R., Kimovski, D., Marrella, A., Leotta, F., Benvenuti, D., & Matskin, M. (2021). DataCloud: Enabling the Big Data Pipelines on the Computing Continuum. *Research Challenges in Information Science: 15th International Conference, RCIS 2021, Limassol, Cyprus, May 11–14, 2021, Proceedings*.

Roman, D., & Vu, K. (2019). Enabling Data Markets Using Smart Contracts and Multi-party Computation. In W. Abramowicz & A. Paschke (Eds.), *Business Information Systems Workshops* (pp. 258–263). Springer International Publishing. https://doi.org/10.1007/978-3-030-04849-5_23

Roseman Labs. (2022). *Easier, safer and more collaboration on healthcare data*. Roseman Labs. https://rosemanlabs.com/blog/zorg_whitepaper.html

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 I. Goldberg & T. Moore (Eds.), *Financial Cryptography and Data Security* (pp. 605–623). Springer International Publishing. https://doi.org/10.1007/978-3-030-32101-7_35

Sayogo, D. S., Zhang, J., Pardo, T. A., Tayi, G. K., Hrdinova, J., Andersen, D. F., & Luna-Reyes, L. F. (2014). Going Beyond Open Data: Challenges and Motivations for Smart Disclosure in Ethical Consumption. *Journal of Theoretical and Applied Electronic Commerce Research*, *9*(2), 1–16. https://doi.org/10.4067/S0718-18762014000200002

Scerri, S., Tuikka, T., de Vallejo, I. L., & Curry, E. (2022). Common European Data Spaces: Challenges and Opportunities. In E. Curry, S. Scerri, & T. Tuikka (Eds.), *Data Spaces: Design, Deployment and Future Directions* (pp. 337–357). Springer International Publishing. https://doi.org/10.1007/978-3-030-98636-0_16

Schomakers, E.-M., Lidynia, C., & Ziefle, M. (2020). All of me? Users' preferences for privacy-preserving data markets and the importance of anonymity. *Electronic Markets*, *30*(3), 649–665.

Schomm, F., Stahl, F., & Vossen, G. (2013). Marketplaces for data: An initial survey. *ACM SIGMOD Record*, *42*(1), 15–26.

Schwinghammer, R., Neuburger, R., & Hess, T. (2022). Towards an Understanding of the Tensions between Data Privacy and Data Monetization: Uncovering the Sweet Spot between the Tensions. *PACIS 2022 Proceedings*.

Sekaran, U., & Bougie, R. (2016). *Research Methods For Business: A Skill Building Approach*. John Wiley & Sons.

Shamir, A. (1979). How to share a secret. *Communications of the ACM*, *22*(11), 612–613. https://doi.org/10.1145/359168.359176

Shen, W., Qin, J., Yu, J., Hao, R., & Hu, J. (2019). Enabling Identity-Based Integrity Auditing and Data Sharing With Sensitive Information Hiding for Secure Cloud Storage. *IEEE Transactions on Information Forensics and Security*, *14*(2), 331–346. https://doi.org/10.1109/TIFS.2018.2850312

Sheng, S., Broderick, L., Koranda, C. A., & Hyland, J. J. (2006). Why johnny still can't encrypt: Evaluating the usability of email encryption software. *Symposium On Usable Privacy and Security*, 3–4.

Smith, H. J., Dinev, T., & Xu, H. (2011). Information privacy research: An interdisciplinary review. *MIS Quarterly*, 989–1015.

Smith, H. J., Milberg, S. J., & Burke, S. J. (1996). Information privacy: Measuring individuals' concerns about organizational practices. *MIS Quarterly*, 167–196.

Son, J.-Y., Tu, L., & Benbasat, I. (2006). A Descriptive Content Analysis of Trust-Building Measures in B2B Electronic Marketplaces. *Communications of the Association for Information Systems*, *18*. https://doi.org/10.17705/1CAIS.01806

Sorescu, A. (2017). Data-Driven Business Model Innovation. *Journal of Product Innovation Management*, *34*(5), 691–696. https://doi.org/10.1111/jpim.12398

Sousa, P. R., Antunes, L., & Martins, R. (2018). The Present and Future of Privacy-Preserving Computation in Fog Computing. In A. M. Rahmani, P. Liljeberg, J.-S. Preden, & A. Jantsch (Eds.), *Fog Computing in the Internet of Things: Intelligence at the Edge* (pp. 51–69). Springer International Publishing. https://doi.org/10.1007/978-3-319-57639-8_4

Spiekermann, M. (2019). Data marketplaces: Trends and monetisation of data goods. *Intereconomics*, *54*(4), 208–216.

Spiekermann, S. (2005). *Perceived Control: Scales for Privacy in Ubiquitous Computing* (SSRN Scholarly Paper ID 761109). Social Science Research Network. https://doi.org/10.2139/ssrn.761109

Spiekermann, S., & Novotny, A. (2015). A vision for global privacy bridges: Technical and legal measures for international data markets. *Computer Law & Security Review*, *31*(2), 181–200. https://doi.org/10.1016/j.clsr.2015.01.009

Spini, G., Mancini, E., Attema, T., Abspoel, M., de Gier, J., Fehr, S., Veugen, T., van Heesch, M., Worm, D., De Luca, A., Cramer, R., & Sloot, P. M. A. (2022). New Approach to Privacy-Preserving Clinical Decision Support Systems for HIV Treatment. *Journal of Medical Systems*, *46*(12), 84. https://doi.org/10.1007/s10916-022-01851-x

Stahl, F., Schomm, F., Vossen, G., & Vomfell, L. (2016). A classification framework for data marketplaces. *Vietnam Journal of Computer Science*, *3*(3), 137–143.

Stefansson, G. (2002). Business-to-business data sharing: A source for integration of supply chains. *International Journal of Production Economics*, *75*(1), 135–146. https://doi.org/10.1016/S0925-5273(01)00187-6

Subramanian, H. (2017). Decentralized blockchain-based electronic marketplaces. *Communications of the ACM*, *61*(1), 78–84. https://doi.org/10.1145/3158333

Sun, S., Cegielski, C. G., Jia, L., & Hall, D. J. (2018). Understanding the Factors Affecting the Organizational Adoption of Big Data. *Journal of Computer Information Systems*, *58*(3), 193–203. https://doi.org/10.1080/08874417.2016.1222891

Susha, I., Flipsen, M., Agahari, W., & de Reuver, M. (2020). Towards Generic Business Models of Intermediaries in Data Collaboratives: From Gatekeeping to Data Control. In G. Viale Pereira, M. Janssen, H. Lee, I. Lindgren, M. P. Rodríguez Bolívar, H. J. Scholl, & A. Zuiderwijk (Eds.), *Electronic Government* (pp. 304–315). Springer International Publishing. https://doi.org/10.1007/978-3-030-57599-1_23

Susha, I., Janssen, M., & Verhulst, S. (2017). Data collaboratives as “bazaars”? A review of coordination problems and mechanisms to match demand for data with

supply. *Transforming Government: People, Process and Policy*, *11*(1), 157–172. https://doi.org/10.1108/TG-01-2017-0007

Svahn, F., Mathiassen, L., & Lindgren, R. (2017). Embracing digital innovation in incumbent firms: How volvo cars managed competing concerns. *MIS Quarterly*, *41*(1), 239–253. https://doi.org/10.25300/MISQ/2017/41.1.12

Svensson, R. B., & Taghavianfar, M. (2020). Toward Becoming a Data-Driven Organization: Challenges and Benefits. In F. Dalpiaz, J. Zdravkovic, & P. Loucopoulos (Eds.), *Research Challenges in Information Science* (pp. 3–19). Springer International Publishing. https://doi.org/10.1007/978-3-030-50316-1_1

Teece, D. J. (2010). Business models, business strategy and innovation. *Long Range Planning*, *43*(2–3), 172–194.

