Data-Driven Value Creation from Businesses to Consumers: Aligning Strategies to Inform Development and Design of Data-Driven Services

DISSERTATION

zur Erlangung des Grades eines Doktors rer. pol.
der Fakultät III – Wirtschaftswissenschaften,
Wirtschaftsinformatik und Wirtschaftsrecht der
Universität Siegen

vorgelegt von

Thomas Neifer

Datum der Disputation: 17.07.2025

Erstkorrektor: Prof. Dr. Gunnar Stevens
Zweitkorrektor: Prof. Dr. Alexander Boden
Dekan der Fakultät III: Prof. Dr. Marc Hassenzahl

Dedication

For my beloved mother † and father

who never stopped believing in me.

Abstract

Data-driven business models are increasingly changing how businesses generate value, shape customer relationships, and secure competitive advantages. In doing so, businesses face the challenge of integrating data operationally and using it strategically. At the same time, data-driven services increasingly penetrate consumers' lives, influencing and collecting data from their consumption practices. This raises the question of how data-driven services can be designed to create added value for businesses and consumers. Based on the three derived concepts of Business-Data, Consumer-Data, and Business-Consumer Alignment, this thesis addresses this interface by examining the perspectives of businesses and consumers on data-driven services.

As a coordination process between business and data strategy, two research papers on Business-Data Alignment examine the challenges businesses face in developing data-driven services. Three papers focus on Consumer-Data Alignment by analyzing the way consumers interact with data-driven services and how data influences their everyday practices. In addition, two research papers develop methodological tools to support Business-Consumer Alignment as a cooperative coordination process between businesses and consumers to achieve common goals and values in designing data-driven services.

This thesis bridges theory and practice by deriving action-oriented principles and tools to address strategic and operational challenges in designing data-driven services, contributing to the research domains of Information Systems, Consumer Informatics, Recommender Systems, and Reputation Systems.

- **Information Systems:** This work expands the understanding of data-driven value creation by specifying the strategic link between business and data strategies through Business-Data Alignment. From the perspective of businesses, technical challenges, such as the integration of heterogeneous data sources, and organizational hurdles, such as a lack of data literacy, are addressed. The developed Data Service Canvas and Data Storytelling Process represent practical frameworks that support companies in strategically planning and communicating data-driven business models and services.

- **Consumer Informatics:** This thesis examines the consumer perspective to show how data-driven services shape consumers' everyday consumption practices and how consumers can be involved as active designers of user-centered data-driven services. Concrete design principles based on consumers' needs and practices are also derived for the domains of food and mobility.

- **Recommender Systems:** This work conducts a case study in the food domain in the Recommender Systems research area to show how personalized and value-oriented recommendations can be designed to better support critical consumption practices, such as sustainable or ethical consumption.

- **Reputation Systems:** In the research domain of Reputation Systems, this thesis explores trust-building in the sharing economy, specifically in peer-to-peer carsharing, by examining an algorithm-based trust mechanism to complement user ratings. Evaluating a prototypical algorithm-based reputation system shows how algorithmic ratings can be made transparent and understandable and how such systems can contribute to fair and comprehensible assessments.

Acknowledgements

> *"It's a dangerous business, Frodo, going out your door. You step onto the road, and if you don't keep your feet, there's no telling where you might be swept off to."*
> — **Bilbo Baggins** [640]

The journey of a PhD student often feels like stepping onto an uncertain path filled with unexpected challenges and discoveries. During this academic adventure, I was fortunate to have outstanding mentors: Gunnar Stevens, Alexander Boden, and Andreas Gadatsch. Your support, insightful feedback and openness to discussing big ideas as well as big concerns have been invaluable. You have given me the opportunity to develop academically through conferences and research projects and shape my professional path. I am deeply grateful for this guidance and trust.

> *"Nine companions... So be it! You shall be the Fellowship of the Ring!"*
> — **Elrond** [286]

Research is rarely a solitary endeavor, and this work would not have been possible without the incredible team of co-authors, colleagues, and friends who accompanied me on this journey: Paul, Dennis, Margarita, Veronika, Christina, Lukas, Felix, Jenny, and Argang. Like the Fellowship of the Ring, each of you has brought unique strength, expertise, and unwavering commitment to our shared goals. Whether through critical discussions, shared enthusiasm for deadlines the next morning, or simply being a reliable presence in challenging times, you made this journey unforgettable. I am proud to have you as my friends, co-authors, and fellow sufferers.

> *"Toss me. [...] I cannot jump the distance. You'll have to toss me. Don't tell the elf!"*
> — **Gimli** [287]

In academia, there are moments when the gap to the next bridge feels impossibly wide – the distance too far, the jump too uncertain. In such moments, it takes someone willing to give you that metaphorical "toss" over the abyss,

placing their trust in your ability to land safely on the other side. I am deeply grateful to the mentors who gave me that decisive push: Prof. Dr. Johannes Natrop, whose clarity and ability to explain complex topics enabled me to understand challenging concepts; Prof. Dr. Dirk Schreiber, whose tireless support guided me through key challenges; Prof. Dr. Dr. Franz W. Peren, with whom I had the privilege of co-authoring my first scientific book; and Prof. Dr. Reiner Clement †, who inspired me to pursue an academic career. Without your encouragement, I would not have dared to make that jump.

> *"A hunted man sometimes wearies of distrust and longs for friendship."*
> — **Aragorn** [640]

The road to completing a dissertation is often long and demanding, and moments of doubt are inevitable. In those moments, the steadfast friendship and encouragement of Martin, Arif, Vincent, and Christoph were a source of strength. Your support went beyond words; it was a constant reminder that I was not walking this path alone, even in the most challenging times.

> *"I can't carry it for you, but I can carry you [...]."*
> — **Samwise Gamgee** [640]

I want to express my deepest gratitude to my girlfriend. In moments of doubt and exhaustion, your unwavering support, love, and belief in me carried me forward when my strength faltered. I am endlessly grateful for that.

> *"Fly, you fools!"*
> — **Gandalf the Grey** [640]

Lastly, to those who have ventured through these pages: academia, much like the Mines of Moria, is filled with unseen depths and overwhelming challenges. Not every obstacle must be faced head-on – sometimes stepping back, reflecting, and choosing a different path is the wisest choice. So if you find yourself cornered by deadlines, buried under pages, or pursued by a metaphorical Balrog (reviewer 2) – remember these words.

Related Publications

Parts of this dissertation have already been published as conference or journal papers.

Section 5

Neifer, T., Lawo, D., Bossauer, P., & Gadatsch, A.
Decoding IPaaS: Investigation of User Requirements for Integration Platforms as a Service
Proceedings of the 18th International Conference on e-Business (ICE-B), 2021, CC-BY-NC-ND 4.0

DOI: 10.5220/0010626000470055

Section 6

Neifer, T., Lawo, D., Stevens, G., Boden, A., & Gadatsch, A.
Recommender Systems in Food Retail: Modeling Repeat Purchase Decisions on Transaction Data of a Stationary Food Retailer
Proceedings of the 18th International Conference on e-Business (ICE-B), 2021, CC-BY-NC-ND 4.0

DOI: 10.5220/0010553600250036

Section 9

Lawo D., Neifer T., Esau M., and Stevens, G.
Buying the 'Right' Thing: Designing Food Recommender Systems with Critical Consumers
CHI Conference on Human Factors in Computing Systems (CHI '21), Yokohama, Japan

DOI: 10.1145/3411764.3445264

Section 10

Bossauer, P., Neifer, T., Stevens, G., & Pakusch, C.
Trust versus Privacy: Using Connected Car Data in Peer-to-Peer Car-sharing
CHI Conference on Human Factors in Computing Systems (CHI '20), Honolulu, HI, USA

DOI: 10.1145/3313831.3376555

Section 11

Neifer, T., Bossauer, P., Pakusch, C., Boehm, L., & Lawo, D.
Trust-Building in Peer-to-Peer Carsharing: Design Case Study for Algorithm-Based Reputation Systems
Computer Supported Cooperative Work (CSCW), 33(2), 137-171, 2024, CC-BY 4.0

DOI: 10.1007/s10606-022-09461-4

Section 14

Neifer, T., Lawo, D., & Esau, M.
Data Science Canvas: Evaluation of a Tool to Manage Data Science Projects
Proceedings of the 54th Hawaii International Conference on System Sciences (HICSS), 2021, CC-BY-NC-ND 4.0

DOI: 10.24251/HICSS.2021.657

Section 15

Neifer, T., Lawo, D., Bossauer, P., Esau, M., & Jerofejev, A. M.
Data Storytelling as a Critical Success Factor for Data Science
HMD Praxis der Wirtschaftsinformatik, 57, 1033-1046, 2020,
Translated from German, CC-BY 4.0

DOI: 10.1365/s40702-020-00662-3

Moreover, these publications contribute to the presented topic. However, they are not included as chapters of this thesis.

Technical Enablers and Challenges of Data-Driven Services

- Neifer, T., Bossauer, P., Lawo, D., Volkening, R., & Gadatsch, A. (2022). The Role of Marketplaces for the Transformation from Robotic Process Automation to Intelligent Process Automation. In Proceedings of the 19th International Conference on Smart Business Technologies, ISBN 978-989-758-587-6, ISSN 2184-772X, pages 15-25. DOI: 10.5220/0011146200003280

- Lawo, D., Esau, M., Neifer, T., & Stevens G. (2021). Probing Integrated Household Information Systems for Integrated Food Practices. In Proceedings of the 5th International Conference on Computer-Human Interaction Research and Applications, ISBN 978-989-758-538-8, pages 15-23. DOI: 10.5220/0010643600003060

- Neifer, T., Schmidt, A., Bossauer, P., & Gadatsch, A. (2019). Data Science Management: Planung, Steuerung und Kontrolle von Data Science. Rethinking Finance, (3).

- Bossauer, P., Schreiber, D., Neifer, T., Pakusch, C., & Stevens, G. (2020). Dezenetralisierung der Sharing Economy - Potentiale Blockchain-basierter Sharing-Plattformen. In Proceedings of the 15th International Conference on Wirtschaftsinformatik (WI).
  DOI: 10.30844/wi_2020_b3-bossauer

- Bossauer, P., Neifer, T., Pakusch, C., & Staskiewicz, P. (2019). Using Blockchain in Peer-to-Peer Carsharing to Build Trust in the Sharing Economy. In Proceedings of the 14th International Conference on Wirtschaftsinformatik (WI). DOI: 10.25819/ubsi/1016

- Stevens, G., Bossauer, P., Neifer, T., & Hanschke, S. (2017). Using Shopping Data to Design Sustainable Consumer Apps. In Proceedings of the 2017 Conference on Sustainable Internet and ICT for Sustainability (SustainIT). DOI: 10.23919/SustainIT.2017.8379810

User Needs and Perceptions of Data-Driven Services

- Lawo, D., Neifer, T., Esau-Held, M., & Stevens, G. (2023). Digital Sovereignty: What it is and why it matters for HCI. In Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems. DOI: 10.1145/3544549.3585834
- Esau, M., Lawo, D., Neifer, T., Stevens, G., & Boden, A. (2023). Trust your guts: Fostering Embodied Knowledge and Sustainable Practices through Voice Interaction. Personal and Ubiquitous Computing, 27(2). DOI: 10.1007/s00779-022-01695-9
- Lawo, D., Neifer, T., Esau, M., Engelbutzeder, P., & Stevens, G. (2021). Scan & Go: Understanding Adoption and Design of Smartphone-based Self-Checkout. In Proceedings of the 18th International Conference on e-Business (ICE-B). DOI: 10.5220/0010625701830194
- Pakusch, C., Neifer, T., Bossauer, P., & Stevens, G. (2018). P2P-Carsharing: Motive, Ängste und Barrieren bei der Teilnahme - Eine Explorative Studie. Internationales Verkehrswesen, 70(4).

List of Figures

1 Structure of this Thesis 6
2 Bridging Function of Principles Between Theory and Practice . 7
3 Research Domains 9
4 From Data to Wisdom 16
5 Taxonomy of Data-Driven Business Models 20
6 Co-Evolution of Aligning Strategic and Operational Level . . . 26
7 Business-Data Alignment on a DDBM Level 28
8 Reputation System Reference Model for P2P Carsharing . . . 38
9 Organization of the Conducted Studies in this Thesis 49
10 Relationship structure of an IPaaS market 70
11 Purchase Events as Independent Stochastic Processes 91
12 Empirical Bayesian Model. 98
13 Distribution of Repeat Customer Probability. 100
14 Distribution of Time Intervals (t) for Repeated Purchases . . . 100
15 MPG-Model 102
16 Workflow of F-RS4CC 128
17 Search, Recommendation & Shopping List 129
18 Conceptualization of Sliders during the Workshop 130
19 Area of Negotiability 164
20 Different Driver Profiles Presented to the Participants 180
21 Clickable Prototype 189
22 Data Science Process. 224
23 Data Science Canvas. 227
24 The Three Key Elements of Data Storytelling 239
25 Genres of Narrative Visualization 241
26 Analysis of App Usage to Derive the Relevant Target Group . 244
27 Data Storytelling Process 249
28 Business-Consumer Alignment Canvas 274
29 Triangulation of BDA, CDA, and BCA 275

List of Tables

1 Approaches to Personalised Recommender Systems 30
2 Contribution and Methodology of the Case Studies in Part II . 51
3 Contribution and Methodology of the Case Studies in Part III . 53
4 Contribution and Methodology of the Case Studies in Part IV . 55
5 Overview of the Interview Participants 72
6 Overview of Results 73
7 Model Evaluation. 102
8 Example Rules of F-RS4CC 130
9 Participants 151
10 Examples of Information Types 153
11 Examples for the Classification of Information 155
12 Examples of Information Purposes 160
13 List of Categories of Measurement Criteria of a Driving Score 181
14 Overview of the First Cycle of the DSR Project 218
15 Concept Matrix. 220
16 Overview of Participants 229
17 Story Scaffolds of Data Stories 240
18 Different Purposes of Data Stories 241
19 Important Patterns of Successful Data Stories 242
20 Four Phases of the Conducted Literature Study 242
21 Concept Matrix 243
22 Data-Enhanced Alignment Strategy 270
23 Data-Dominated Alignment Strategy 271

List of Abbreviations

AaS Analytics-as-a-Service

AI Artificial Intelligence

API Application Programming Interface

ATD Aggregate Time Distribution Model

BCA Business-Consumer Alignment

BDA Business-Data Alignment

BI Business Intelligence

BM Business Model

CDA Consumer-Data Alignment

CDO Chief Data Officer

CEO Chief Executive Officer

CF Collaborative Filtering

CMS Content-Management-System

CPO Chief Process Officer

CRM Customer Relationship Management

CI Consumer Informatics

CO2 Carbon Dioxide

DaaS Data-as-a-Service

DDBM Data-Driven Business Model

DDS Data-Driven Service

DMS Document Management System

DSR Design Science Research

DSS Decision Support Systems

ERP Enterprise Resource Planning

F-RS4CC Food Recommender System for Critical Consumers

HCI Human-Computer Interaction

ICT Information and Communication Technologies

IoT Internet of Things

IPA Intelligent Process Automation

IPaaS Integration Platform as a Service

IS Information Systems

IT Information Technology

loc Number of One-Time Repeat Buyers

LSD Logarithmic Series Distribution

MF Matrix Factorization

ML Machine Learning

moc Number of Repeat Buyers

MPG Modified Poisson-Gamma Model

NaN Not a Number

NBD Negative Binomial Distribution

P2P Peer-to-Peer

PD Participatory Design

PDDBM Pure Data-Driven Business Model

PG Poisson-Gamma

RCP Repeat Customer Probability Model

RecSys Recommender System

RNN Recurrent Neural Network

RQ Research Question

SaaS Software-as-a-Service

SD Service Design

SQ3R Survey, Question, Read, Recite, Review

SVD Singular Value Decomposition

UCD User-Centered Design

UI User Interface

UX User Experience

Contents

Dedication ii
Abstract iii
Acknowledgements iv
Related Publications vi
List of Figures x
List of Tables xi
List of Abbreviations xii

I Introduction & Overview 1
1 Introduction 2
1.1 Motivation . . . 2
1.2 Research Domains . . . 8
2 Related Work 13
2.1 Business Models and Services in a Data-Driven World . . . 14
2.2 Business-Data Alignment . . . 24
2.3 Consumer-Data Alignment . . . 32
2.4 Business-Consumer Alignment . . . 40
3 Study Outline & Methodology 48
3.1 Research Framework . . . 48
3.2 Research Activities . . . 55

II Uncovering Problems in Developing Data-Driven Services to Support Business-Data Alignment 63
4 Introduction 64
5 Decoding IPaaS 66
5.1 Introduction . . . 66
5.2 IPaaS Platform Economy . . . 68
5.3 Methodology . . . 71
5.4 Results . . . 73
5.5 Discussion . . . 79
5.6 Conclusion . . . 81
6 Recommender Systems in Food Retail 83
6.1 Introduction . . . 83
6.2 Recommender Systems . . . 85
6.3 Integration of the Time Factor . . . 89
6.4 Repeat Purchase Recommendations . . . 90
6.5 Methodology . . . 93
6.6 Results . . . 99
6.7 Discussion . . . 101
6.8 Conclusion . . . 103
7 Comparative Summary of Findings 105
7.1 Technical and Organizational Challenges . . . 105
7.2 Personalization of Data-Driven Services . . . 107

III Exploring Consumer Practices and Consumer Data Strategies for Consumer-Data Alignment 109
8 Introduction 110
9 Buying the 'Right' Thing 114
9.1 Introduction . . . 114
9.2 Related Work . . . 116
9.3 Empirical Pre-Study . . . 120
9.4 Development of F-RS4CC . . . 127
9.5 Evaluation . . . 133
9.6 Discussion . . . 139
9.7 Conclusion . . . 143
10 Trust versus Privacy 145
10.1 Introduction . . . 145
10.2 Related Work . . . 147
10.3 Methodology . . . 150
10.4 Findings . . . 152
10.5 Discussion . . . 165
10.6 Conclusion . . . 170
11 Trust-Building in Peer-to-Peer Carsharing 172
11.1 Introduction . . . 172
11.2 Related Work . . . 174
11.3 Empirical Pre-Study . . . 179
11.4 Prototyping . . . 187
11.5 Evaluation . . . 191
11.6 Discussion . . . 197
11.7 Conclusion . . . 204
12 Comparative Summary of Findings 206
12.1 Data Strategies Shaping Consumer Practices . . . 206
12.2 Human-Algorithm Interaction . . . 207
12.3 Co-Creating Value, Acceptance, and Trust . . . 209
12.4 Implications for Support Instruments for Businesses . . . 210

IV Aligning Businesses and Consumers: Instruments to Support Data-Driven Value Creation 212
13 Introduction 213
14 Data Science Canvas 216
14.1 Introduction . . . 216
14.2 Data Science Canvas Design & Development . . . 219
14.3 Demonstration & Evaluation . . . 228
14.4 Discussion & Implications . . . 233
14.5 Conclusion . . . 235
15 Data Storytelling 236
15.1 Introduction . . . 236
15.2 Data Storytelling . . . 238
15.3 Methodology . . . 241
15.4 The Success Factors of Data Storytelling . . . 243
15.5 Conclusion and Outlook . . . 249

16 Comparative Summary of Findings **251**
16.1 Data Science Canvas as a Structured Framework 251
16.2 Data Storytelling as a Bridging Method 253

V Research Outcome **255**
17 Discussion **256**
17.1 Business-Data Alignment through a Practice Lens 259
17.2 Consumer-Data Alignment: Flexibility and Negotiation 262
17.3 Co-Creating Data-Driven Value-Added Services 271
18 Conclusion **277**
18.1 Contributions 281
18.2 Limitations 283
References **285**

Part I

Introduction & Overview

1 Introduction

1.1 Motivation

Data has become a critical success factor for companies to benefit from digitalization. Businesses must face the challenge of generating added value from this data. In this context, advanced statistical and machine learning methods are becoming increasingly important due to the volume, velocity, and variety of data [499]. Due to the dynamics of today's markets, intense competition, and pressure to innovate, decision-makers need to evaluate the opportunities in their data and the existing expertise in their companies and unite them to create value [442]. To drive digital transformation, decision-makers need to understand how to scale the value of their data assets and leverage the potential of analytics within their business strategy [472]. Therefore, business models are instruments for explaining the gap between business strategy and its processual implementation [658, 430, 166].

Data transformation into value-creating business models and services requires expert knowledge [67]. However, a study shows that only 24% of decision-makers consider themselves data literate, although this competence is perceived as increasingly relevant [501]. Furthermore, despite its high relevance, the implementation of a data strategy is not yet mature in many companies [140]. It requires an active data-driven corporate culture to enable data-driven value creation [252].

Since digitalization – as an IT-enabled use of digital technologies to enhance and innovate business models [503] – is advancing, Business-IT Alignment is a topic of Information Systems research [24, 580] that deals with the ensuring of IT initiatives to be both timely and appropriate in a synergistic way with the overall business strategy and objectives [386]. However, because data is crucial to business success, so-called Data-Driven Business Models (DDBMs) emerged. DDBMs describe the utilization of Data-Driven Services (DDSs) – defined as "the use of data and analytics to support the decision-making process of the customer via data and analytics-based features and experiences in the form of a stand-alone offering or bundled with an existing product or service" [560]. These services are used to enhance an existing business model

or to innovate new (pure) DDBMs [270].

Research has shown that when organizations see data strategy as an IT strategy, it leads to failure of strategy implementation, loss of management confidence, and inability to generate valuable results [209]. Therefore, there is a need to rethink Business-IT Alignment and emphasize a new term for the interaction between business and data in the form of Business-Data Alignment (BDA). Business-Data Alignment as a term is rarely used in academic literature, and a definition is pending, although it is seen as "fundamental to ensure further exploitation of big data [...]" [696].

As more and more consumers use digital channels and touchpoints to connect with businesses, such as using social media to rate products or services or utilizing digital services from retailers to purchase products for recipes in cooking applications like Chefkoch [242], they derive increasing value from these services. At the same time, businesses face the challenge of integrating data from various applications, channels, and social media, along with the growing number of cloud applications connected to a DDBM, to enhance DDS [270, 178].

In line with the definition of a DDS, it should not only be considered valuable for businesses. As consumers serve as suppliers of data to generate and drive DDBMs [409], and DDSs are increasingly finding their way into consumers' lives [75], the consumer's perspective – as claimed by User-Centered Design (UCD) [4] – must also be considered when designing a suitable DDBM or DDS because consumers would only tend to share (personal) data if there is also an added value for them [35]. Therefore, this thesis further motivates the term Consumer-Data Alignment (CDA), which can be interpreted as a composition of consumer strategy and data strategy that consumers must harmonize to benefit from their provided data through appropriate and sufficient added-value DDSs. Supporting consumers in aligning their consumer strategy with their data strategy in the context of DDSs is essential, as consumers are usually unclear about the use of their data, do not have a data strategy, or do not implement it consistently [321, 451, 289], let alone make cost-benefit considerations when sharing their data concerning the expected added value of the DDS [344, 318, 8], as they do not understand or even recognize the collection

and processing of their data [289]. Regarding the consumer strategy, however, consumers pursue individual goals and want to realize these with data disclosure [321], whereby achieving goals is strongly influenced by the respective consumption practices [182, 672]. Accordingly, consumers must be empowered to understand and perceive DDS to achieve their goals and best support their consumption practices in developing and designing DDS [374].

The question arises regarding how consumers and businesses can create value with DDS's complex and innovative technology. Therefore, this thesis motivates the term Business-Consumer Alignment (BCA), which methodically describes the coordination between the interests of both stakeholders. While many DDS developments in businesses are still predominantly technology-centered and focus on the company's needs [246, 673], UCD distinguishes itself by actively involving the user in the development and design processes [189]. The consumer strategy is usually regarded in the literature as an "organizational strategy that is developed based on insights about consumers" [245], which focuses on the user as a customer and is business-oriented [246]. Accordingly, UCD must not be limited by businesses to the focus on the user as a customer but must take into account the user's entire consumption practices, as "consumers do not choose brands, they choose lives" [200], since digitalization ensures that consumers' living environments are increasingly permeated with DDS [75]. Due to the gap between innovation pressure, the integration of the consumer in design processes, and missing expert knowledge in developing and managing DDS in businesses [501], there is a practical need for tools that can support companies in their data-driven value-creation processes, DDBM and DDS development [333].

Therefore, this thesis can be categorized into Information Systems and Consumer Informatics research fields. While Information Systems mainly focuses on the corporate perspective in the context of the development and design of DDSs, Consumer Informatics – as a new field closely linked to Information Systems and Human-Computer Interaction (HCI) – deals with the development and design of DDS at the interface between businesses and consumers, taking a holistic view of consumers and thus expanding Information Systems research beyond the corporate perspective. Central consumer fields of Consumer Informatics, which are currently being powerfully penetrated by digi-

talization and in which DDS play a central role, are, in particular, food and mobility [75], examined in this thesis's context. Mobility and food, in particular, are to be united under the banner of sustainable consumption, which is currently to be promoted by digital artifacts such as DDSs [75, 212, 256].

Personalized recommender systems are under discussion to support food consumption practices through DDS, which are traditionally based on food popularity or similarity and, therefore, do not work well for sustainable critical consumption practices, e.g., vegan food recommendation. Hence, the case study in the food context deals with developing a recommender system for critical consumers and deriving design principles for such DDS based on users' practices, perceptions, and needs.

A prominent example in the mobility sector for achieving sustainable mobility practices is shared mobility, which stands out in the context of the sharing economy; for example, by breaking up car-ownership and car-based mobility and sharing the private vehicle in peer-to-peer (P2P) carsharing [256, 136]. However, a central problem in P2P carsharing is the lack of peer trust [602]. Traditional reputation mechanisms for building trust, such as user ratings, are increasingly losing acceptance and effectiveness. At the same time, digitalization enables the provision of algorithm-based reputation systems by integrating diverse sensor technology in cars. In contrast to user ratings, these are free of social bias and enable additional insights and features for building trust among peers [602]. For this reason, the case studies in the mobility sector are dedicated to developing and designing an algorithm-based reputation system that is intended to improve trust building in P2P carsharing and thus support the transition to sustainable mobility practices.

Having identified the need to support businesses regarding the development of value-added DDBMs and DDSs, which qualifies as a "heretofore unsolved and important business problem" [266] where "concepts and methods of how new services can be designed from this data is [sic] still developing" [70], and consumers in understanding and developing for participation in Business-Consumer Alignment [266, 478], my research questions are:

RQ1: What problems do businesses face within their practices in Business-Data Alignment and data-driven value creation, and what are recommenda-

tions for action to overcome these barriers?

RQ2: How do consumers perceive DDSs within Consumer-Data Alignment, and what supportive factors should be considered in a user-centered design of DDSs?

RQ3: How can businesses and consumers be supported in Business-Consumer Alignment by suitable tools and processes to create value for both?

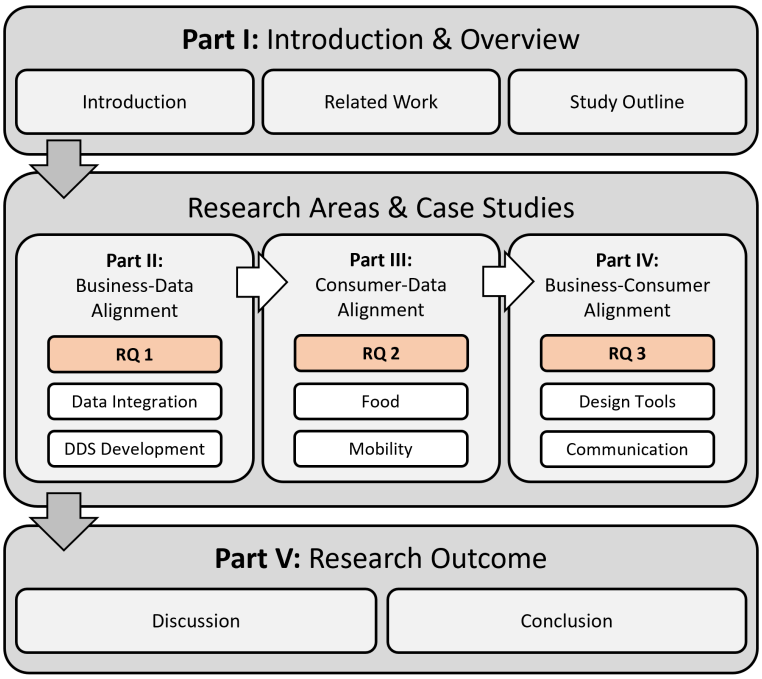


Figure 1: Structure of this Thesis

Acting as a bridge between theory and practice (see Fig. 2), this thesis derives principles based on the theoretical foundations behind data-driven value creation, which induce general, rather abstract and action-oriented guidelines from theory and practice to support practical implementation in flexible contexts [327, 583, 101]. Finally, methods are derived from the principles, which represent specific and context-related tools and serve as practical implementations. This thesis creates a framework that makes it possible to strategically and practically design DDBMs and DDSs in terms of Business-Data, Consumer-Data, and Business-Consumer Alignment [583, 101, 30].