Teece, D. J. (2018). Business models and dynamic capabilities. *Long Range Planning*, *51*(1), 40–49. https://doi.org/10.1016/j.lrp.2017.06.007

Teece, D. J., & Pisano, G. (1994). The Dynamic Capabilities of Firms: An Introduction. *Industrial and Corporate Change*, *3*(3), 537–556. https://doi.org/10.1093/icc/3.3.537-a

Tilson, D., Lyytinen, K., & Sørensen, C. (2010). Research Commentary—Digital Infrastructures: The Missing IS Research Agenda. *Information Systems Research*, *21*(4), 748–759. https://doi.org/10.1287/isre.1100.0318

Tiwana, A. (2014). *Platform Ecosystems: Aligning Architecture, Governance, and Strategy*. Elsevier. https://doi.org/10.1016/C2012-0-06625-2

Tiwana, A., Konsynski, B., & Bush, A. A. (2010). Research Commentary—Platform Evolution: Coevolution of Platform Architecture, Governance, and Environmental Dynamics. *Information Systems Research*, *21*(4), 675–687. https://doi.org/10.1287/isre.1100.0323

Tõldsepp, K., Pruulmann-Vengerfeldt, P., & Laud, P. (2012). *Deliverable d1.2: Requirements specification based on the interviews*. UaESMC project. http://usable-security.eu/files/d12final.pdf

Van Alstyne, M. W., & Lenart, A. (2020). Using data and respecting users. *Communications of the ACM*, *63*(11), 28–30. https://doi.org/10.1145/3423998

Van Alstyne, M. W., Petropoulos, G., Parker, G., & Martens, B. (2021). "In situ" data rights. *Communications of the ACM*, *64*(12), 34–35. https://doi.org/10.1145/3491270

van de Ven, M., Abbas, A. E., Roosenboom-Kwee, Z., & de Reuver, G. A. (2021). Creating a Taxonomy of Business Models for Data Marketplaces: 34th Bled eConference – Digital Support from Crisis to Progressive Change. *Proceedings 34th Bled EConference – Digital Support from Crisis to Progressive Change*, 313–325. https://doi.org/10.18690/978-961-286-385-9.23

van den Broek, T., & van Veenstra, A. F. (2018). Governance of big data collaborations: How to balance regulatory compliance and disruptive innovation. *Technological Forecasting and Social Change*, *129*, 330–338.

van der Wel, I. (2021). *Multi-Party Computation: Van wantrouwen naar vertrouwen in het delen van persoonsgegevens op datamarktplaatsen* [Bachelor's thesis]. Delft University of Technology.

van Egmond, M. B., Spini, G., van der Galien, O., IJpma, A., Veugen, T., Kraaij, W., Sangers, A., Rooijakkers, T., Langenkamp, P., Kamphorst, B., van de L'Isle, N., & Kooij-Janic, M. (2021). Privacy-preserving dataset combination and Lasso regression for healthcare predictions. *BMC Medical Informatics and Decision Making*, *21*(1), 266. https://doi.org/10.1186/s12911-021-01582-y

van Schaik, P., Jansen, J., Onibokun, J., Camp, J., & Kusev, P. (2018). Security and privacy in online social networking: Risk perceptions and precautionary behaviour. *Computers in Human Behavior*, *78*, 283–297. https://doi.org/10.1016/j.chb.2017.10.007

Venkatesh, V., Brown, S. A., & Bala, H. (2013). Bridging the Qualitative-Quantitative Divide: Guidelines for Conducting Mixed Methods Research in Information Systems. *MIS Quarterly*, *37*(1), 21–54.

Verschuren, P., & Doorewaard, H. (2010). *Designing a research project* (Vol. 2). Eleven International Publishing.

Vimercati, S. D. C. di, Foresti, S., Livraga, G., & Samarati, P. (2021). Toward Owners' Control in Digital Data Markets. *IEEE Systems Journal*, *15*(1), 1299–1306. https://doi.org/10.1109/JSYST.2020.2970456

Virkar, S., Viale Pereira, G., & Vignoli, M. (2019). Investigating the Social, Political, Economic and Cultural Implications of Data Trading. In I. Lindgren, M. Janssen, H. Lee, A. Polini, M. P. Rodríguez Bolívar, H. J. Scholl, & E. Tambouris (Eds.), *Electronic Government* (pp. 215–229). Springer International Publishing. https://doi.org/10.1007/978-3-030-27325-5_17

Volgushev, N., Schwarzkopf, M., Getchell, B., Varia, M., Lapets, A., & Bestavros, A. (2019). Conclave: Secure multi-party computation on big data. *Proceedings of the Fourteenth EuroSys Conference 2019*, 1–18. https://doi.org/10.1145/3302424.3303982

Walsh, J. M., Varia, M., Cohen, A., Sellars, A., & Bestavros, A. (2022). Multi-Regulation Computing: Examining the Legal and Policy Questions That Arise From Secure Multiparty Computation. *Symposium on Computer Science and Law (CSLAW '22)*.

Wang, T., Duong, T. D., & Chen, C. C. (2016). Intention to disclose personal information via mobile applications: A privacy calculus perspective. *International Journal of Information Management*, *36*(4), 531–542.

Weick, K. E. (1995). What Theory is Not, Theorizing Is. *Administrative Science Quarterly*, *40*(3), 385–390. https://doi.org/10.2307/2393789

Wessels, N., Gerlach, J., & Wagner, A. (2019). To Sell or not to Sell – Antecedents of Individuals' Willingness-to-Sell Personal Information on Data-Selling Platforms. *ICIS 2019 Proceedings*. https://aisel.aisnet.org/icis2019/cyber_security_privacy_ethics_IS/cyber_security_privacy/34

Westin, A. F. (1968). Privacy and freedom. *Washington and Lee Law Review*, *25*(1), 166.

White, A., Daniel, E., Ward, J., & Wilson, H. (2007). The adoption of consortium B2B e-marketplaces: An exploratory study. *The Journal of Strategic Information Systems*, *16*(1), 71–103.

Whitten, A., & Tygar, J. D. (1999). Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0. *USENIX Security Symposium*, *348*, 169–184.

Wiener, M., Mahring, M., Remus, U., & Saunders, C. (2016). Control Configuration and Control Enactment in Information Systems Projects: Review and Expanded Theoretical Framework. *Management Information Systems Quarterly*, *40*(3), 741–774.

Wixom, B. H., & Ross, J. W. (2017). How to Monetize Your Data. *MIT Sloan Management Review*, *58*(3), n/a-13.

Xiong, J., Bi, R., Tian, Y., Liu, X., & Wu, D. (2022). Toward Lightweight, Privacy-Preserving Cooperative Object Classification for Connected Autonomous Vehicles. *IEEE Internet of Things Journal*, *9*(4), 2787–2801. https://doi.org/10.1109/JIOT.2021.3093573

Xu, H., & Dinev, T. (2022). Guest Editorial: Reflections on the 2021 Impact Award: Why Privacy Still Matters. *MIS Quarterly*, *46*(4), xx–xxxii.

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), 1.

Yakoubov, S. (2017). *A Gentle Introduction to Yao's Garbled Circuits*. https://web.mit.edu/sonka89/www/papers/2017ygc.pdf

Yao, A. C. (1982). Protocols for secure computations. *23rd Annual Symposium on Foundations of Computer Science (Sfcs 1982)*, 160–164. https://doi.org/10.1109/SFCS.1982.38

Yao, A. C. (1986). How to generate and exchange secrets. *27th Annual Symposium on Foundations of Computer Science (Sfcs 1986)*, 162–167. https://doi.org/10.1109/SFCS.1986.25

Yee, K.-P. (2002). User interaction design for secure systems. *International Conference on Information and Communications Security*, 278–290.

Zaheer, N., & Trkman, P. (2017). An information sharing theory perspective on willingness to share information in supply chains. *The International Journal of Logistics Management*, *28*(2), 417–443. https://doi.org/10.1108/IJLM-09-2015-0158

Zavolokina, L., Spychiger, F., Tessone, C. J., & Schwabe, G. (2018). *Incentivizing data quality in blockchains for inter-organizational networks–learning from the digital car dossier*.

Zhao, C., Zhao, S., Zhao, M., Chen, Z., Gao, C.-Z., 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

Zhao, L., Lu, Y., & Gupta, S. (2012). Disclosure Intention of Location-Related Information in Location-Based Social Network Services. *International Journal of Electronic Commerce*, *16*(4), 53–90. https://doi.org/10.2753/JEC1086-4415160403

Zhou, T. (2011). The impact of privacy concern on user adoption of location-based services. *Industrial Management & Data Systems*, *111*(2), 212–226. https://doi.org/10.1108/02635571111115146

Zhou, T. (2017). Understanding location-based services users' privacy concern: An elaboration likelihood model perspective. *Internet Research*.