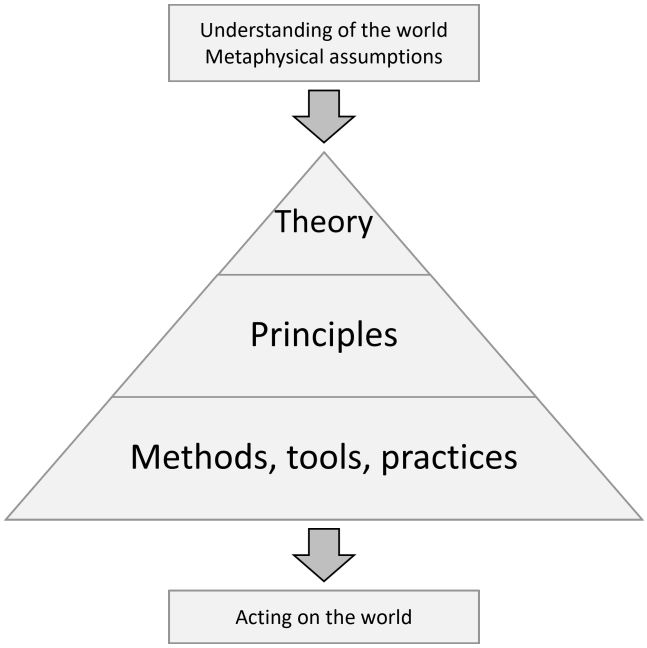


Figure 2: Bridging Function of Principles Between Theory and Practice according to [583, 327]

To answer the research questions and build the bridge between theory and practice on data-driven value creation, this thesis is structured in five parts (see Fig. 1): **Part I** lays the theoretical foundation by setting out the motivation, previous related work, the research methodology and its link to the following parts. To answer the research questions, case studies are then conducted in the three main parts of this thesis: **Part II** is dedicated to RQ 1 and considers case studies on the business perspective on Business-Data Alignment, uncovering the current problems and hurdles in data sourcing and integration for and development of DDS to enable DDBMs. **Part III** deals with answering RQ 2 and sheds light on the consumer perspective. To gain in-depth insights into consumer perceptions, needs, and obstacles in the context of Consumer-Data Alignment, this thesis presents three case studies using UCD methods in the application fields of Consumer Informatics, food, and mobility. **Part IV** synthesizes the findings of the previous parts to answer RQ 3 and develops suitable design and communication tools that support businesses and consumers in creating value-added DDS to achieve a successful Business-Consumer Alignment. Finally, **Part V** discusses the findings compared to the

related work and addresses limitations and future work.

1.2 Research Domains

This thesis contributes to Information Systems and Consumer Informatics research alongside data-driven value creation. In addition, this thesis contributes to the research fields of Recommender and algorithm-based Reputation Systems in the context of DDSs (see Figure 3).

As an interdisciplinary field of research, Information Systems deals with the development, use, and effects of Information and Communication Technologies (ICT) in organizations and society [343]. It examines, mainly from a business perspective, how information is collected, processed, integrated, stored, and communicated to support decision-making processes and enable value creation [506]. Therefore, there is a need to consider Business-Data Alignment to enable data-driven value creation with DDS from a business perspective.

Consumer Informatics is a relatively new field of research that deals with consumers' use of ICT and their impact on consumption and society. It, therefore, focuses in particular on the aspect that information technologies developed by businesses also find their way into private households and have a substantial impact on their consumption and living environments, which means that it is closely linked to Information Systems and HCI [683, 600]. Consumer Informatics is grounded on the interdisciplinary consumer research approach and complements it by analyzing and designing digital systems in private living environments. The Consumer Informatics application fields are housing, mobility, food, and finance. As mentioned, this thesis will focus on mobility and food, as these are currently being heavily penetrated with DDS under the banner of sustainable consumption [75]. Consumer-Data Alignment is, therefore, of central relevance to supporting the user by making informed decisions and strengthening the position in Business-Consumer Alignment when participating in the development and design of user-centered DDSs. The following paragraphs present the specific problems of the addressed Consumer Informatics fields of food and mobility, specifying their respective hurdles

concerning DDSs. Therefore, figure 3 highlights the respective topics of Consumer-Data Alignment and DDSs examined within the case studies in part III.

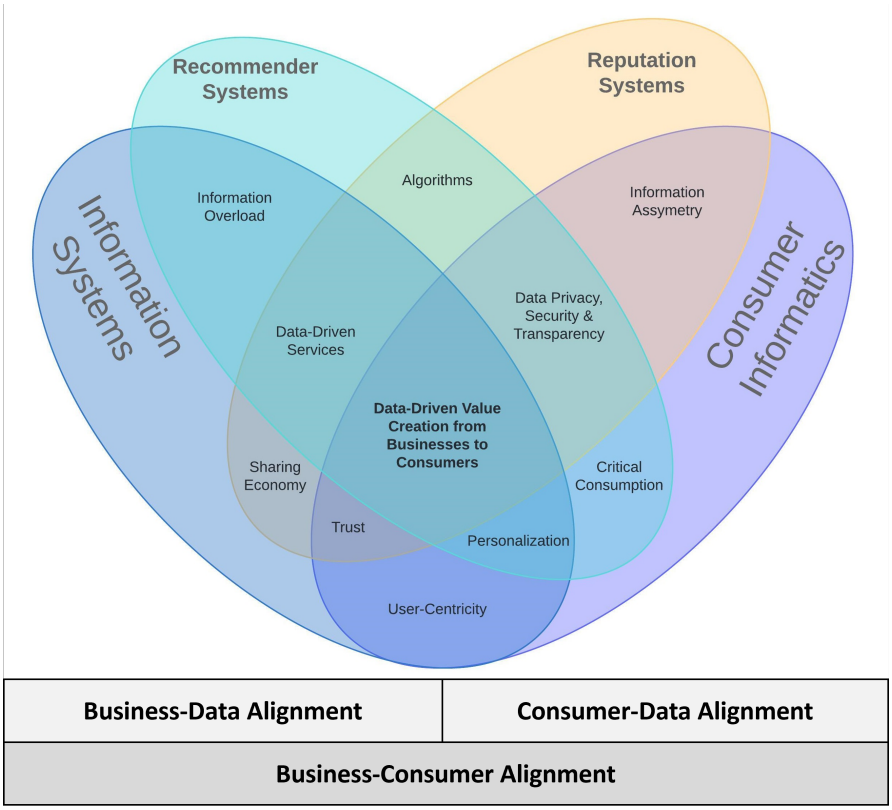


Figure 3: Research Domains

Consumers are confronted with increasing information regarding product availability, dietary choices, and sustainability aspects in the food sector. This abundance of data contributes to ”information overload”, a vital issue in Information Systems research that makes it difficult for individuals to process all available information and make informed decisions [187]. Recommender Systems aim to address this issue by focusing on individual preferences and offering personalized suggestions to reduce information overload [292, 325]. As an essential component of digital platforms and DDSs, these systems leverage machine learning and artificial intelligence to handle large amounts of data and present relevant recommendations to consumers [14, 153, 73]. Food Recommender Systems are designed primarily from a business perspec-

tive to stimulate consumption [586] instead of additionally supporting sustainable niche practices like veganism. The necessary user-centricity can be seen at the intersection of Consumer Informatics and Information Systems. However, the user-centered development of digital services requires understanding such practices [465, 672]. Therefore, this thesis examines critical consumption in the food sector, i.e., conscious and sustainable consumption [224].

The focus on reducing environmental impact has further led to the rise of various innovative business models in the context of sustainable mobility, including peer-to-peer (P2P) carsharing. Driven by the principles of the sharing economy [597, 188], P2P carsharing promotes the efficient use of private vehicles by enabling car rental as a private resource between renters and rentees, thus reducing the need for personal car ownership and supporting more sustainable mobility practices [40]. By leveraging underutilized private resources, P2P carsharing aligns with broader sustainability goals, encouraging reduced carbon emissions and resource consumption. However, despite its environmental benefits, P2P carsharing platforms' success depends heavily on trust between participants, as transactions are carried out between private individuals rather than established companies. The necessity of building trust comes from the problem of information asymmetry about peers' true characteristics, reliability, and intentions [219, 602].

Reputation systems are widely implemented in P2P carsharing platforms to address these issues. These systems aim to reduce information asymmetries between peers by providing them with ratings and reviews that reflect the trustworthiness of renters and rentees [81]. Reputation Systems offer a proxy for trustworthiness, enabling peers to make more informed decisions about whom to engage with on the platform. However, one key issue with Reputation Systems is the presence of social bias, where certain user groups may be unfairly targeted with negative reviews or assessments based on discriminatory attitudes rather than objective experience [250]. To counteract this, recent research examines the role of DDSs like Algorithm-based Reputation Systems that utilize objective data rather than subjective user reviews to assess trustworthiness. In P2P carsharing, car telematics data - such as driving behavior and vehicle usage - can generate algorithm-based ratings, offering a more objective and data-driven measure of trustworthiness, helping to mit-

igate potential bias and discrimination [602]. By integrating such objective data into Reputation Systems, P2P carsharing platforms can foster greater trust and inclusivity by helping consumers make informed decisions while continuing to support sustainable mobility practices.

As recommender systems and algorithm-based reputation systems are algorithms that process personal data, data protection and data security must also be taken into account from the consumer's perspective. It is necessary to examine which data is made available by users under which circumstances to enable DDSs [290, 639, 289]. Transparency and traceability of the algorithms and systems used also play a key role in supporting decisions [366]. Users must understand how recommendations and ratings are generated and the data used. Only then will they be able to make informed decisions and retain control over their data [100]. The issue of fairness and discrimination through algorithmic systems is also increasingly being discussed [42] and makes it necessary to ensure that recommender and reputation systems do not reproduce and reinforce existing inequalities and prejudices [204].

To enable data-driven value creation from businesses to consumers - as a central and overarching topic across all addressed research fields - both the technical and user-centered aspects must be considered [237]. To this end, Information Systems research sheds light on how this value creation can be designed and optimized - primarily from a business perspective [381] and the inclusion of Consumer Informatics ensures that the needs and preferences of consumers are integrated and that a corresponding user-centric development takes place to create added value on both sides [683, 600, 572].

Cooperation and exchange between the various stakeholders play a critical role here. Only through an interdisciplinary and participatory approach can the complex challenges of creating data-driven value-added services succeed [381]. This thesis is intended to play a vital bridging role and bring together the different perspectives.

Data skills and technical capabilities are still required to create user-centric value-added services, which are currently somewhat underdeveloped, especially on the part of businesses [501]. This results in the need to support businesses in developing and designing data-driven value-added services [333,

70]. Accordingly, this thesis focuses on developing tools to support businesses in creating and communicating user-centered DDSs.

2 Related Work

The transition from traditional business models (BM) to DDBMs represents a paradigm shift in the business world. This evolution is reflected in how organisations create and capture value and services' roles in these models. In the context of DDBMs, DDSs are gaining importance as they offer new ways of reaching the customer and gaining deeper insights into customer needs and market dynamics by analysing big data. Therefore, the concepts of business models, DDBMs, DDSs and data-driven value creation are introduced in section 2.1, and the data-driven transition is shown. Due to the complexity of DDSs, the need to rethink the alignment of business and IT strategy in Business-IT Alignment is justified with a new term called Business-Data Alignment. The term is derived and formed from the traditional Business-IT Alignment in section 2.2. Supporting businesses in Business-Data Alignment requires understanding the challenges faced in developing and designing DDSs. One of the issues associated with development is the integration of data from different sources to enable DDSs. To this end, data integration platforms are being introduced in section 2.2.1 that can support the integration challenges of businesses and offer added value. Another problem is implementing suitable methods and models to develop DDSs. Therefore, recommender systems are described on the business side in section 2.2.2, which can offer enhanced added value for customers in food retailing.

In addition to businesses, DDSs should offer added value for consumers, especially as they are increasingly finding their way into their lives and are only made possible by disclosing their data. CI, therefore, calls for a UCD, presented in section 2.4.1, of DDSs developed in a participatory manner by businesses and consumers as part of the Business-Consumer Alignment. To do this, companies need to understand the needs and obstacles of consumers, who, in turn, need to align their consumer strategy with their data strategy as part of the Consumer-Data Alignment. Consumer-Data Alignment is derived in section 2.3. To understand consumers and their obstacles in the development and design of DDSs and the alignment process in the context of Consumer-Data Alignment, the CI application fields of sustainable food and mobility are considered in the context of case studies. In the food sector,

recommender systems for critical consumers are examined in section 2.3.3, which, as DDSs, support the practices of a sustainable diet. In the mobility sector, section 2.3.4 presents the basics of algorithm-based reputation systems, which are intended to support trust-building in P2P carsharing to promote more sustainable mobility.

Business-Consumer Alignment is then presented in section 2.4, which describes the alignment process of businesses and consumers in the context of the development and design of DDSs. Finally, section 2.4.2 derives tools to support Business-Consumer Alignment, which are intended to help businesses and consumers in the context of UCD methods.

2.1 Business Models and Services in a Data-Driven World

2.1.1 Business Models

In general, business models describe the value-adding activities of a company, its business performance, and the opportunities for achieving competitive advantages [713].

Over time, different definitions of business models have emerged from various disciplines and their respective perspectives [403]. In particular, the disciplines of management [713, 610], information systems [658, 10], sustainability management [385], and innovation management [403] characterize the term.

Massa et al. [403] define a business model about the achievement of corporate goals as "[...] a description of an organization and how that organization functions in achieving its goals [...]." Stewart and Zhao [610] focus on profit generation, according to which a business model describes "[...] how a firm will make money and sustain its profit stream over time." Veit et al. [658, 430, 166] further emphasize the business model" as an intermediary between a company's strategy and its business processes".

According to Teece [628], a business model consists of three components: "The design or architecture of the value creation, delivery and capture mechanisms of a firm." The decomposition of a business model into components

is also supported by Osterwalder and Pigneur [462] or Johnson and Lafley [294]. The key elements here are value proposition, value creation and value capture [628, 462, 464, 430]. Value proposition includes defining stakeholders and satisfying their needs through products, services, or other solutions [20]. Value creation defines the activities that are undertaken to fulfill the value proposition. This involves using the company's resources and expertise [462]. Value capture deals with the revenue model and cost structure and shows how the company's offering is transformed into revenue [462]. Accordingly, it can also be understood as an interplay of different activities to satisfy the needs of stakeholders [712].

A business model is an instrument for describing, renewing, and evaluating the projects of start-ups and organizations [658], explaining the gap between business strategy and its processual implementation in a fast, complex technological progress [20]. Due to the complexity of such business models under the conditions of digital transformation, the business model is becoming the focus of research [658], as the constantly evolving competition between companies in the future will mainly take place at the business model level, it needs to be taken into account throughout the whole process of developing or optimizing a business model [213]. This also explains the growing research interest in DDBMs, which will be described below.

2.1.2 Data-Driven Business Models

In times of constantly growing data volumes, data is becoming increasingly important for companies and their competitiveness [253]. As described in the previous section, business models – as an instrument for explaining the gap between business strategy and its processual implementation – play a central role for companies to remain competitive in a complex market environment [213].

As a result, companies are increasingly trying to harness the potential of data by expanding their existing business model with data-driven elements or developing complete DDBMs [682].

A basic understanding of data should be given at the beginning for a more

detailed description of the problems behind DDBMs and their definition (see Fig. 4).

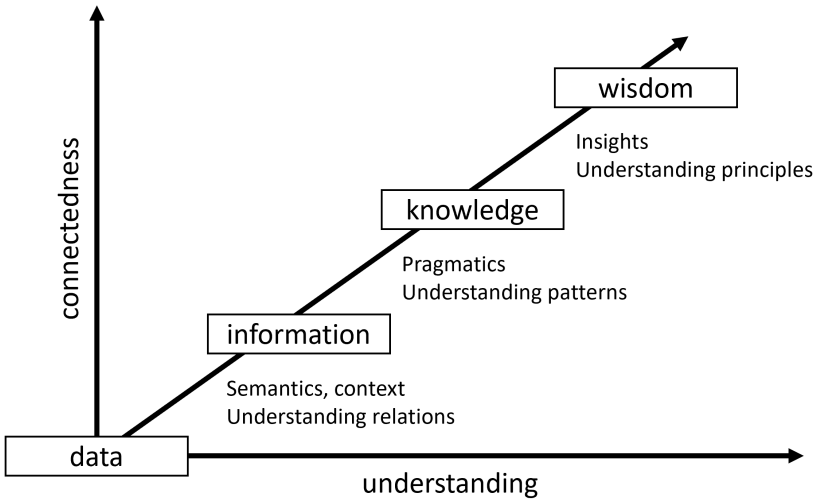


Figure 4: From Data to Wisdom according to [6, 55]

Data consists of different signs that follow a syntax. The syntax refers to the structure of the data and enables human users or computer systems to read, interpret, and understand it [208]. Since data has no context or enables interpretation [233], it has no meaning or value. They are unorganized, unprocessed [37, 116] and represent recorded descriptions of events, things, activities, and transactions [343]. It is information when data is contextualized using semantics so that it is meaningful and helpful to a person [343]. Information is usually aggregated data that supports decision-making [37]. If the information is linked, knowledge is created from it that instead of data as ”a property of things”, is ”a property of people that predisposes them to act in a particular way” [74]. This requires a pragmatic relationship between the information to ‘convey understanding, experience, accumulated learning, and expertise as they apply to a current problem or activity’ [651]. As the highest level of abstraction, wisdom represents ”the ability to see beyond the horizon” [37] through knowledge of the underlying principles from which insights can be gained [6, 55, 253].

The scientific literature presents various properties to describe data further, often called the ”4Vs”. These are volume, variety, velocity, and veracity [682]. Volume and veracity are more of a technical nature and, from the point of view of business informatics, do not lead to any significant changes in their architecture or treatment [487]. From a socio-technical perspective, however, the variety of data and its processing speed (velocity), i.e., more qualitative aspects, are of particular relevance [487, 442, 443]. The diversity of data creates new challenges for companies when it comes to processing and integrating it, as primarily structured internal company data is now combined with external data. The data sources are primarily IoT sensor data of intelligent machines, and increasingly human data - as a ‘walking data generator’ [409] - in the context of HCI [487, 443, 409].

Accordingly, the mere availability of data does not represent added value for companies. Only when at least information is generated from the data can the company benefit from it [487]. In line with the frequent statement that data is the new oil [524], the question arises as to how the value of data can be determined. Günther et al. [235] derive two socio-technical characteristics that make data valuable:

- Portability: Data can be used in different application contexts. On a technological level, this requires a digital infrastructure that enables use in other contexts. Furthermore, this impacts working practices and organizational models, affecting the social level.
- Interconnectivity: Data can be linked with other data, gaining new insights. This depends on advanced technical methods (e.g., algorithms, data integration platforms) and human collaboration. On an organizational level, it enables new business models and involves social and ethical considerations, e.g., data privacy.

DDBMs must be able to answer questions about how portability and interconnectivity can be used to gain knowledge and thus generate value from data.

Numerous definitions exist to describe DDBMs in more detail. At the meta-level, the three components of value proposition, value creation, and value

capture can be derived from the definitions in analogy to the components of a BM [561].

The value proposition is primarily defined by offering innovative services and smart services. In particular, the aim is to achieve customer loyalty by providing added value for customers [281]. Hunke et al. state that "[...] business models exploiting big data and analytics are expected to be one of the next frontiers of servitization." [281]. Similarly, Zolnowski et al. define DDBMs in terms of "[...] data acting as an enabler for the development of innovative services." [711].

In addition to the term "Data-Driven Business Model", there are other terms that are used as a more precise specification of the design of DDBMs [270] or an exact designation of the service [124]. It is interesting to note that, according to Hilbig et al., a "pure data-driven business model" (PDDBM) consists only of services and no longer offers any actual products and is very strongly focused on the user [270]. A PDDBM is defined as a "[...] business model [that] uses data as a key resource to generate any type of digital services by means of key processes such as data aggregation, data generation, data analytics, data exchange, data processing, data interpretation, data distribution and data visualization to create value for customers, users or stakeholders and to capture revenue. PDDBMs always rely on digital data and digital technology as fundamental enablers." [270]. Furthermore, based on Chen et al. [124], Schüritz and Satzger [561] define the service models "Data-as-a-Service" (DaaS) as "[Services where] DaaS providers gives its customers access to and/or aggregates a wide range of data [...]." and "Analytics-as-a-Service (AaaS) as "[...] services that offer a wide range of analytics services on demand".

However, the term (digital) business model innovation also includes definitions for DDBMs that address transforming a (traditional) product-based business model by enriching it with DDSs. According to Sorescu, (Data-Driven) Business Model Innovation "[...] can be designed around a process of collecting, organizing and summarizing internal or external data, to simplify the market research process and increase the probability of identifying an unmet consumer need or improving product assortment, product recommendations,

and promotional efforts." [593]. Sathananthan et al. also see four characteristics to specify the innovation process of (data-driven) digital business models [551]:

1. The motivating factor for the digital transformation of a business model is "[...] often relatable to interaction and/or offer to the customer"
2. "data is a key resource"
3. "data can be internally or externally derived"
4. "repurpose derived data/insights enables new innovation."

The requirements and activities for value creation in the context of DDBMs can also be derived from the above definitions. To realize data-driven value creation, digital data is the central resource [252, 711, 561], which can be obtained internally within the company [551] or externally through customers, products or services [551, 593]. Value-adding activities include, in particular, data aggregation [124, 270], data analysis [124, 270, 281, 561, 593], data processing [593], data interpretation [270], data exchange [270] and data visualization [270]. Other cooperation companies can also carry out individual activities as part of an ecosystem [661].

Hilbig et al. (2018) provide a taxonomy for differentiating the maturity level of DDBMs in three stages, which are differentiated by the dimensions of the degree of utilization of the data and the degree of digitalization [270]. Among other things, this shows the extent to which data is used to create value (see Fig. 5).

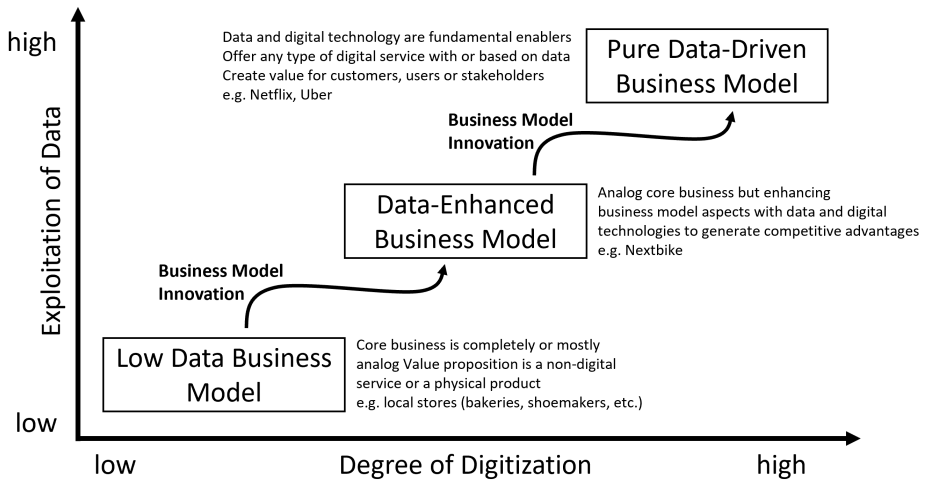


Figure 5: Taxonomy of Data-Driven Business Models according to [270]

Value is captured in the context of DDBMs by providing added value for customers, users, and stakeholders or by optimizing internal business processes [270, 593, 658]. Schuritz and Satziger (2016) also show that DDBMs do not necessarily have to cover all of these capturing components [561].

As it is usually almost impossible for established companies with physical goods to switch to a PDDBM, but these companies also inevitably have to reposition themselves by realigning their business model, this thesis focuses on the innovation part of the business model and defines Data-Driven Business Model Innovation as the transformation of a traditional, often product-based business model through the strategic, systematic use of data to create new value propositions, improve decision-making, optimize processes, and offer innovative DDSs. Data is accepted as a vital resource for developing DDSs or improving existing offerings by integrating data, analyses, insights, and technologies to respond to changing market needs and customer requirements. At its core, it is about aligning business strategy with data strategy and ensuring that data collection, processing, and utilization directly support the company's overall objectives, improve competitiveness, and enable scalable, consumer-centric solutions.

According to the characteristics of Sathananthan et al., [551], data is emphasized as a core resource for innovation and, following Zolnowski et al. [711], as an enabler for innovative services. Furthermore, the concept of Hilbig et al. [270], according to which development from a traditional, primarily product-based low data to a PDDBM without a direct physical product offering is supported by data innovation. Data-Driven Business Model Innovation helps to use the internet or external data to fulfill market and customer needs in a more targeted manner. Servitization and the associated improvement in customer loyalty through personalization and the creation of added value for customers, according to Hunke et al. [281], is also considered. Overall, the definition supports the idea of Business-Data Alignment by aligning business and data strategies to implement the business objectives effectively. Business-Data Alignment creates the basis for data to function as an integral part of the value chain.

For innovation of the business model for the increased use of data, DDSs - as initiated above - are necessary as a value proposition [711, 281]. A more precise categorization of DDSs is provided in the next section. In the context of value capture, this thesis focuses more on creating added value for the customer and less on optimizing internal business processes [593, 270].

The following section generally describes data-driven value creation as necessary for successful DDBMs. Since the creation of added value in the context of DDBMs is realized via provided services, such as recommender systems, section 2.1.3 then introduces DDSs in general and the DDSs specifically relevant to this thesis in the form of recommender and reputation systems.

2.1.3 Data-Driven Value Creation

Data-driven value creation uses data to make better decisions, optimize processes, and create personalized experiences for customers [409]. By analyzing customer data, companies can identify patterns and trends that enable them to better tailor products and services to their customer's needs and preferences [28]. This should increase customer satisfaction and loyalty, which in turn increases the company's competitiveness [492]. Customers should also benefit from data-driven value creation. By analyzing customer data,

companies can create personalized offers and recommendations tailored to the customer's individual preferences [534]. It is therefore also highly relevant that consumer privacy is safeguarded [290] and that consumers are made aware of potential nudging mechanisms [558] through DDSs, as otherwise, it is a double-edged sword that only promises superficial added value [559]. At first glance, disclosing potential nudging mechanisms appears negative for companies, as consumers may not respond to nudging. However, a disclosed nudging mechanism, which does not seem profitable, can still be beneficial by increasing customer loyalty and the number of users [47].

However, to fully utilize data-driven value creation, companies must introduce an active data-driven corporate culture, provide the necessary technological infrastructure and skills, and ensure data protection and security [346]. Businesses must effectively process, analyze and interpret large amounts of data to gain meaningful insights [211, 346]. However, many companies and decision-makers do not yet have the necessary understanding of data, which is why increasing data literacy and still missing suitable tools to support decision-making are needed to make the right decisions as part of the development of DDBMs and DDS [501, 70].

The next section describes DDSs as the actual value-creating part of DDBMs and the specific manifestations relevant to this thesis.

2.1.4 Data-Driven Services

A class of information systems known as Decision Support Systems (DSS) intends to aggregate, store, retrieve, and process data analytically to help decision-makers in organizational settings. The large corpus of literature on the topic demonstrates the significant attention that such systems have received in information systems research [120, 409, 675]. Scientific discourse specifies a tripartite structure as the foundation for DSS development to optimize operational efficiency. This framework includes the steps before gathering data, concluding analysis, and taking action [675].

In addition to using data to optimize internal processes or efficiencies, businesses have recognized the utilization of data to develop or enhance their ser-

vices and, in turn, provide substantial value for customers [151, 61]. Many new terms and concepts are being presented due to the field's infancy. The concept of "smart services" defines services supported by smart products equipped with interfaces, sensors, actuators, and connectivity capabilities that allow them to interact and communicate autonomously [61, 161, 476]. Another term is "analytics-based services", which utilise analytics and data to increase customer value by empowering them to make better decisions, solve complex problems, and enhance their ability to accomplish their objectives with greater effectiveness or efficiency [280]. According to Hunke et al. [280], using data and analytics to create unique value propositions for customers within services is strategically significant. This can be achieved by giving customers access to functions driven by data and analytics to improve their decision-making abilities as part of stand-alone items or integrated parts of already available products or services. In the current research, the term "Data-Driven Service" is consistently acknowledged, notwithstanding the diversity of terminology and conceptual frameworks [279, 562, 297, 280]. This thesis adheres to Schüritz et al.'s definition of DDSs, which is "the use of data and analytics to support the decision-making process of the customer via data and analytics-based features and experiences in the form of a stand-alone offering or bundled with an existing product or service" [560]. This definition considers the consumer and makes it possible to differentiate between PDDBMs and data-enhanced DDBMs.