Ziefle, M., Halbey, J., & Kowalewski, S. (2016). Users' Willingness to Share Data on the Internet: Perceived Benefits and Caveats. *IoTBD*, 255–265.

Zobiri, F., Gama, M., Nikova, S., & Deconinck, G. (2022). A Privacy-Preserving Three-Step Demand Response Market Using Multi-Party Computation. *2022 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, 1–5. https://doi.org/10.1109/ISGT50606.2022.9817546

Zöll, A., Olt, C., & Buxmann, P. (2021). Privacy-sensitive Business Models: Barriers of Organizational Adoption of Privacy-Enhancing Technologies. *ECIS 2021 Research Papers*. https://aisel.aisnet.org/ecis2021_rp/34

Zott, C., & Amit, R. (2017). Business Model Innovation: How to Create Value in a Digital World. *GfK Marketing Intelligence Review*, *9*(1), 18–23.

Zrenner, J., Möller, F. O., Jung, C., Eitel, A., & Otto, B. (2019). Usage control architecture options for data sovereignty in business ecosystems. *Journal of Enterprise Information Management*.

Appendices

A. Interview protocol with MPC experts and practitioners

Introduction

Thank you for your participation in this interview. The purposes of this interview are to clarify the conceptual foundations of MPC and other relevant techniques as well as exploring the possibility of MPC implementation in data markets.

Before we start the interview, I will briefly explain the concept of data markets using a short PowerPoint presentation so that we have the same understanding of data markets. Then, halfway through the interview, I will illustrate one possible use-case of MPC in data markets based on my interpretation.

Recording instructions

Our conversation will be recorded with your consent. The purpose of recording is to carry on an attentive conversation while getting the details at the same time. The interview recording will be confidential and will only be used for academic purposes.

Interview questions

Note: Question 1-9 emerged after the fourth interview with MPC experts. To narrow down the focus, we stopped asking question 12-18 from the fifth interview onwards.

1. What are the differences and similarities between MPC and other techniques such as homomorphic encryption, private set intersection, and Blockchain? Which one is the most relevant in the context of data markets?
2. When you talk to business actors, do they want to know the detailed process of how MPC works? Or maybe just the simple one is enough?

3. What are the value propositions offered by MPC? How do you explain the value proposition of MPC to business actors? Why should they use MPC for data sharing?
    a. What about the value proposition in terms of giving more control over data?
    b. What about the value proposition in terms of trust toward data requesters/consumers?
    c. What are the risks that can be tackled using MPC?
4. How about GDPR compliance? Does data sharing with MPC comply with GDPR? What about consent to use data for other purposes? And anonymization?
5. What are the requirements to perform computation? Should the data be standardized?
6. How do data providers input their data? Web-based, or apps, or embedded automatically in their database?
7. Can you think of a real-world use case of MPC that is already implemented, especially in the mobility/automotive industry? Or maybe in data markets?
8. How do you see the potential of implementing MPC in data markets? Is it possible to create data markets that follow the MPC protocol?
9. How about the verification/validation of MPC? How can we be sure that the process is done with MPC?
10. What does the architecture of data markets look like with MPC in place? Do we configure/install MPC in each data provider or only in a central system/intermediaries?
11. One important aspect of MPC is that you must always be online (synchronous), but this might not be the case in data markets. Some data providers might not have the capacity for that and prefer to input their data whenever they have time. Is it possible to perform the computation asynchronously?
12. Which statement is correct?
    a. With MPC, intermediaries (third-party) are not needed for data sharing.

b. With MPC, it is still possible to have intermediaries, but there is no need to trust them.
c. None of the above is correct.

13. Another assumption about MPC is that the data is not exchanged and only "exposed", meaning that the raw data will not leave the owner. Is this correct?
14. What kind of data can we "share" with MPC? Is it only a numerical dataset (e.g., a list of location coordinates, mobile phone numbers, distance covered, etc.)? How about non-numeric data (e.g., health records, text)?
15. Can we do computation with dynamic, real-time data? Or only static data/datasets?
16. Is the MPC implementation only limited to data sharing in a clear usage context and pre-defined use case? This means that we know who are the data buyer and the purpose of sharing/using the data. (**Note:** Think about, for instance, heat maps of popular pick-up spots for ride-hailing vehicles. The usage context is clear (e.g., to create heat maps), and the actors involved are also clear (ride-hailing companies and transport authorities)).
17. In data markets, we typically do not know who will buy the data and what they will do with the data. Does MPC provide a suitable solution in data markets for this setting.
18. Another assumption of MPC is that computation involves at least two parties that "expose" their data to get useful insights. Is this true? How about in the data market? Would it be possible for data buyers to get "data insights" without "exposing" their data? Do all parties (including data buyers) need to participate in the computation?
19. Do you have other remarks that you want to share?

Thank you note

Once again, thank you for your time in participating in this interview. I will guarantee the full confidentiality of your personal information. The recording will be transcribed and send back to you for your review.

Regards,
Wirawan Agahari (Aga) - PhD researcher, TU Delft

B. Interview protocol with business actors in the automotive industry

Introduction

Thank you for your participation in this interview. The purposes of this interview are to explore:

1. why companies are willing/unwilling to participate in data sharing facilitated by data markets;
2. whether a privacy-preserving technology called Multi-Party Computation (MPC) could create more (or less) feeling of control over company data while participating in data sharing via data markets.

The context of this interview is about **business-to-business data sharing**. This means that, unless stated otherwise:

1. your answers are referring to **the point of view of your company**.
2. data sharing refers to **data sharing with other business actors via data markets**.
3. We will focus on **data possessed by your company that might be valuable for other companies, but not being shared at the moment**.

Before we start the interview, I will briefly explain the concept of data markets using a short PowerPoint presentation. By doing this, I can make sure that we have the same understanding of data markets.

Recording instructions

Our conversation will be recorded with your consent. The purpose of recording is to carry on an attentive conversation while getting the details at the same time. The interview recording will be confidential and will only be used for academic purposes.

Interview questions

1. How familiar are you with data markets? Has your company ever share data via data markets?
2. Do you have data that you think is valuable for other business actors, but which you are not sharing now via data markets? What kind of data is it, and why your company is not sharing this data?

Note: In the rest of the interview, **we will focus on those data, so please keep them in mind.** Please assume also that you will get something back from sharing these data, for instance, a payment through data markets or access to other actors' data that is valuable to you.

3. What risks might emerge if your company would start to share those data in data markets? How do these risks play a role in your company's decision to share those data?
4. What about trust towards other business actors? How does it play a role in your company's decision to participate in data sharing?
5. Besides risks and trust, are there other reasons why you are not sharing those data at the moment?
6. What kind of control over those data would you want while sharing in data markets, and why?

Before we move to the next questions, I will briefly explain Multi-Party Computation (MPC) to help you understand how it works. Then, I will illustrate how MPC could be

implemented in data markets. Therefore, please answer the following questions **under the assumption of MPC implementation in data markets**. The previously described context still applies, unless stated otherwise. **Remember to focus on the data as answered in Question 2.**

7. How familiar are you with MPC technology?
8. With MPC in place, how likely would you share those data in data markets? Why?
9. With MPC in place, do you expect to:

    a. Have more or less control over those data while sharing in data markets? Why?
    b. Encounter more or fewer risks of sharing those data in data markets? Why?
    c. Have more or less trust towards other business actors while sharing those data in data markets? Why?

10. Do you have other remarks that you want to share?

<u>Thank you note</u>

Once again, thank you for your time in participating in this interview. I will guarantee the full confidentiality of your personal information. The recording will be transcribed and send back to you for your review.

Regards,
Wirawan Agahari (Aga) - PhD researcher, TU Delft

C. Experiment setup

Introduction

Dear participants,

Thank you for making time to take part in this survey. Your contribution is greatly appreciated!

You are being invited to participate in a research study titled **"Does multi-party computation (MPC) enhance control in data sharing through data marketplaces? An experimental study."** This study is conducted by Wirawan Agahari, a Ph.D. researcher at Delft University of Technology (TU Delft), supervised by Dr. Ir. Mark de Reuver and Dr.-Ing. Tobias Fiebig. The Human Research Ethics Committee of TU Delft approved this study.

The purpose of this research study is **to investigate the impact of a privacy-enhancing technology called Multi-Party Computation (MPC) on individuals' control over data and willingness to share data in data marketplaces.** This study will take you **approximately 20 minutes** to complete. Your answer will remain anonymous, cannot be traced back to you, and will only be used for research purposes. **Your participation in this study is entirely voluntary, and you can withdraw at any time.**

We believe there are no known risks associated with this research study; however, as with any online-related activity, the risk of a breach is always possible. To the best of our ability, your answers in this study will remain confidential. We will minimize any risks by **only storing data at remote, protected storage at TU Delft, only accessible by project members, as well as abstaining from both distributing data to others or retrieving it on personal devices.**

An anonymized, non-reducible version of this dataset will be publicly available through 4TU Research Data Repository. Before publication, we will drop any personal data.

For any further inquiries, please refer to:

Wirawan Agahari (Aga)
Ph.D. researcher
Delft University of Technology
w.agahari@tudelft.nl

Please check the first box to give permission to process your data for this research:

- *I acknowledge that I have read and understood this introduction, and **I hereby give consent** that my survey data will be processed for this research.*
- ***I do not consent**, and I do not wish to participate in this study.*

Case description: sharing driving data from the connected cars via data marketplaces

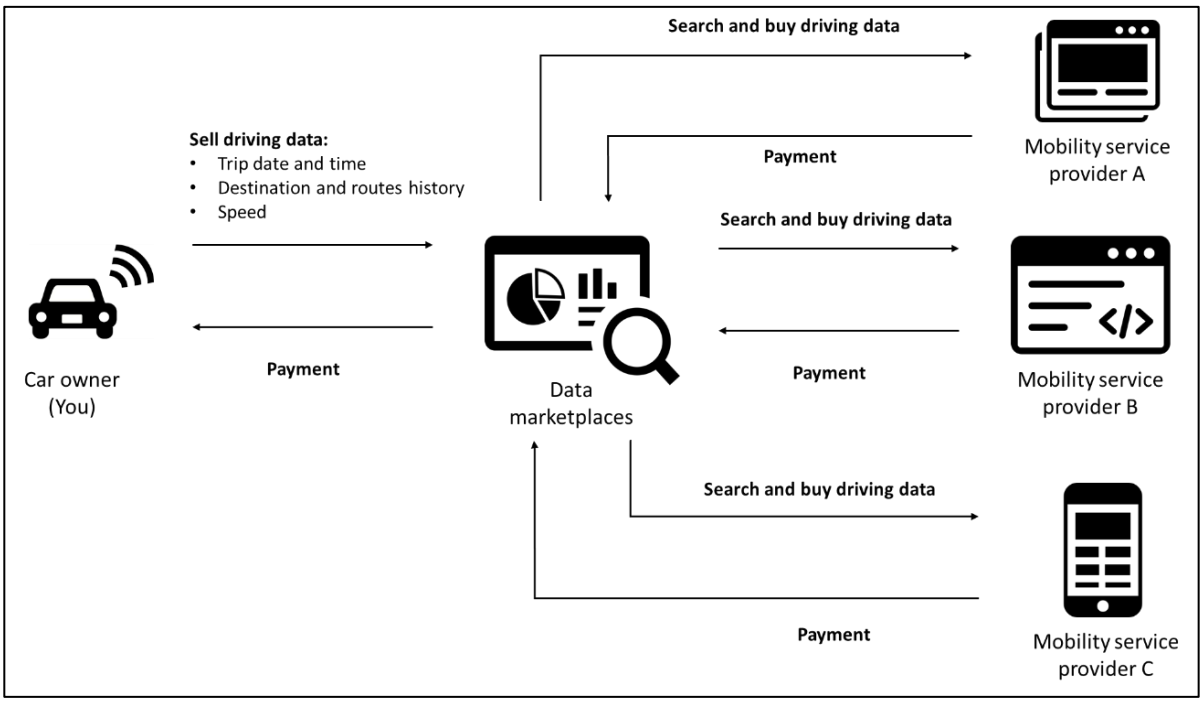
Imagine that you own a car that you regularly use to travel to work or other places. The car is connected to the Internet. Then, suppose that several mobility start-up companies want to offer various applications to improve driving behaviour and ultimately become better and safer drivers. **You do not know those companies, and you have not used their applications before.**

To achieve their goals, those companies would like to buy the data from your car. You can sell your car data to them via a **data marketplace:** a website where you can sell your car data to those mobility start-up companies. Please assume that it takes you very little effort to share your data on the data marketplace; all you need to do is register and click "upload."

To offer their services, those companies would like to **buy the following data from you** on the data marketplace:

- **Trip date and time:** when are you driving to a particular destination
- **Destination and routes history:** GPS information on where you are driving
- **Speed:** how fast are you driving

A simplified illustration is provided in the figure below (you can zoom the image if you are using a smartphone/tablet).



In the following questions, we will present a drawing of the data marketplace. Then, we will ask your perception on whether you would offer your data to those mobility start-up companies through the data marketplace based on the design presented. You may assume that when you make your data available, those companies would be able to **describe your driving behaviour** and **suggest better and safer driving advice to you (if you use their apps in the future).** Furthermore, by selling your data, **you will get financial compensation.**

Treatment 1: Trusted Third Party (TTP)

A technology to share your confidential information in a data marketplace is a **Trusted Third Party (TTP)**. The TTP stores, processes, and analyses your data on a central system.

Example:

Four colleagues go out to dinner, and they have agreed that the person with the highest salary should pay the bill. However, for privacy reasons, they do not want to disclose their salary to each other. So, they decided to ask the waiter to be the Trusted Third Party (TTP). They individually tell their salary to the waiter. Then, the waiter gives the bill to the colleague who has the highest salary, without the colleagues knowing what their salary is from each other. As a TTP, the waiter will know all of their salaries and compare them before giving the bill.

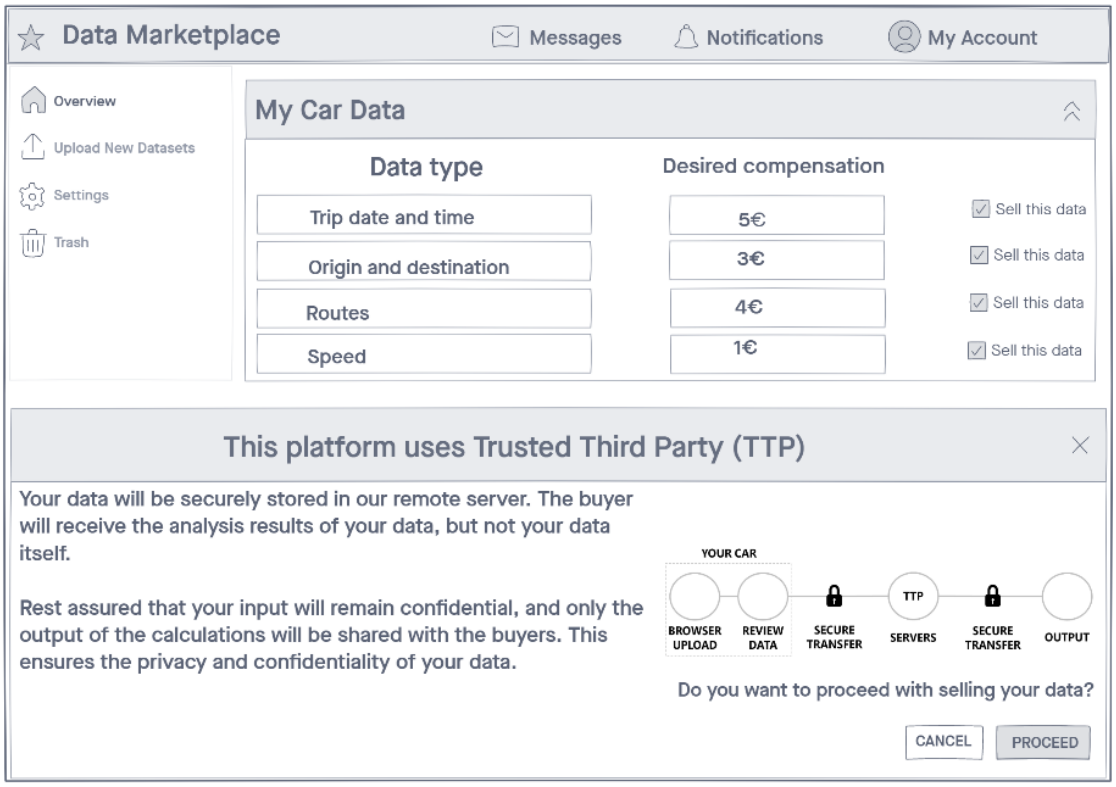
In short: TTP has access to the data of all data providers. But it will tell data buyers only the answers to the questions that data buyers have.