This thesis analyses two specific DDSs currently highly relevant in the context of the digital consumption fields of food and mobility under consideration: recommender and algorithm-based reputation systems. Recommender systems play a central role in the personalization of consumers' food decisions [193], and algorithm-based reputation systems shape the development of trust in the mobility sector, in which sustainable practices such as peer-to-peer carsharing are to be strengthened by these systems, mainly due to information asymmetries [469]. The aim is to derive insights into how DDSs can be designed user-centered to support businesses and consumers in this process with suitable tools. This is necessary because, according to Blöcher et al. [70], companies in particular "often lack the knowledge of how to fully take advantage of data and analytics" and "concepts and methods of how new

services can be designed from this data is [sic] still developing".

2.2 Business-Data Alignment

In earlier definitions, business-IT alignment is described as the strategic application of information technology (IT) to support and enhance business strategies, goals, and requirements. This involves ensuring that IT initiatives are both timely and appropriate in a synergistic way with the overall business objectives [386]. Other definitions describe it at a strategic level as the extent to which IT objectives are aligned to support and are supported by business objectives [130, 620]. Furthermore, earlier research mainly focused on aligning both strategies rather than including the operational aspects [48]. The viewpoint is strongly presented in Carr's critical article "IT Doesn't Matter," where the author emphasizes the increasing importance of IT for maintaining competitiveness. Nonetheless, he posits that IT will diminish in strategic significance as it becomes universally utilized, negating any competitive advantage [108, 159].

In advancing digitalization, the focus of business-IT alignment has evolved. It is now widely recognized that IT should no longer be perceived simply as a support factor but rather as an active driver that significantly influences and shapes the business and its objectives. [381, 96]. As Davenport and Westerman (2018) emphasize, the digital transformation of business-IT alignment is "multifaceted and diffuse, and doesn't just involve technology. Digital transformation is an ongoing process of changing the way you do business" [154].

As socio-technical systems, the triangulation between people, technology, and tasks is central to the design of information and IT systems [329]. Accordingly, Charoensuk et al. (2014) argue that in addition to the strategic level, the operational level must also be considered [118]. Furthermore, it follows that a distinction should be made between a social and a technical component: While the social component comprises the expertise, skills, and abilities of employees [329, 536], the technical component describes the IT architecture, infrastructure, software and databases [548, 329].

With their business model ontology and the business model canvas based on

it, Osterwalder and Pigneur (2003, 2010, 2011) promote alignment between information systems and business alignment at the business model level, as the modeling of social systems or an ontology reveals the relevant elements and their interdependencies, makes them communicable and shareable, creates a common understanding of the business logic and facilitates change [460, 462, 463].

Gajardo and La Paz (2019) point out that business-IT alignment is still an unsolved problem, as most models assume static rather than dynamic alignment, and ”these models do not guide organizations on how to improve in that process”. Nevertheless, many organizations often ignore the scientifically proven ways of aligning business strategy with IT and IS to drive innovation and improvement, instead continuing to focus on aligning IT with business strategy [480, 209].

Digitalization also adds new factors that make the problem more complex [209]. Alignment is no longer taking place between business and IT strategies. However, these are being integrated into a digital strategy, resulting in a ”blurring of the distinction between business and IT strategies” and requiring the operational level to be aligned accordingly [699]. In the literature, this is also formulated as a co-evolution process, which brings the external (strategic) and internal (operational) levels closer together (see Fig. 6).

Amarilli et al. (2023) differentiate between six factors, each with different characteristics, which influence the alignment process, but some of which share intersections and correlate with each other [508, 24]. They include (1) top management leadership and commitment, (2) communication structures, (3) relational mechanisms, (4) operational conditions, (5) environmental factors, and (6) shared domain knowledge [24]. It turns out that many sources are mainly concerned with the operational conditions, which include the flexibility of the IT infrastructure in terms of its connectivity, compatibility and modularity, governance processes, personnel competencies, and skills [444, 282, 540, 24]. The focus is on communication in general, which should ensure understanding and trust [386, 156] via a common language [508]. The focus on operational conditions is also supported by other sources that define the flexibility of the IT infrastructure, in particular through connectivity and

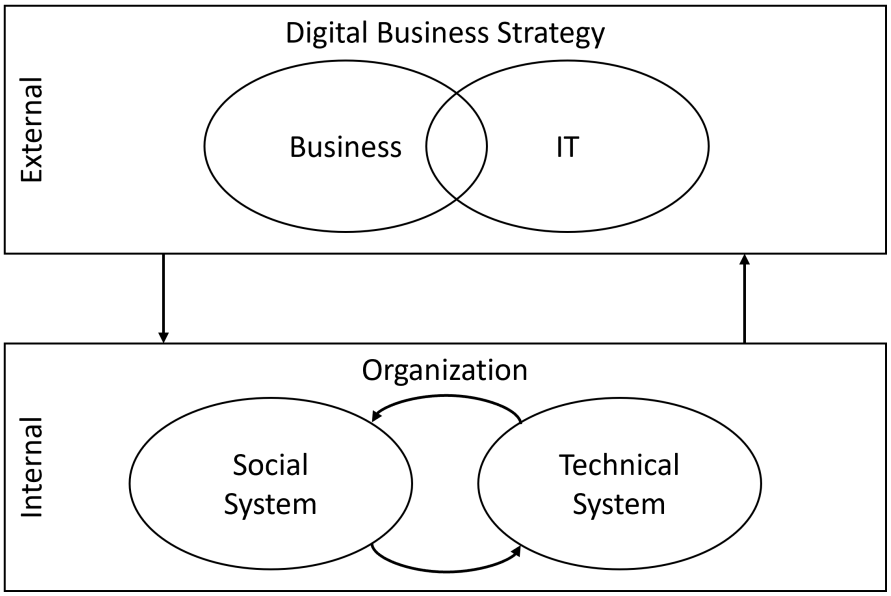


Figure 6: Co-Evolution Process of Aligning Strategic and Operational Level according to [24]

integration, consisting of both process and data integration [510].

The term ”business-IT alignment” comes from when IT was the focus. The merging of It with business strategy to form a digital strategy has shown that digitalization is leading to a new focus on alignment [24]. The changed requirements for aligning strategies are reflected in the evolving leadership responsibilities of the Chief Digital Officer. Distinguished from the domain of the Chief Information Officer, the Chief Digital Officer's purview does not encompass the implementation and upkeep of IT operations. Instead, the role emphasizes strategically driving the digital transformation within the organization to promote IT innovations and foster digital business models [207].

Similarly, due to the increasing relevance of data to corporate decisions, the role of the Chief Data Officer, who is responsible for the utilization of data to create Data-Driven Business (Model) Innovation, has emerged [364]. Accordingly, the data strategy has also emerged, ensuring that data is increasingly becoming the company's focus and that a data-driven corporate culture actively promotes the data strategy in corporate decisions and business models [236]. Furthermore, research has shown that when organizations see data

strategy as an IT strategy, it leads to failure of strategy implementation, loss of management confidence, and inability to generate valuable results. With this in mind, a distinction should be made between a technology-driven and a business-driven data strategy. While the technology-driven data strategy focuses in particular on data management and the technological possibilities for achieving ”data quality, integrity, accessibility and security”, the business-driven data strategy sees data as ”core and precious organisational assets that are fundamental to the survival and success of a business” [236].

However, IT work also differs from the work of a data scientist, for example, at an operational level. While IT employees essentially deal with providing the required technology [207], the data scientist concentrates on utilizing the data and gaining insights from the data [153]. The skills required, therefore, also differ significantly, so a data scientist must have profound knowledge in maths, statistics, and domains in addition to knowledge of programming and computer science [467]. Furthermore, it is evident that the consequence of digitization is the availability of data [700] and that IT involves storing, providing, and managing the data as well as the technology to process it [695, 172]. Data science is the processing of this data into information [497, 172].

It is relevant to emphasize that a new term for the interaction between business and data in the form of Business-Data Alignment is required. The academic literature rarely uses the term, and no definition exists. Lee et al. (2014) point out in the context of the Chief Data Officer role that this is responsible for strengthening Business-Data Alignment, but without defining this more precisely [364]. Yang (2020) sees Business-Data Alignment as ”fundamental to ensure further exploitation of big data in their organization” [696].

In the face of a pending definition, this thesis, based on Amarilli et al. (2023) [24] and Hilbig et al. (2018) [270], characterize Business-Data Alignment as the merging of IT, digital and business strategy into a data strategy that must be aligned with the operational business. In the context of DDBMs, a distinction can be made at the strategic business model level between an alignment of the data strategy with the business strategy in the form of data-enhanced business models and an alignment of the business strategy with the data strategy in the form of pure DDBMs (see Fig. 7).

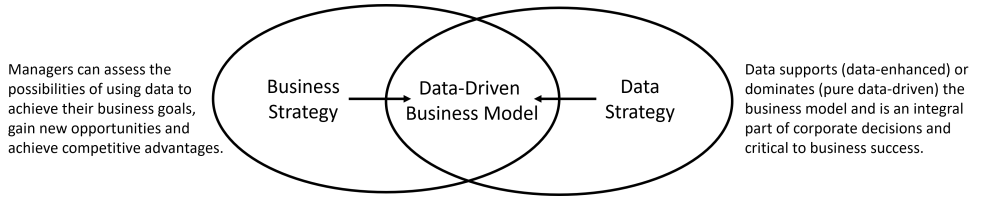


Figure 7: Business-Data Alignment on a Data-Driven Business Model Level according to [461]

Since data integration is a vital part of successful Business-Data Alignment from a technical perspective, enabling the development of DDSs [510], the following section 2.2.1 presents the related work to the case study on data integration practices. As DDSs represent an innovative use of data to create value-added offerings and businesses often lack knowledge about handling data [501, 333, 70], section 2.2.2 introduces the case study basics regarding developing a recommender system in a business setting as a form of a DDS.

2.2.1 Integrating Data for DDSs

To enable DDBMs and DDSs, companies must collect data. In today's diverse landscape of cloud-based application system ecosystems, it is increasingly difficult for companies to combine the required data from various applications to make it usable [248] for DDSs. Connectors are needed to link the individual, primarily Software as a Service (SaaS) solutions, which connect the Application Programming Interfaces (APIs) with the company system and thus enable data exchange [370]. As this is a complex process, Integration Platforms as a Service (IPaaS) models attempt to provide a "suite of cloud services enabling development, execution and governance of integration flows connecting any combination of on premises and cloud-based processes, services, applications and data within individual, or across multiple, organizations" [482]. By connecting businesses with SaaS providers, IPaaS platforms are multi-sided markets aiming to reduce the integration effort for companies by developing, managing, and executing integration processes [401]. Nevertheless, there are issues regarding their strategical and practical implementa-

tion [178].

One main task of IPaaS solutions is to organize data synchronization and mapping across the various applications and systems, including data preparation, transformation, and migration into existing infrastructures [708]. Therefore, IPaaS reduce the number of necessary connectors through the platform structure by enabling the development of only one connector to all application and system integrations [493]. Furthermore, this allows the orchestration of the various applications along business processes, supporting the development of individual connectors, data mappings, and the execution of automated processes [178].

Research on IPaaS mainly focuses on technical or legal aspects, which is why we need a more nuanced understanding of companies' adoption factors. Ebert et al. (2017), therefore, claim that "critical success factors for using IPaaS [...] have not been investigated" [178]. Understanding companies' current practices is essential to derive adoption factors and hurdles to their data integrations [178].

2.2.2 Recommender Systems in Food Retailing

Food retailing is undergoing significant changes that go far beyond digitization. As consumers' lifestyles shift towards healthier, more ecological, ethical, and social values, food increasingly becomes an expression of individual identity. This leads to an increasing variety of food products and an information overload for consumers [283, 352, 605], especially in online retailing. Ensuring a personalized customer experience, transaction, behavioral, and evaluation data collected in online retail is used to provide relevant content through recommender systems [619].

2.2.2.1 General Recommender systems support users in their (future) decisions based on their previous usage history and that of other users. In principle, recommender systems can be differentiated into non-individual, collaborative, content-based, knowledge-based, demographic, and hybrid filters (see also Table 1) [13, 73, 477], whereby the term 'filter' defines the process

of identifying the most likely matching options that meet the preferences or needs of a user from a variety of possible choices (e.g. food products).

Ansatz	**Methodik**	**Input**
Collaborative	Recommendations are made using a collaborative approach that uses the ratings and actions of similar users as well as the behavior of the active users themselves.	User ratings from active and all users
Content-based	Recommendations are based on the items (or their properties) that an active user has preferred in the context of their previous ratings and behavior.	User ratings and item properties
Knowledge-based	Recommendations are generated based on the explicit specification of desired item properties. This requires domain knowledge that supports the user information.	item properties, domain knowledge

Table 1: Approaches to Personalised Recommender Systems according to [13]

While non-individual recommendations are the same for all users and therefore have no personalization of the respective products (e.g., the most frequently clicked products) [13], collaborative filtering examines the preferences of different users based on their consumption and usage behavior to identify similar people or items (e.g., films on Netflix). Recommendations are either based on the similarity of two items in terms of their user ratings (”item-based”) or based on the ratings of similar users (”user-based”) [376, 549, 707].

Content-based filters generate recommendations by analyzing the properties and content of consumed items. For instance, users may receive suggestions for items similar to previously consumed ones, such as articles on science and technology [73, 132, 656]. Knowledge-based filters are based on the satisfaction of customer needs through specifically defined product features.

Explicit rules are used to generate recommendations (e.g. specifications such as size, min., and max. price, or zip code when buying a house) [13].

Demographic filtering generates recommendations by analyzing diverse socio-demographic attributes of users, including factors like age, gender, or nationality [636, 477]. Hybrid approaches are a combination of different filters. These are often used to circumvent problems with individual methods and to increase the accuracy and efficiency of filtering [13, 636, 73].

2.2.2.2 Repeat Purchase Recommender Systems In recent years, collaborative and content-based recommendation processes have become established - primarily driven by online retail - which utilize historical purchase data [89, 376]. Nevertheless, recommender systems are not omnipresent in food retailing as the decisions of consumers are often habituated, and the question is often not what to buy but when to buy it [296, 182]. Therefore, a repeat purchase can be described as any situation in which a customer buys more than one product unit [62]. Ehrenberg (2000) discusses three ways of describing repeat purchases: Customers may purchase an item in more than one purchase in a given period, purchase an item in more than one time period, or purchase multiple units in the same purchase [182].

Established studies and models for repeat purchase recommendations concentrate on specific products or brands [192, 431]. However, there are first approaches for recommender systems for repeat purchase behaviour using online tracked data [62, 165].