In this design, you can share your driving data via a **TTP-based data marketplace**. The data marketplace is managed by a platform operator that is also responsible for processing all data.

After you agree to sell your driving data, **it will be stored in the remote server of data marketplaces.** Then, when the mobility service providers buy your driving data, they have to tell the platform operator which analyses they want to perform. After that, the platform operator performs the analysis and sends the results to those companies (as data buyers). This way, those companies only see the analysis results based on your driving data and not your driving data itself. **Only the platform operator can view your data because it performs the analyses.**

A screenshot preview of the TTP-based data marketplace is presented below (you can zoom the image if you are using a smartphone/tablet).

To convince you that the analysis of your data is confidential, a disclaimer has been added that explains how TTP works.



- ✓ Please tick here if you have **read the description** and **understand the scenario** of sharing driving data through TTP-based data marketplaces

You have read the scenario of sharing driving data through **TTP-based data marketplaces.** The code for this scenario is **A1**.

Please enter the code here:

Treatment 2: Multi-Party Computation (MPC)

A technology to share your confidential information in a data marketplace is **Multi-Party Computation (MPC)**. With MPC, your data is encrypted, meaning that **your data is being changed so that it cannot be read without knowledge of the secret key that has been used to change your data.** Your encrypted data is then shared and can be used to perform meaningful calculations. Think of MPC as a black box that calculates a specific function. Parties discretely share their input with MPC, then the output follows from the function without revealing the input.

Example:

Four colleagues go out to dinner, and they have agreed that the person with the highest salary should pay the bill. However, for privacy reasons, they do not want to disclose their salary. That's why they decide to use MPC. They individually enter their salary into the MPC application. This application indicates which colleague has the highest salary, without the colleagues knowing what their salary is from each other. The application itself does not see this data either because the salaries are first encrypted before the analysis is performed.

In short: MPC is a protocol that creates knowledge for all parties via a function without releasing the underlying data.

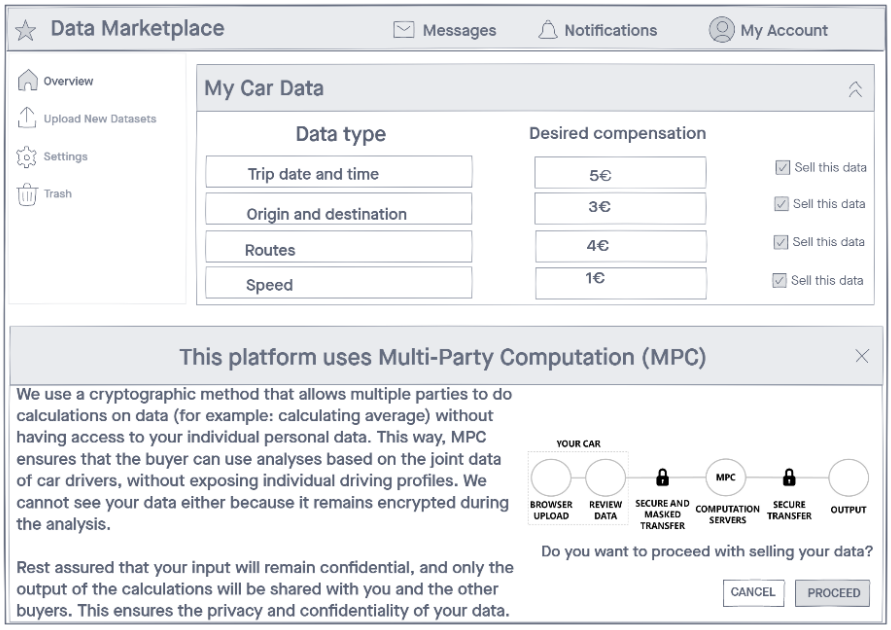
In this design, you can share your car data via an **MPC-based data marketplace**. The data marketplace is managed by a platform operator, but **they do not have access to your data.** They only connect buyers (in this case, mobility service providers) and sellers (in this case, you).

After you agree to sell your driving data, it will be encrypted and stored in your car. Then, when mobility service providers buy your driving data, they have to tell the platform operator which analyses they want to perform. **You can choose to exclude**

your driving data if you wish. Then, using MPC, the platform operator performs the analysis and sends the results to those companies (as data buyers). This way, those companies only see the analysis results based on your driving data and not driving car data itself. **The platform operator does not have access to your driving data because it is encrypted and stored in your car.**

A screenshot preview of the MPC -based data marketplace is presented below (you can zoom the image if you are using a smartphone/tablet).

To convince you that the analysis of your data is confidential, a disclaimer has been added that explains how MPC works.



- ✓ Please tick here if you have **read the description** and **understand the scenario** of sharing driving data through MPC-based data marketplaces

You have read the scenario of sharing driving data through **MPC-based data marketplaces**. The code for this scenario is **B2**.

Please enter the code here:

Treatment 3: Data-Computation-Protection (DCP/fictious technology)

A technology to share your confidential information in a data marketplace is **Data-Computation-Protection (DCP)**. With DCP, your data is encrypted, meaning that **your data is being changed so that it cannot be read without knowledge of the secret key that has been used to change your data**. Then, your encrypted data is shared and can then be used to perform meaningful calculations. Think of DCP as a black box that calculates a specific function. Parties discretely share their input with DCP, then the output follows from the function.

Example:

Four colleagues go out to dinner, and they have agreed that the person with the highest salary should pay the bill. However, for privacy reasons, they do not want to disclose their salary. That's why they decide to use DCP. They individually enter their salary into the DCP application. This application indicates which colleague has the highest salary, without the colleagues knowing what their salary is from each other. The application itself does not see this data either because the salaries are first encrypted before the analysis is performed.

In short: DCP is a protocol that creates knowledge for all parties via a function without releasing the underlying data.

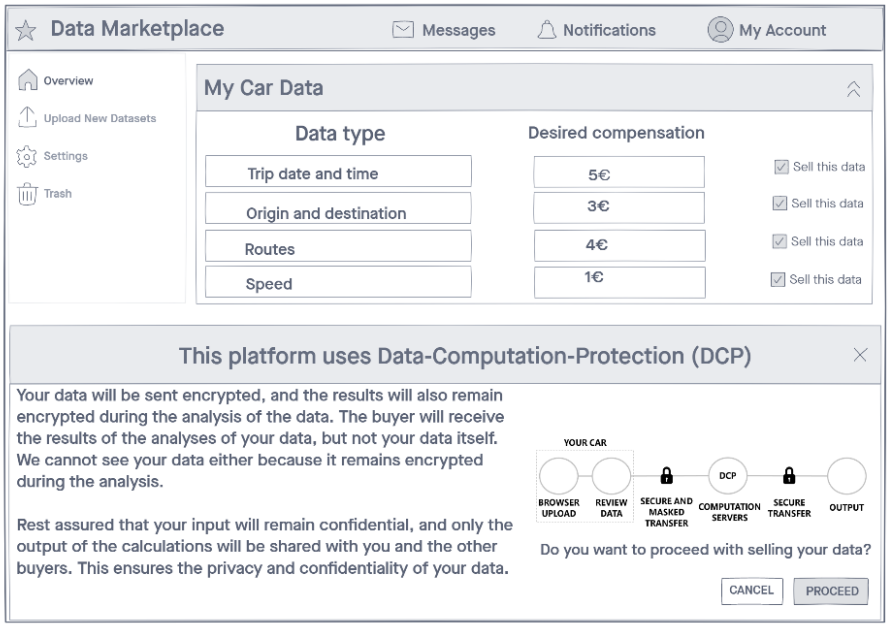
In this design, you can share your driving data via **a DCP-based data marketplace**. The data marketplace is managed by a platform operator, but **they do not have access to your data**. They only connect buyers (in this case, mobility service providers) and sellers (in this case, you).

After you agree to sell your driving data, it will be encrypted and stored in your car. Then, when mobility service providers buy your driving data, they have to tell the platform operator which analyses they want to perform. **You can choose to exclude**

your driving data if you wish. Then, using DCP, the platform operator performs the analysis and sends the results to those companies (as a data buyer). This way, those companies only see the analysis results based on your driving data and not your driving data itself. **The platform operator does not have access to your driving data because it is encrypted and stored in your car.**

A screenshot preview of the DCP-based data marketplace is presented below (You can zoom the image if you are using a smartphone/tablet).

To convince you that the analysis of your data is confidential, a disclaimer has been added that explains how DCP works.



- ✓ Please tick here if you have **read the description** and **understand the scenario** of sharing driving data through DCP-based data marketplaces

You have read the scenario of sharing driving data through **DCP-based data marketplaces**. The code for this scenario is **C3**.

Please enter the code here:

D. Datasets

The datasets and supplementary materials for each study conducted as part of this dissertation are published in the 4TU Research Data Repository. To access these datasets, please refer to the DOI below:

1. Datasets (grounded table) and a short presentation used in Chapter 3: https://doi.org/10.4121/19995101
2. Datasets (grounded table) and a short presentation used in Chapter 4: https://doi.org/10.4121/20406330
3. Datasets (anonymized survey data) resulted from the pre-study in Chapter 5: https://doi.org/10.4121/19403534
4. Datasets (anonymized survey data) resulted from the main study in Chapter 5: https://doi.org/10.4121/20939428

Summary

Digitalization enables massive data generation that can benefit businesses and society, especially when shared with other parties. The emergence of data marketplaces, which rely on a Trusted Third Party (TTP), could play a key role in enabling data sharing to generate meaningful insights and enable new opportunities for businesses and individual consumers. However, many companies refrain from data sharing due to, among others, fear of losing control over data, while consumers are also concerned that their privacy might be compromised. Thus, new approaches are needed to balance data utilization, privacy, and control while minimizing dependency on a TTP.

Advances in privacy-enhancing technologies (PETs)—technical means to protect sensitive data through, for instance, anonymization, encryption, and secure computation, without losing its functionality—could address data sharing barriers. One example that is the focus of this study is Multi-Party Computation (MPC), a cryptographic technique that enables joint computation to generate insights while keeping the input data private. Compared to a TTP, MPC is a different data sharing approach in which (1) the data stays with the owner, (2) the computation process is distributed without an intermediary, and (3) only the results are shared. These differences challenge the current understanding of why and how businesses and consumers share data. Nevertheless, whether businesses and individuals would be more willing to share data with MPC in place is unclear, as less attention is given to the socio-technical implications of MPC on data sharing decisions in data marketplaces and its antecedents. As such, it is necessary to understand the potential impact of MPC use in data marketplaces on antecedents of data sharing decisions by businesses and individual consumers.

This study aims to theorize the socio-technical implications of MPC on sharing through data marketplaces, by investigating how MPC potentially impacts data sharing antecedents by businesses and individuals. To fulfill the research objective, we developed three research questions:

1. What types of data sharing antecedents could be impacted by MPC use in data marketplaces?
2. What could be the impact of MPC use in data marketplaces on antecedents of data sharing by businesses?
3. What could be the impact of MPC use in data marketplaces on antecedents of data sharing by consumers?

We used a mixed-method research design to answer research questions and fulfill the research objective. We first reviewed the current landscape of MPC and data marketplaces to understand the capabilities and the limitations of MPC that might be beneficial in addressing challenges and barriers in data marketplaces adoption. Prior studies suggest that MPC could be implemented in different architectural paradigms ranging from private servers to cloud-assisted models, which come with their trade-offs regarding security requirements and robustness toward adversaries. From the literature, we also identified why data marketplaces are struggling to maintain a strong position in the market, such as fear of losing control over data, privacy concerns, data valuation difficulties, and legal uncertainty. Combining insights from MPC and data marketplaces literature, we found that MPC could enable new architectural approaches for data marketplaces and change the value proposition for actors in the ecosystem. Further, MPC use in data marketplaces could address data sharing barriers, allowing further utilization and value creation from data to generate meaningful insights and stimulate innovation.

Next, we conducted a business model analysis to identify data sharing antecedents that could be impacted by MPC use in data marketplaces. We collected data from 15

MPC experts and practitioners through semi-structured interviews, which were then structured using the Unified Business Model framework. Findings show that MPC could change the architecture of data marketplaces into peer-to-peer or cloud-based models, which ultimately could enable new value propositions by facilitating distributed, trustless, and privacy-enhancing data sharing that maintains control and reduces risks. These value propositions allow MPC to potentially impact control, privacy, trust, and risks as antecedents of data sharing decisions in data marketplaces. We used these findings to narrow down our focus in the next research phase.

Based on the findings of the business model analysis, we investigated the potential impact of MPC use in data marketplaces on control, trust, and risks as antecedents of business data sharing. We conducted semi-structured interviews with 23 experts and practitioners in the automotive industry. We found that MPC use in data marketplaces could enhance technology-based control by data providers, shift the importance of trust into trust in the technology, and shift the relevance of competitiveness and data misuse risks from data providers to data buyers. These impacts are expected to materialize under three conditions: the presence of clear benefits, strong organizational readiness, and low data sensitivity as a starting point.

After that, we investigated the potential impact of MPC use in data marketplaces on control, privacy, trust, and risks as antecedents of consumer data sharing. We conducted a between-subject experiment via an online crowdsourcing platform with 1457 participants by comparing consumers' willingness to share data through data marketplaces that use a TTP, MPC, and fictitious technology DCP. Our findings show that consumers who are asked to share through data marketplaces that use MPC perceive higher control over data, lower privacy concerns, and lower risks than those who are asked to share through TTP. We also found that consumers perceive higher trust in data buyers while sharing through data marketplaces that use MPC than TTP. Moreover, those who are asked to share through data marketplaces that use MPC are

also more willing to share data than those who are asked to share through TTP. However, there is no significant difference between MPC and TTP regarding trust in data marketplace operators. Furthermore, comparing MPC with fictitious technology DCP, our findings show no differences between the two conditions for all antecedents that we investigate.

This research is one of the first to theorize the potential impact of MPC on data sharing decisions. In this way, we contribute to the socio-technical understanding of MPC, which is currently dominated by technical perspectives to improve efficiency and scalability. We also contribute to the data marketplaces literature by adding insights on how MPC could address data marketplaces challenges to boost its adoption and enable the data economy. This research is also relevant to practitioners. We show that MPC should be positioned as a data collaboration tool besides being an instance of PETs. We also show that MPC could enable companies to adopt privacy-friendly business models. At the same time, MPC challenges the relevance of trusted intermediaries by allowing distributed computation between interested parties, thus leveling the playing field in the data economy. Nevertheless, our research raises the importance of effectively communicating MPC to improve transparency and knowledgeability. Furthermore, practitioners should be aware that MPC should be complemented with other governance mechanisms to ensure a trustworthy data sharing environment.

Finally, this research provides a foundation for future studies to understand the socio-technical implications of MPC on data sharing decisions. Future research could expand our findings from the qualitative study by testing the resulting propositions and the nomological net between data sharing antecedents using Structural Equation Modelling (SEM). Future research could also incorporate three conditions of benefits, readiness, and data sensitivity as either moderating effects or boundary conditions that should be kept constant. From the design perspective, future research could employ Design Science Research Methodology (DSRM) to develop an artifact or

working prototype of MPC-based data marketplaces, which can be evaluated to examine the impact of MPC on decisions to share data through data marketplaces by businesses and consumers. Moreover, future research could go beyond data marketplaces domain and evaluate the impact of MPC on data sharing decisions in different domains, such as digital health, finance, and energy. Furthermore, future research might benefit from linking MPC with other theories like data platform openness, dynamic capabilities, data-driven business models, collective action, and the unifying theory of polycentric information commons.