Bhagat et al. (2018) assume that the temporal dependencies between repeat purchases depend on when a product was last purchased and how quickly the customer runs out of it [62]. Accordingly, a customer will repurchase a product if he has already purchased it. However, it can also be stated that the attractiveness of a product for a repeated purchase directly following the purchase of this product is rather low. Therefore, a repeated purchase can be seen as a periodic phenomenon subject to a specific time interval. Personalization can be implemented via a Bayesian Hierarchical Model, which enables individual purchase rates [62, 129, 158, 229]

As the models discussed tend to be business- or technology-orientated, since

they hardly have the ‘average customer’ in mind, the following section on Consumer-Data Alignment presents recommender systems for critical consumers that also take into account the complex practices of (niche) consumers [226, 586, 599, 303].

2.3 Consumer-Data Alignment

Since the term Consumer-Data Alignment has not been defined in the literature in this way, this thesis motivates derivation via Business-Data Alignment as a composition of a consumer strategy and a data strategy, which must be harmonized. The relevance of the alignment of the consumer and data strategy by the consumer is especially evident in the limited awareness and understanding of consumers to assess the added value of DDS, which is caused by a lack of transparency about data collection and processing [397], the functionality [387] as well as a perceived lack of control over their data [397, 113].

Following the derivation of the terms consumer and data strategy in section 2.3.1 and 2.3.2 and their link to the application fields of food and mobility, the specific basics of the case studies on recommender systems for critical consumers examined in the context of Consumer-Data Alignment in section 2.3.3 and algorithm-based reputation systems in section 2.3.4 are presented.

2.3.1 Consumer Strategies through Consumer Practices

Most literature that deals with the topic of a ”consumer strategy” is mainly related to marketing research or from an institution-based perspective, which sees it as something that needs to be defined from a business perspective or derived from an institutional strategy [245, 202, 399, 201, 507, 382]. It is described as a conceptual framework designed to derive consumer needs, preferences, and desires to understand and fulfill them. The aim is to build long-term relationships with consumers to ensure the success of the company and customer satisfaction and loyalty [41, 622]. In this regard, consumer behavior research aims to provide an understanding of consumers by analyzing the processes involved in selecting, purchasing, using, or disposing of products and services to satisfy needs and identify preferences [591]. Customer

journey mapping, which describes consumers' interactions with the company across different phases and touchpoints, is also crucial in this research [365].

Research on food consumption behavior suggests that consumers choose a habituated set of products when planning their purchases, which is usually not updated and tends not to need to be accompanied by a DDS such as a recommender system. While planning serves as orientation, strategy is how a goal can be achieved. Accordingly, planning and strategy are closely linked [29]. Orientation is sufficient in the case of a habituated product set, as the routines of the consumers already define the strategy, but without thinking this through strategically [182]. However, suppose consumers develop new practices or want to achieve goals, and the product set under consideration changes accordingly. In that case, it is necessary to change routines and use DDS as possible touchpoints as a strategic tool to implement these practices. Personalized recommender systems support this by reducing the complexity of the search, setting up suitable product alternatives and filtering information to improve the search for niche products in dominant consumer infrastructures, which also strengthens the needs of critical consumers, for example. It can be summarized that the strategy of consumers is thus shaped on a meta-level by the actual consumption practices and behaviors of consumers. Warde (2016) [672] also argues that "consumption occurs as items are appropriated in the course of engaging in particular practices and that being a competent practitioner requires the appropriation of the requisite services, possession of appropriate tools, and devotion of a suitable level of attention to the conduct of the practice. Such a view [...] also emphasizes that practices are internally differentiated and dynamic."

In mobility research, we see that current consumer mobility practices are primarily relying on motorized private transport, which is responsible for a significant proportion of energy consumption and CO2 emissions in terms of sustainability [348, 689, 206]. In addition, this practice results in other problems such as congestion, noise pollution, and pollution. However, the consumer strategy here is aimed at maintaining comfort and flexibility and therefore forgoes the use of more sustainable mobility alternatives, such as public transport [689, 223, 582, 655]. This is due to the often poor networking of mobility services, which makes their use cumbersome and unattractive

[309]. To overcome these mobility problems, shared mobility is currently becoming increasingly relevant [79], which also addresses the inefficient use of resources in private cars, which remain unused most of the time [223, 689].

Shared mobility can reduce private transport by allowing several people to share a vehicle, thus reducing the number of vehicles required and resulting in less traffic and emissions. Vehicles are utilized more, which can contribute to more careful use of resources [79, 684]. In this context, P2P carsharing platforms have emerged that focus on the direct rental of vehicles between private individuals via a digital platform and thus aim to change mobility practices towards a more sustainable form [684]. P2P carsharing platforms bring private car owners and potential renters together, with most steps of the rental process, from vehicle search and booking to payment, handled via the platform [82]. However, building trust between users is a crucial challenge of P2P carsharing platforms, as there are usually no prior relationships between them [602]. This poses a significant challenge, especially in the online context, as there are no usual trust factors, such as personal interaction or getting to know each other directly. Users must rely on the information they receive online via the platform [82, 602].

To address the problem of trust and promote more sustainable mobility practices, P2P platforms rely on reputation mechanisms. These mechanisms are intended to create transparency and enable users to assess the trustworthiness of other users better [259, 633]. One established mechanism is user ratings, where users can rate the other party after completing a transaction. These ratings are usually aggregated in the form of a rating and displayed to other users to help them make a decision [602]. However, user ratings are susceptible to manipulation, e.g., through fake ratings or one-sided negative ratings due to reciprocity effects [179, 107, 638, 420, 413]. DDS in the form of algorithm-based reputation systems can build trust and transform mobility practices to maintain comfort and flexibility, as they represent a rating based on actual driving data collected via telematics more objectively and less susceptible to manipulation [602]. However, as research on algorithm-based reputation systems is still limited, it is unclear how users would perceive and design such algorithm-based reputation systems to increase their acceptance and use [602].

2.3.2 Uncertainty Regarding Consumer Data Strategies

In the age of big data, people are characterized as "walking data generator[s]" [409]. Activities leave behind data that reveal our preferences, views, behaviors, and intentions. Individuals and businesses can benefit from harnessing this data by processing and analyzing it using data analytics and providing DDSs [409]. Nevertheless, there is also a risk of abuse of personal data and the associated loss of autonomy [356, 137]. However, due to the development and increasing penetration of information technologies into consumers' lives, the collection and use of personal data often goes unnoticed [488, 113]. In market terms, this phenomenon can be described as asymmetric and incomplete information, which unsettles consumers about the true intentions (hidden characteristics) of companies and other stakeholders [135, 439, 8]. Acquisti et al. (2015) [8] differentiate between two fundamental problems concerning uncertainty in the context of consumers' data strategy: On the one hand, damages caused by erroneous behavior can rarely be tangibly measured by consumers [488] and, on the other hand, privacy is almost exclusively associated with trade-offs [344, 8]. It is pointed out here that creating comprehensibility and increased visualization of corresponding decisions can positively affect usage intentions [650]. Therefore, several approaches rely on visualizations and dashboard concepts, which help users to understand which data is collected [113, 488]. Nevertheless, these concepts only allow users to see what happens after they engage with a DDS and not before using it.

Consumers are usually unclear about their respective data strategies or do not implement them consistently. In particular, the privacy paradox - the discrepancy between consumers' intentions to make great efforts to protect their privacy and their actual behavior, which is usually characterized by the disregard of data protection efforts in the case of expected cost or benefit advantages - can be cited here [321, 451, 289]. However, cost-benefit considerations are only sometimes straightforward, as consumers cannot rationally grasp the necessary calculations [344, 318, 8] and many factors play a role in consumer behavior, such as social norms and emotions [7]. It also shows that many decisions in the area of privacy are context-dependent and that cues in their living environment guide consumers. Thus, privacy strategies "differ according to

time and place [...] and vary according to situation" [8]. Research claims for data literacy so that consumers can develop a data strategy and realize the value of their data. Consumers, therefore, need to understand and know how a DDS works, which data is analyzed, how a company uses it in which context, and for which purpose [488, 113].

The following sections 2.3.3 and 2.3.4 therefore introduce the basics of recommender systems for critical consumers and algorithm-based reputation systems for the case studies conducted in this thesis in part III and the analysis for Consumer-Data Alignment.

2.3.3 Recommender Systems for Critical Consumerism

Critical consumption addresses the environmental or social problems of the food system, such as organic food, the reduction of animal-based foods and thus the promotion of plant-based foods, or fair trade [224]. From a design perspective, Håkansson et al. [239] describe that supporting the personalization of the 'right' values is necessary instead of enforcing a normative understanding. However, the complexity of sustainable practices is particularly evident in cultural aspects as well as the personal situation, e.g., budget, allergies, and surrounding infrastructures with different types of food retailers and tens of thousands of products [224, 109, 475].

Recommender systems reduce complexity and information overload in the decision-making process [13]. Currently, however, these are primarily designed to serve the interests of suppliers by increasing and supporting consumption [226, 586]. They do not promote critical consumption, nor do they support the complexity and diversity of diets and practices, e.g., they do not work well for vegan or vegetarian diets [644], as they predominantly prioritize popular foods rather than focusing on niches and do not consider important critical implications such as allergies. They predominantly reinforce current consumer behavior rather than helping to explore and make actionable new goals and support critical consumers with their complex values, ethics, and constraints [599, 303].

Trang et al. [643] differentiate between four types in the domain of food rec-

ommender systems. The first type generates recommendations of recipes or products based on their similarity to consumer preferences [643, 183, 203, 255]. Type two incorporates the recommendations of (health) professionals and their ratings [3, 653]. Type three establishes a link between consumer food preferences and nutritional recommendations, and type four generates recommendations for consumer groups, e.g., families [643, 184]. Content-based, collaborative filtering or hybrid approaches are mainly used here [643, 648]. The recommendations are primarily based on the popularity or similarity of foods and focus on the 'average customer'. For this reason, food recommender systems do not work well for critical consumers, as they do not correspond to the average. Therefore, food recommender systems must work better for critical consumers as they are not average [644].

Therefore, current approaches start with integrating domain knowledge in the context of knowledge-based recommender systems to recommend food based on a holistic user profile that takes individual health aspects into account [436]. However, there is still a research gap for food recommender systems to consider the preferences of individual consumers in a profile [643] and incorporate personalized critical consumer values and ethics [303], which is a problem of missing Business-Consumer Alignment. Karpati et al. [303] already refer to the first food recommender systems that use ethical approaches but still need to be personalized. Therefore, consumers will only find the "right" product if they share the values of the recommender system provider. Furthermore, the underlying ethical guidelines, algorithms, and business models are often opaque [303, 150]. For this reason, transparency and trust are advocated to personalize consumers' ethical values and needs [303].

2.3.4 Algorithm-Based Reputation Systems in Peer-to-Peer Carsharing

Since a growing number of consumers nowadays are giving up individual ownership for demand-based access to goods [222, 317, 358, 455], business models have emerged where consumers pay for this access on-demand [684, 466]. A well-known example in this context is carsharing [434], which includes peer-to-peer (P2P) platforms in addition to a commercial business [684]. As multi-sided markets, these P2P platforms orchestrate rentals be-

tween providers of rental vehicles and renters. Therefore, the lack of trust in other ”buyers” or ”sellers” causes the rejection of these P2P platforms. [83, 469]. Transparency, reputation, and trust are crucial to reduce the information asymmetry about renters' true behavior with the vehicle and, thus, to promote the success of the P2P sharing economy. [53, 84, 259].

Reputation systems collect, aggregate and distribute data about an entity to predict and characterize this entity [513]. By providing this reputation data, users receive decision support regarding the trust associated with this entity. This is complemented by a socially corrective function that incentivizes positive and disincentivizes negative reputation and promotes positive behavior. Reputation data is exchanged between platform users to mimic characteristics of a long-term peer relationship [513, 512].

A reference model for reputation systems by Hendrikx et al. (2015) [262] can be adapted to the context of P2P carsharing (see Fig. 8).

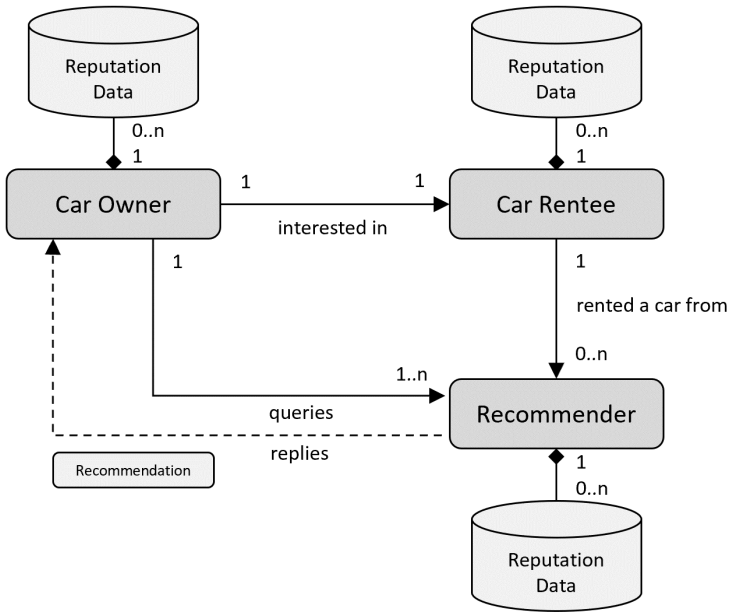


Figure 8: Reputation System Reference Model for P2P Carsharing Platforms adapted from [262]

The interactions of the parties involved on a P2P carsharing platform can be described as follows: The car owner (trustor) is the party that wants to inter-

act with a target entity, the car renter (trustee), and must trust it [163, 328]. To do this, the car owner must check the reputation of the car rentee, for which it first accesses the self-provided information from the car owner and the internal reputation data about this person if this is available from a previous interaction. If there were no previous interactions, the car owner would consult recommenders for their opinion [531, 539] who have had prior interactions with the car rentee. This recommender is an entity that either draws on reputation data from its own experience with the car rentee based on historical interactions or as a system that has observed an interaction between two parties or gathers information from other sources [380]. If this recommender has adequate information about the car rentee, it responds with a recommendation, also called feedback or rating. Utilizing this recommendation, the car owner can make a trust decision [262].

Reputation data can be differentiated into self-provided, peer-provided, and computationally provided information [602]. Showing self-provided (e.g., personal information) or peer-provided information (e.g., reviews and ratings) are traditional mechanisms to increase trust [57, 278, 301]. Nevertheless, especially peer ratings consist of various biases [179, 107, 638]. Therefore, computational-provided information could serve as a third trust mechanism [210, 249, 604, 669] as they ground on objective measures [602].

Resnick and Zeckhauser (2002) [513] describe that trust systems must 1) provide information for distinguishing between trustworthy and untrustworthy actors, 2) promote integrity and trustworthiness, and 3) deter untrustworthy people.