Samenvatting

Digitalisering maakt het massale genereren van gegevens mogelijk die ten goede kan komen aan bedrijven en de samenleving, vooral wanneer ze worden gedeeld met andere partijen. De opkomst van “datamarktplaatsen”, die afhankelijk zijn van een Trusted Third Party (TTP), zouden een sleutelrol kunnen spelen bij het mogelijk maken van het delen van data om zinvolle inzichten te genereren en nieuwe kansen voor bedrijven en individuele consumenten mogelijk te maken. Veel bedrijven zien echter af van het delen van data, onder meer uit angst de controle over data te verliezen, terwijl ook consumenten bang zijn dat hun privacy in het gedrang komt. Er zijn dus nieuwe benaderingen nodig om gegevensgebruik, privacy en controle in evenwicht te brengen en tegelijkertijd de afhankelijkheid van een TTP te minimaliseren.

Voordelen in Privacy-Enhancing Technologies (PETs)—technische middelen om gevoelige gegevens te beschermen door bijvoorbeeld anonimisering, codering en veilige berekeningen, zonder de functionaliteit te verliezen—zou belemmeringen voor het delen van gegevens kunnen wegnemen. Een voorbeeld dat centraal staat in dit onderzoek is Multi-Party Computation (MPC), een cryptografische techniek die gezamenlijke berekeningen mogelijk maakt om inzichten te genereren terwijl de invoergegevens privé blijven. Vergeleken met een TTP is MPC een andere aanpak voor het delen van gegevens waarbij (1) de gegevens bij de eigenaar blijven, (2) het rekenproces wordt verspreid zonder tussenpersoon en (3) alleen de resultaten worden gedeeld. Deze verschillen vormen een uitdaging voor het huidige begrip van waarom en hoe bedrijven en consumenten gegevens delen. Desalniettemin is het onduidelijk of bedrijven en individuen meer bereid zouden zijn om gegevens te delen met bestaande MPC's, aangezien er minder aandacht wordt besteed aan de sociaal-technische implicaties van MPC op beslissingen over het delen van gegevens op datamarktplaatsen en de antecedenten daarvan. Als zodanig is het noodzakelijk om de potentiële impact van MPC-gebruik op datamarktplaatsen te begrijpen op

antecedenten van beslissingen over het delen van data door bedrijven en individuele consumenten.

Deze studie heeft tot doel de sociaal-technische implicaties van MPC op het delen via datamarktplaatsen te theoretiseren, door te onderzoeken hoe antecedenten van het delen van gegevens door bedrijven en individuen mogelijk worden beïnvloed door MPC. Om aan de onderzoeksdoelstelling te voldoen, hebben we drie onderzoeksvragen ontwikkeld:

1. Welke soorten antecedenten voor het delen van gegevens kunnen worden beïnvloed door MPC-gebruik op datamarktplaatsen?
2. Wat zou de impact kunnen zijn van MPC-gebruik op datamarktplaatsen op antecedenten van het delen van data door bedrijven?
3. Wat zou de impact kunnen zijn van MPC-gebruik op datamarktplaatsen op antecedenten van het delen van data door consumenten?

We gebruikten een mixed-method onderzoeksontwerp om onderzoeksvragen te beantwoorden en uiteindelijk de onderzoeksdoelstelling te bereiken. We hebben eerst het huidige landschap van MPC- en datamarktplaatsen beoordeeld om inzicht te krijgen in de mogelijkheden en beperkingen van MPC die nuttig kunnen zijn bij het aanpakken van uitdagingen en belemmeringen bij de acceptatie van datamarktplaatsen. Eerdere studies suggereren dat MPC kan worden geïmplementeerd in verschillende architecturele paradigma's, variërend van privéservers tot cloudondersteunde modellen, die gepaard gaan met hun afwegingen met betrekking tot beveiligingsvereisten en robuustheid ten opzichte van tegenstanders. Uit de literatuur hebben we ook vastgesteld waarom datamarktplaatsen moeite hebben om een sterke positie in de markt te behouden, zoals angst om de controle over data te verliezen, zorgen over privacy, problemen met de waardering van data en rechtsonzekerheid. Door inzichten uit de literatuur over MPC en datamarktplaatsen te combineren, ontdekten we dat MPC nieuwe

architecturele benaderingen voor datamarktplaatsen mogelijk zou maken en de waardepropositie voor actoren in het ecosysteem zou veranderen. Verder zou het gebruik van MPC's op datamarkten belemmeringen voor het delen van data kunnen wegnemen, waardoor verder gebruik en waardecreatie van data mogelijk wordt om zinvolle inzichten te genereren en innovatie te stimuleren.

Vervolgens hebben we een bedrijfsmodelanalyse uitgevoerd om antecedenten voor het delen van gegevens te identificeren die kunnen worden beïnvloed door MPC-gebruik op gegevensmarkten. We verzamelden gegevens van 15 MPC-experts door middel van semi-gestructureerde interviews, die vervolgens werden gestructureerd met behulp van het Unified Business Model-raamwerk. Bevindingen tonen aan dat MPC de architectuur van datamarktplaatsen zou kunnen veranderen in peer-to-peer- of cloudgebaseerde modellen, die uiteindelijk nieuwe waardeproposities mogelijk zouden kunnen maken door gedistribueerde, betrouwbare en privacybevorderende gegevensdeling mogelijk te maken die de controle behoudt en risico's vermindert. Deze waardeproposities stellen MPC in staat om controle, privacy, vertrouwen en risico's mogelijk te beïnvloeden als antecedenten van beslissingen over het delen van gegevens op datamarkten. We gebruikten deze bevindingen om onze focus in de volgende onderzoeksfase te verfijnen.

Op basis van de bevindingen van de analyse van het bedrijfsmodel onderzochten we de mogelijke impact van MPC-gebruik op datamarkten op controle, vertrouwen en risico's als antecedenten van het delen van bedrijfsgegevens. We hebben semigestructureerde interviews gehouden met 23 experts in de auto-industrie. We ontdekten dat MPC-gebruik op datamarkten de op technologie gebaseerde controle door dataproviders zou kunnen verbeteren, het belang van vertrouwen zou kunnen verschuiven naar vertrouwen in de technologie en de relevantie van concurrentievermogen en risico's op datamisbruik zou kunnen verschuiven van dataproviders naar datakopers. Deze effecten zullen zich naar verwachting manifesteren onder drie voorwaarden: de aanwezigheid van duidelijke voordelen, een

sterke paraatheid van de organisatie en een lage gegevensgevoeligheid als uitgangspunt.

Daarna onderzochten we de mogelijke impact van MPC-gebruik op datamarkten op controle, privacy, vertrouwen en risico's als antecedenten van het delen van consumentendata. We hebben een online experiment uitgevoerd via een online crowdsourcingplatform met 1457 deelnemers door de bereidheid van consumenten om gegevens te delen te vergelijken via datamarktplaatsen die een TTP, MPC en fictieve technologie genaamd DCP gebruiken. Onze bevindingen laten zien dat consumenten die worden gevraagd om gegevens te delen via datamarktplaatsen die MPC gebruiken, meer controle over gegevens, minder zorgen over privacy en lagere risico's ervaren dan degenen die worden gevraagd om te delen via TTP. We ontdekten ook dat consumenten meer vertrouwen in datakopers ervaren wanneer ze gegevens delen via datamarktplaatsen die MPC gebruiken dan TTP. Bovendien zijn degenen die worden gevraagd om te delen via datamarktplaatsen die MPC gebruiken, ook meer bereid om gegevens te delen dan degenen die worden gevraagd om te delen via TTP. Er is echter geen significant verschil tussen MPC en TTP wat betreft het vertrouwen in exploitanten van datamarktplaatsen. Bovendien laten onze bevindingen, wanneer we MPC vergelijken met fictieve technologie DCP, geen verschillen zien tussen de twee condities voor alle antecedenten die we onderzoeken.