In P2P carsharing, data from connected car technology [144, 238, 357, 424, 603, 618], e.g., monitoring driving behavior, could build trust by algorithmically processing it into a rating. Such an algorithm-based rating could increase confidence in the potential rentee based on objective standards. However, this information is currently only used by car insurance companies for risk assessment and individual insurance tariffs [417], where defensive driving is promoted, offering the ”good” drivers cost reductions [238, 432, 677, 701].

Algorithm-based ratings require telematics data provided by the trustees concerned. This raises the question of what data is disclosed to be able to rent a

car and poses challenges in terms of privacy [606].

2.4 Business-Consumer Alignment

Business-Data Alignment deals with aligning the business strategy with a company's data strategy. This occurs at both strategic and operational levels (see Fig. 6). The internal organizational level reflects the business's dependencies as a technical and social system [24]. The social system comprises a company's human resources and represents the know-how and skills of the employees [80]. Meanwhile, the technical system defines the available technical resources of a company, including its data capabilities [617, 472]. Companies need to know the potential in their data assets and the existing know-how of their employees and understand how they can leverage this with analytics to develop successful value-added business models and services [472, 442]. As many decision-makers do not yet consider themselves competent enough in dealing with data [501] and the implementation of a data strategy is often not yet sufficiently in place [140], suitable concepts and methods are needed, how companies can be supported in the development and design of DDS to overcome the problems associated with the alignment of business and data strategy [266, 70] and to be able to communicate these correctly in the context of the organization as a social system [24].

At the same time, Consumer-Data Alignment represents the alignment of the consumer strategy, which is characterized by its consumption practices [672], with the data strategy. As already described in section 2.3, deriving a data strategy for the consumer is a challenge. Businesses must be aware of the expected added value resulting from the use of a DDS to support consumption practices as part of the consumer strategy. Furthermore, an increased data literacy and knowledge of data sourcing, processing, functionality, and control is required [397, 113, 488].

Overall, it can be deduced that an often implicit [596] negotiation process takes place between consumers and businesses, which aims to harmonize the interests of both parties [446]. This thesis defines Business-Consumer Alignment as negotiation processes - as "dynamic processes in which the parties

involved communicate to exchange offers, make concessions, [...] or otherwise influence each other to reach an agreement" [197] - contributing to DDSs to create added value for both sides in the area of conflicting interests [196, 446].

Since Business-Consumer Alignment to find a joint fit is a complex process that requires both sides to understand the added value to be created in the context of DDSs and, thus, Business-Data on the one hand and Consumer-Data Alignment on the other, suitable methods and tools to support the negotiation process for both businesses and consumers are still lacking [70, 125]. UCD [666, 4, 488] is a classic instrument for taking consumer interests into account in the development and design process to support the negotiation process in the context of Business-Consumer Alignment and to create joint added value. Section 2.4.1 therefore introduces the basics of UCD and derives the necessity of strengthening consumer integration within the framework of Business-Consumer Alignment. To support businesses in the context of data-driven value creation, section 2.4.2 also presents the basics of the tools developed as part of this thesis.

2.4.1 User-Centered Service Design

In the 1980s, Donald Norman coined the term user-centered design (UCD) [453] as an iterative framework and philosophy aiming to integrate the requirements and goals of the end users into the design process. The understanding of UCD in HCI has evolved from a narrow view in software engineering psychology to an approach that requires "more collaborative efforts to understand what it means to develop systems that people value" [302]. To this end, UCD calls for the continuous involvement of users to understand their behaviors, needs, and preferences and to consider them when developing solutions [400, 4]. UCD encompasses various design processes but only postulates that users *are* involved in the design process. However, the degree to which users are engaged in the design process in practice, depends on the company and the designers. Requirements engineering is often carried out here, and the involvement in the design of the solutions remains limited [4]. In the context of the negotiation process of Business-Consumer Alignment, the

user thus tends to play a passive role, and the focus remains on the fulfillment of business goals, taking into account usage requirements and feedback, but without a deeper integration of consumer practices into the strategic design [4, 400, 5].

To further democratize design processes, Participatory Design (PD) emerged as a collaborative approach focusing on user participation in an organizational and social context [231]. With its roots in Scandinavia, the first PD concepts related to Marxist theory explained that ”emancipatory practice must not only aim at changing the use of artifacts, but also their technical design, the design process, and the relation between the design process and the labor process the artifacts are designed for. When these changes violate the production process as a valorization process there are social conflicts inherent in the capitalist mode of production” [181]. By considering conflicts as native design characteristics, detecting social and organizational obstacles is improved [231]. PD initially emerged in the context of work environments in which workers were empowered to participate in system development [581, 68], but evolved - also in the wake of the increasingly blurring boundaries between consumers and producers (also neologistically called ‘prosumers’) in the context of the advancing internet growth - to consider a broader view of users and application contexts [588]. In this way, PD democratizes the understanding of the role of consumers as equal partners in the design process, whereby they are actively involved in all phases.

As more and more (data-driven) services permeate the lives of consumers, service design (SD) is becoming increasingly relevant [657]. SD emerged in the 1980s in marketing and management from the increasing importance of services and the need to create positive experiences [390, 576]. At the same time, designers used this term to describe processes that support the conception, design, and development of services [398, 545]. SD views services as complex systems consisting of various elements such as processes, touchpoints, technologies, and employees and aims to ensure that all aspects and the interactions that take place there are harmonized [611] to focus on joint value creation through active user participation [657]. SD draws on PD methods to understand users’ needs and supplements these with strategies for designing services, including developing customer journeys, service blueprints,

and other tools for visualizing services [611].

Because PD promotes joint negotiation [71], both the company's strategic goals and the needs and expectations of consumers can be incorporated into a dynamic, reciprocal negotiation process of Business-Consumer Alignment through "direct interaction [...] to articulate, create, and develop users' ideas and visions" [565]. SD can also expand the understanding of participation in terms of the design of DDSs, as it goes beyond the consideration of user integration and does not regard users as merely passive recipients of services but recognizes that they actively contribute to value creation by contributing resources and participating in a joint design [611, 454, 657].

Therefore, e.g., Saad-Sulonen et al. [537] see the combination of PD and SD as "an opportunity to combine existing participatory and service design approaches to participation in the way they weave connections between design, IT, digitization and democracy [...]". The result is that consumers contribute to requirements, better understand the business strategy's impact, and play an active role in shaping it. This method supports and corresponds to the idea of shared value creation and makes it possible to develop DDSs that take both business and consumer strategy into account [588].

The design and development of DDSs pose new challenges for both businesses and consumers, particularly about the necessary data literacy in PD for both designers and consumers [448]. There is a '[...] lack of practical tools and methods that take into account the complexity of the service and the inherent dynamics of the design process' [114] and the creation of a shared understanding and a suitable language for mutual exchange and learning [448].

However, companies still need to take advantage of the opportunities that arise from user-centered development and, therefore, tend to exhibit an unacceptable level of user-friendliness, which means that products and services do not last long on the market [667]. Accordingly, Norman (2002) [452] points to the need for a philosophy that considers the needs and preferences of consumers and focuses on usefulness and comprehensibility in the development of products and services, which is referred to as user-centered design. In principle, UCD can be applied to any products and services, although in practice, it is mainly the development and design of information objects and systems that

are considered [254, 384]. Since information systems define the interaction between humans and technology to fulfill a task, various research domains are concerned with improving this interaction. In particular, human-computer interaction (HCI) has established itself as a field of research [520].

According to Lowdermilk (2013) [384], ”Usability practices could be implemented in everything from a toaster to a door knob, and even the packaging of both.” This view goes hand in hand with user-centered design [452] and defines the degree to which a consumer can use a system, product, or service to achieve their goals as effectively, efficiently, and satisfactorily as possible in a specific context [60]. Agarwal and Venkatesh (2002) [12] emphasize that usability is associated with positive experiences and outcomes that promote the use of a system due to a positive attitude towards it.

The Association for Computing Machinery describes HCI as ”a discipline concerned with the design, evaluation and implementation of interactive computing systems for human use and with the study of major phenomena surrounding them” [267] and includes UCD, which serves as a software design method for developers and designers to incorporate the needs of users in the development of applications to promote their user experience (UX) [384]. According to Ritter et al. (2014) [520], UX encompasses ”emotions, beliefs, preferences, perceptions, physical and psychological responses, behaviors, and accomplishments that occur before, during and after use” intending to help improve the user’s understanding of complex systems, products, and services, especially since in most cases no documentation is available [520].

For the development of DDSs, it is relevant to understand that traditional development and design processes are essentially characterized by the internal corporate vision. Accordingly, an internal architecture is defined first, and only then the user-relevant elements are created, whereby DDSs can be seen more as a compilation of components and functions [666]. Furthermore, this traditional approach is subject to linear, rigid, and inflexible development. In contrast, UCD postulates a focus on the user and pursues a flexible and iterative design process [520]. To this end, user feedback must serve as input for design at various points in the design and development process [666].

UCD is a paradigm that represents the interface between business and con-

sumer and is, therefore, of central relevance for the alignment between business and consumer [72]. This relevance is also evident from the studies in section III, in which various UCD methods were used to develop and evaluate DDSs in a user-centered manner.

Although UCD is widely acknowledged, there is a discrepancy between academic frameworks and the actual implementation in business practice. The academic frameworks for UCD see users as an integral part of all phases of the design process. In organizations, however, user feedback is often only collected at specific points in the design process, such as the user interface, and user involvement in strategic planning and the early stages of design does not take place [553, 523, 275]. It is more applied as a linear method that collects user feedback at the end of the design process instead of establishing a continuous and proactive integration in an iterative process. The relatively superficial application of UCD results from frequent practical limitations, such as budget constraints and the need to deliver quick solutions [553]. In addition, the necessary expertise for the holistic implementation of UCD is often lacking. It is crucial to promote communication and dialogue with users and to bridge the gap between theory and practice [92, 339].

The following section 2.4.2 introduces the basics of tools intended to support the design and development of DDSs. In particular, creating a shared understanding and adequate communication between all stakeholders involved in the design and development process with different data competencies [501] plays a central role.

Canvas models contribute to the simplification of complex problems by visualizing them and making their elements easier to understand [464]. The first canvas models have also been discussed in participatory design [585]. It can be seen that these canvas models as boundary objects [598], in addition to the possibility of joint design with consumers, also take into account the convivial exchange of knowledge and show how systems are developed and how their respective elements are related to each other [585, 330].

However, the scientific literature still lacks Canvas tools for designing with data and aligning businesses with consumers in DDS development. Therefore, this thesis focuses on canvas models to support Business-Data, and Business-

Consumer Alignment (see sections 14 and 17.3), as well as on a process to communicate with data (see section 15).

2.4.2 Tools to Support the Development and Communication of DDSs

As a critical resource, data is an essential component of modern innovation strategies for the development of DDBMs [252]. Business-Data Alignment represents how organizations align business value propositions and strategy with their data resources to develop new DDBMs and DDSs and connect them to the operational level. In addition to the challenges of gathering and integrating the necessary data to create and implement DDSs, their development and implementation also confront businesses with significant challenges, especially as this requires expert knowledge [67]. Decision-makers must be able to manage and evaluate opportunities arising from data and the knowledge in their business and create value with these resources [442, 472]. Decision-makers require data literacy and basic knowledge of data analysis next to an active data-driven corporate culture to enable data-driven value creation and business models [252]. As decision-makers lack data competencies [501], there is a practical need for tools that support creating value with data [333] since this is a 'heretofore unsolved and important business problem' [266].

As highlighted in section 2.1.1 and 2.1.2, several elements of a traditional business model play a role in the implementation of such models, including the representation of value creation, realization, and the cost structure [711]. Various instruments exist to support businesses in developing new business models, mainly based on simplifying the creation and communication of business models by helping to visualize and understand their elements. Therefore, the 'Business Model Canvas' developed by Alexander Osterwalder is a famous example [462]. A canvas model intends to summarize a complex problem as clearly and simply as possible and guide solution-finding.

Based on the current scientific literature, it is clear that conventional methods for developing and communicating (data-driven) business models are often inadequate, especially when it comes to establishing the link between data and value propositions [462, 561, 252, 334]. There is a lack of concepts and tools considering the specific requirements of DDBMs [333].

First canvas approaches for DDBMs like the Data Canvas [404] or the canvas for data-driven ideation workshops [330] focus on data and algorithmic value creation, approaches like the Data Insight Generator include the customer's value proposition more strongly [335]. Nevertheless, there is no holistic canvas approach that covers all relevant parts of business models and DDBMs [562, 711].

Specific problems of Business-Data Alignment that need to be considered in a holistic canvas approach are mainly data gathering [333], data integration of internal and external data sources in existing corporate infrastructures (see section 2.2) and (existing) operational processes [711, 404], data quality, confidence and integrity as a measure of trustworthiness, correctness and timely availability [99], and aligning data with value concepts [334].

Furthermore, as DDBMs require close collaboration between data scientists and business stakeholders, interdisciplinary communication of data requires data literacy and skills in data storytelling, as stories help to understand complex issues and arouse corresponding emotions to solve them [155]. Although more significant amounts of data and more complex analysis methods offer new possibilities for capturing facts, this also increases the complexity of analyzing, interpreting, understanding, and communicating [176, 272, 443].

Nevertheless, despite its increasing importance for business decisions [394], data storytelling is still a niche topic in the scientific literature. In particular, there is a lack of systematic research into success factors for effective communication with data, which helps companies communicate DDSs and DDBMs to their stakeholders.

3 Study Outline & Methodology

3.1 Research Framework

To bridge the gap between theory and practice (see Fig. 2), the terms Business-Data, Consumer-Data, and Business-Consumer Alignment are derived and coined from theory, which are intended to serve as principles for the development and design of DDBMs and DDSs and represent abstract constructs. These are examined in an application-oriented manner in the case studies to develop tools based on the resulting findings intended to support the user-centered development of data-driven value-added services.

To support companies in Business-Data Alignment, it is necessary, on the one hand, to understand the challenges that companies face in their technical development and design as outlined in section 2.2. On the other hand, section 2.3 shows that the needs and perceptions of consumers within Consumer-Data Alignment must be taken into account to create added value on both sides [433].

Concerning the technical development of DDSs, the complexity of the data infrastructure due to mostly heterogeneous data sources and formats means that, in addition to the identification and collection of relevant data, it is necessary to integrate this data into the corporate information architecture (see section 2.2.1) [147, 121]. In addition, companies face the challenge of developing suitable algorithms for their specific use case and require appropriate expertise in decision-making positions to assess them (see section 2.2, especially 2.2.2 on the complexity of choosing a suitable model within the case study in the food domain) [501].