Dit onderzoek is een van de eerste waarin de potentiële impact van MPC op beslissingen over het delen van gegevens wordt getheoretiseerd. Op deze manier dragen we bij aan het socio-technisch begrip van MPC, dat momenteel wordt gedomineerd door technische perspectieven om de efficiëntie en schaalbaarheid te verbeteren. We dragen ook bij aan de literatuur over datamarktplaatsen door inzichten toe te voegen over hoe MPC de uitdagingen van datamarktplaatsen zou kunnen aanpakken om de acceptatie ervan te stimuleren en de data-economie mogelijk te maken. Dit onderzoek is ook relevant voor beoefenaars. We laten zien dat MPC moet worden gepositioneerd als een hulpmiddel voor gegevenssamenwerking naast een

instantie van PET's. We laten ook zien dat MPC bedrijven in staat kan stellen om privacyvriendelijke bedrijfsmodellen te adopteren. Tegelijkertijd daagt MPC de relevantie van vertrouwde tussenpersonen uit door gedistribueerde berekeningen tussen geïnteresseerde partijen mogelijk te maken, waardoor het speelveld in de data-economie gelijk wordt. Desalniettemin wijst ons onderzoek op het belang van effectieve communicatie over MPC om de transparantie en kennis te verbeteren. Bovendien moeten beoefenaars zich ervan bewust zijn dat MPC moet worden aangevuld met andere bestuursmechanismen om een betrouwbare omgeving voor het delen van gegevens te waarborgen.

Ten slotte biedt dit onderzoek een basis voor toekomstige studies om de sociaal-technische implicaties van MPC op beslissingen over het delen van gegevens te begrijpen. Toekomstig onderzoek zou onze bevindingen uit de kwalitatieve studie kunnen uitbreiden door de resulterende proposities en het nomologische net tussen antecedenten voor het delen van gegevens te testen met behulp van Structural Equation Modeling (SEM). Toekomstig onderzoek zou ook drie voorwaarden van voordelen, gereedheid en gegevensgevoeligheid kunnen omvatten als modererende effecten of randvoorwaarden die constant moeten worden gehouden. Vanuit het ontwerpperspectief zou in toekomstig onderzoek Design Science Research Methodology (DSRM) kunnen worden gebruikt om een artefact of werkend prototype van MPC-gebaseerde datamarktplaatsen te ontwikkelen, die kunnen worden geëvalueerd om de impact van MPC op beslissingen om gegevens te delen via datamarktplaatsen door bedrijven te onderzoeken en consumenten. Bovendien zou toekomstig onderzoek verder kunnen gaan dan het domein van datamarktplaatsen en de impact van MPC op beslissingen over het delen van gegevens in verschillende domeinen evalueren, zoals digitale gezondheid, financiën en energie. Verder zou toekomstig onderzoek baat kunnen hebben bij het koppelen van MPC aan andere theorieën, zoals openheid van dataplatforms, dynamische mogelijkheden,

datagestuurde bedrijfsmodellen, collectieve actie en de verenigende theorie van polycentrische informatie gemeenschappen.

About the author

Wirawan Agahari (Aga) was born in Jakarta, Indonesia, on 21 August 1990. In 2012, Aga obtained his BEng in Telecommunication Engineering from Bandung Institute of Technology (ITB) in Bandung, Indonesia. He then worked for 1.5 years in Telkomsel, one of the largest telecom operators in Indonesia, as a project management staff in Jakarta, Indonesia. He was awarded a scholarship from the Indonesia Endowment Fund for Education (Lembaga Dana Pengelola Pendidikan/LPDP) in 2014 to continue his studies at Delft University of Technology, where he received his MSc degree in Engineering and Policy Analysis in 2016. His master's thesis investigated the impact of digital healthcare platforms on the capabilities of the Dutch elderly to achieve independent living. He subsequently worked as a Senior Research Associate at the Centre for Innovation Policy and Governance (CIPG) in Jakarta, Indonesia, from 2017 to 2019, where he conducted various research and advisory projects in the field of information and social change through collaboration with, among others, government institutions, non-profit technology companies, and civil society organizations.

In March 2019, Aga joined the ICT section at the Faculty of Technology, Policy and Management at Delft University of Technology. His PhD was part of the Safe-DEED project (Safe Data-Enabled Economic Development) funded by EU Horizon 2020. His research has been published in conferences (ECIS, Bled eConference, ITS Biennial Conference) and journals (Telematics and Informatics, Journal of Theoretical and Applied Electronic Research, Electronic Markets), and he was awarded best Research-in-Progress paper at the European Conference on Information Systems (ECIS) 2022. Aga also taught courses at the BSc (e.g., Governance Specialization I&C) and MSc levels (e.g., Digital Platform Design, Design Innovation 4.0 in Supply Networks) and supervised five master theses. Further, he held management responsibility as a colloquium coordinator of the ICT section from 2019-2020.

List of publications

Agahari, W., Fiebig, T., & de Reuver, M. *Does multi-party computation impact consumers' willingness to share data? An experimental study*. [Manuscript in preparation].

Abbas, A. E., **Agahari, W.**, Ofe, H., Zuiderwijk, A., & de Reuver, M. (2023). Toward Sovereign Data Exchange Through a Meta-Platform for Data Marketplaces: A Preliminary Evaluation of the Perceived Efficacy of Control Mechanisms. In *Proceedings of the 36th Bled eConference, Slovenia*.

de Reuver, M., Ofe, H., **Agahari, W.**, Abbas, A. E., & Zuiderwijk, A. (2022). The openness of data platforms: a research agenda. In *Proceedings of the 1st International Workshop on Data Economy* (pp. 34-41).

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*, 1-26.

Agahari, W., & de Reuver, M. (2022). Rethinking consumers' data sharing decisions with the emergence of multi-party computation: An experimental design for evaluation. *ECIS 2022 Research-in-Progress Papers*. 25.

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.

Agahari, W., de Reuver, M., van der Wel, I., & van Aalst, C. (2021). *D2.7 User experiment report v3*. Safe Data-Enabled Economic Development (Safe-DEED).

Agahari, W., Petronia, M., & de Reuver, M. (2021). Cutting out the trusted third party in business-to-business data exchange: A quantitative study on the impact of multi-party computation on firms' willingness to share sensitive data in supply chains. In *Proceedings of the ITS Biennial Conference 2021*.

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. In *Proceedings of the 34th Bled eConference, Slovenia*.

Agahari, W., Dolci, R., & de Reuver, M. (2021). Business model implications of privacy-preserving technologies in data marketplaces: The case of multi-party computation. *ECIS 2021 Research Papers*. 59.

Agahari, W. (2021). Platformization of data sharing: Multi-party computation (MPC) as control mechanism and its effect on firms' participation in data sharing via data marketplaces. In *ECIS 2021 Doctoral Consortium*.

Petronia, M., **Agahari, W.**, & De Reuver, M. (2020). *D2.6 User experiment report v2*. Safe Data-Enabled Economic Development (Safe-DEED).

Breitfuss, G., & Fruhwirth, M., Disch, L., de Reuver, M., & **Agahari, W.** (2020). *D2.3 Business model decision support tool*. Safe Data-Enabled Economic Development (Safe-DEED).

De Reuver, M., **Agahari, W.**, Dolci, R., Breitfuss, G., & Fruhwirth, M. (2020). *D2.2 Business models for use cases and generic business models*. Safe Data-Enabled Economic Development (Safe-DEED).

Susha, I., Flipsen, M., **Agahari, W.**, & de Reuver, M. (2020). Towards generic business models of intermediaries in data collaboratives: From gatekeeping to data control. In *the International Conference on Electronic Government* (pp. 304-315).

Nikou, S., **Agahari, W.**, Keijzer-Broers, W., & de Reuver, M. (2020). Digital healthcare technology adoption by elderly people: A capability approach model. *Telematics and Informatics*, 101315.

Agahari, W. (2020). Platformization of data sharing: Multi-party computation (MPC) as control mechanism and its effect on firms' participation in data sharing via data marketplaces. In *Proceedings of the 33rd Bled eConference, Slovenia (Doctoral Consortium)*.

De Reuver, M., Fiebig, T., **Agahari, W.**, & Faujdar, V. (2019). *D2.4 User experiment report*. Safe Data-Enabled Economic Development (Safe-DEED).

Swamy, J.K.N., de Reuver, M., **Agahari, W.**, & Fiebig, T. (2019). *D2.1 Threat and incentive model.* Safe Data-Enabled Economic Development (Safe-DEED).

Agahari, W., de Reuver, M., & Fiebig, T. (2019). Understanding how privacy-preserving technologies transform data marketplace platforms and ecosystems: The case of multi-party computation. In *6th Innovation in Information Infrastructure (III) Workshop*.