The design of DDSs must continue to be aligned with users' needs. Therefore, Business-Consumer Alignment provides the opportunity to enable personalized and contextualized experiences that also represent added value for consumers [538]. As consumer protection organizations claim, transparent information about data usage must be provided, and consumers must be given control over their data [50, 356, 623, 525, 261] to support Consumer-Data Alignment (see section 2.3). According to section 2.4.1, a user-centered de-

sign approach enables companies to develop DDSs harmonizing technical possibilities with users' needs. By involving users in the development process, for example, through interviews and focus groups, valuable insights can be gained into their requirements and perceptions [611]. In this way, companies can design DDSs that offer added value for the company and consumers.

Nevertheless, UCD approaches differ between theory and practice. While theory claims users to be an integral part of all phases, businesses often use user feedback only at specific points in the design process [553, 523, 275]. Furthermore, developing and designing DDSs requires specialized expertise and close collaboration between data analysts and business experts [153]. Therefore, tools are needed to support businesses and consumers when developing DDSs [501, 333].

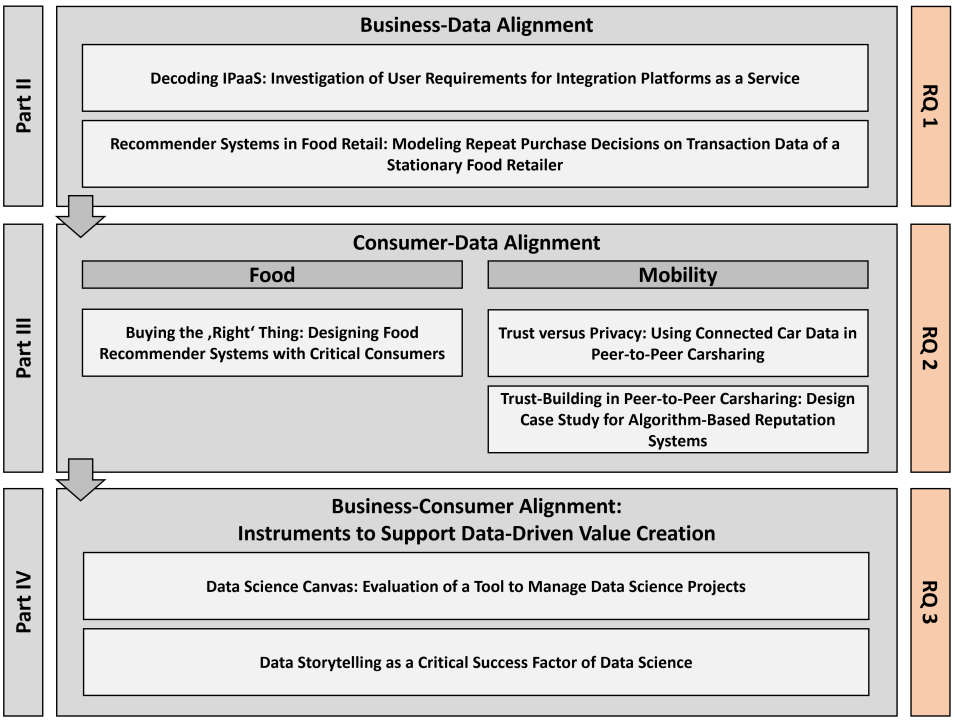


Figure 9: Organization of the Conducted Studies in the Structure of this Thesis

Part II deals with the answer to RQ1 and thus the perspective of companies

that develop and implement DDSs (see Figure 9). The aim is to analyze the factors that enable the development of these services and the challenges that arise from their technical implementation. Two methods are used for this purpose: expert interviews to reveal data sourcing and integration hurdles and the development and comparison of various algorithms suitable for the inherent problem in food retail of modeling repeat purchase recommendations as a specific case study for implementing a personalized DDS in the form of a recommender system.

The expert interviews in the case study *Decoding IPaaS: Investigation of User Requirements for Integration Platforms as a Service* in section 5 analyze the challenges that companies face when integrating data from different applications and sources and derive success factors for Integration Platforms as a Service, which are used by businesses as decision-making guidelines in the context of Business-Data Alignment. Therefore, ten Software as a Service provider experts from various business areas, such as Customer Relationship Management, Enterprise Resource Management, and Business Intelligence, were interviewed. The case study offers insights into practical hurdles and opportunities for companies to improve data integration.

In contrast, the case study *Recommender Systems in Food Retail: Modeling Repeat Purchase Decisions on Transaction Data of a Stationary Food Retailer* in section 6 examines the development and validation of algorithms for modeling repeat purchase decisions using real transaction data from a stationary food retailer. This study focuses on developing time-sensitive models to predict repeat purchases as an example of a DDS, emphasizing the practical application of transaction data to improve personalization in offline retail environments as data-enhanced business models (see section 2.1.2).

These two case studies collectively reveal insights into the principles of Business-Data Alignment (see Tab. 4): Since *Decoding IPaaS: Investigation of User Requirements for Integration Platforms as a Service* highlights the technical and organizational challenges in enabling seamless data integration across platforms, systems, and data sources, *Recommender Systems in Food Retail: Modeling Repeat Purchase Decisions on Transaction Data of a Stationary Food Retailer* demonstrates the practical implications of using integrated

transaction data for DDSs.

Case Study	Section 5: Decoding IPaaS	Section 6: Recommender Systems in Food Retail
Topic	Data Integration	DDS Development
Theory	Section 2.2, esp. 2.2.1	Section 2.1.2 & 2.2, esp. 2.2.2
Principles for BDA	Technical and organizational principles for data integration as an enabler of BDA	Findings on the development of a personalized DDS for a data-enhanced business model as a concrete operationalization of BDA
Methodology	Expert Interviews	DDS Development, Validation
Link to CDA	-	Uncovering the problems of a non-user-centered development of DDS as motivation and transition to part III, esp. sect. 9
Methods (Link to BCA)	Provides the technical basis for BCA, uncovering relevant elements for consideration in tools to support DDS development, esp. for section 14	Connects personalized services with strategical goals, uncovering relevant elements for consideration in tools to support DDS development, esp. for section 14

Table 2: Contribution and Methodology of the Case Studies in Part II on Business-Data Alignment and the Bridging Between Theory and Practice through Principles (see Fig. 2)

In part III, the focus shifts to the users' perspective of DDSs and addresses RQ2. The aim is to examine consumers' needs and perceptions in food and mobility domains (see section 1.2) and derive user-centered design principles for DDSs.

For this purpose, the case studies in part III are based on the user-centered methodology of the design case study [691], which is suitable for deriving the perceptions and perspectives of consumers about DDSs and the resulting findings for the user-centered design of these services. Design case studies are characterized by their perspective and creative nature. Combining the analysis of actual usage contexts with developing and evaluating design solutions enables a close link between research and practice insights [691]. However, a long-term evaluation of the appropriation of the artifacts developed in this thesis still needs to be conducted. Furthermore, interviews and focus groups are conducted to identify initial user needs and requirements and thus gain a deeper understanding for the subsequent design case studies [428, 338].

The case study *Buying the 'Right' Thing: Designing Food Recommender Sys-*

tems with Critical Consumers in section 9 examines how a personalized food recommender system supports critical consumers in making the 'right' purchasing decisions in the context of Consumer-Data Alignment. As also revealed in section 6, recommender systems usually only represent the 'average' consumer and neglect other niche practices. In the food domain, this applies to critical consumers such as vegans. Therefore, a user-centered food recommender system focused on ethical and sustainable values was designed and evaluated. Based on 24 pre-study interviews with consumers about their current consumption practices, the recommender system was developed as part of a co-creation workshop and evaluated by ten consumers concerning practical requirements and appropriation.

The case study *Trust versus Privacy: Using Connected Car Data in Peer-to-Peer Carsharing* in section 10 sheds light on the alignment processes of data sharing and the perceived added value of consumers to create trust through telematics data in peer-to-peer carsharing (see section 1.2). Three focus groups and eight interviews were used to identify principles characterizing Consumer-Data Alignment. The findings of the study form the basis for the development of an algorithm-based reputation system in the subsequent case study *Trust-Building in Peer-to-Peer Carsharing: Design Case Study for Algorithm-Based Reputation Systems* in section 11. The understanding of consumers of algorithm-based reputation systems was explored in a preliminary study using 16 interviews. Based on this, a prototype was co-designed, which was finally evaluated by 12 consumers regarding the creation of trust in peer-to-peer carsharing.

The findings of the three studies from part III offer principles for Consumer-Data Alignment and Business-Consumer Alignment. While *Buying the 'Right' Thing: Designing Food Recommender Systems with Critical Consumers* emphasizes the user-centered design of DDSs to promote sustainable and ethical consumption practices, *Trust versus Privacy: Using Connected Car Data in Peer-to-Peer Carsharing* focuses on the negotiation processes between data disclosure and trust for the acceptance of DDSs. In addition, *Trust-Building in Peer-to-Peer Carsharing: Design Case Study for Algorithm-Based Reputation Systems* shows not only user-centered design principles for DDSs but also how the negotiation processes of consumers in data-driven value-added

services can be created. Overall, it becomes clear how data-driven innovations and user-centered approaches can be combined to develop DDSs that meet both the needs of consumers and the strategic goals of businesses.

Case Study	Section 9: Buying the 'Right' Thing	Section 10: Trust versus Privacy	Section 11: Trust-Building in P2P Carsharing
Topic	Food	Mobility	Mobility
Theory	Section 2.3, esp. 2.3.3	Section 2.3, esp. 2.3.4	
Principles for CDA	Provides insights into the user-centered design of DDSs in the food domain	Illustrates the negotiation processes of consumption practices and data strategy between added value and privacy of DDS	Provides insights into the user-centered design of DDSs in the mobility domain
Methodology	Design Case Study (w/o long-term appropriation)	Focus Groups & Interviews	Design Case Study (w/o long-term appropriation)
Methods (Link to BCA)	Shows the complexity and important elements in the user-centered development of value-added DDSs in the food sector, uncovering relevant elements for consideration in tools to support DDS development, esp. for section 14	Provides principles on how to support coordination processes between businesses and consumers, uncovering relevant elements for consideration in tools to support DDS development, esp. for section 14, and revealing the necessity of communication for section 15	Shows the complexity and important elements in the user-centered development of value-added DDSs in the mobility sector, uncovering relevant elements for consideration in tools to support DDS development, esp. for section 14

Table 3: Contribution and Methodology of the Case Studies in Part III on Consumer-Data Alignment and the Bridging Between Theory and Practice through Principles (see Fig. 2)

Part IV brings together the results from part II and III to provide a holistic understanding of Business-Consumer Alignment.

Building on the challenges of Business-Data Alignment and Consumer-Data Alignment, the study *Data Science Canvas: Evaluation of a Tool to Manage Data Science Projects* in section 14 develops an instrument that supports companies in the strategic and operational linking of data with business objectives. By taking into account the identification of data sources and their integration (see section 5) as well as the findings on the development of DDSs from section 6, it is ensured that technological and organizational challenges can be mapped. Furthermore, user-centered factors based on the findings of part III can also be mapped, especially in data collection.

The Design Science Research (DSR) approach was used to develop the Data Science Canvas, which, as a design-oriented research approach, aims to create innovative artifacts such as models, methods, or systems that are geared toward solving practical problems while at the same time generating scientific knowledge [266, 479]. In the beginning, the problem was identified that many existing canvas models that deal with data are rather technically focused and do not offer a holistic approach to the consideration of DDBMs, which motivates the development of a holistic canvas model. The development of the Data Science Canvas artifact was based on a literature review to uncover possible objectives of a solution and an expert workshop with eight participants from different relevant scientific disciplines. The Data Science Canvas was finally evaluated by ten managers and data experts using it in real projects.

The study *Data Storytelling as a Critical Success Factor of Data Science* in section 15 further emphasizes the importance of transparent communication of data in the context of Business-Consumer Alignment, which is crucial for taking consumer concerns such as data protection and trust (see part III) into account. By integrating both perspectives, Business-Consumer Alignment is mapped as a holistic concept for developing DDSs. To this end, the literature review method was used to derive a process for communicating DDSs and the associated use of data. A literature review makes it possible to systematically analyze the current state of research on a specific topic and derive implications for practice [676]. A literature review can help to identify existing approaches, models, and theories, derive success factors, and support transparent, comprehensible, and understandable communication of DDSs [314].

Case Study	Section 14: Data Science Canvas	Section 15: Data Storytelling
Topic	Support of DDBM and DDS Development	Support of Communication with Data
Theory	Section 2.4, esp. 2.4.2	Section 2.4, esp. 2.4.2
Link to BDA	Develops a tool to support the strategic and operational alignment of data with business objectives	Supports the transparent communication of data-driven insights within organizations
Link to CDA	Considers user-centered factors by integrating stakeholder needs and the specification of data sourcing	Strengthens trust and acceptance in DDS through comprehensible and transparent communication for consumers
Principles for BCA	Combines technical insights on data integration and the development of DDSs with organizational and strategic challenges to support BCA	Provides guidelines for communicating with data and thus enables a common understanding in the context of a user-centered development of DDSs
Methodology	Design Science Research	Literature Review
Methods	Provides a tool for the development of DDBMs and DDSs	Provides a process for target group-specific communication with data

Table 4: Contribution and Methodology of the Case Studies in Part IV on Business-Consumer Alignment and the Bridging Between Theory and Practice through Principles (see Fig. 2)

Once the structure of this thesis has been presented, the individual research activities will be defined in more detail in the following.

3.2 Research Activities

3.2.1 Expert Interviews

As a method of qualitative social research, expert interviews enable the collection of specific knowledge and experience from people with particular expertise in a specific area [76, 77, 174]. In contrast to other interview forms, the focus of expert interviews is not on the individual characteristics or attitudes of the interviewees but on their function as holders of specialist knowledge and their integration into organizational contexts [421].

Their position, qualification, or experience characterizes experts and have privileged access to information and decision-making processes in a specific field of action [225]. Their knowledge can take various forms, such as technical, process, or interpretative knowledge [78]. By interviewing experts, direct access to this knowledge can be gained, resulting in insights into complex

contexts and backgrounds that would be difficult to access using other methods [483].

In the context of this thesis, expert interviews in section 5 are conducted to analyze the challenges companies face when integrating data from different sources and applications. By interviewing experts from different company areas, insights can be gained into the technical, organizational, and strategic aspects of data integration, and success factors are derived [121, 147].

Data integration experts have detailed knowledge of data structures and systems in their organizations. They can provide information on their challenges when combining and using data from different sources [533]. They have knowledge of the prevailing corporate integration practices and processes and can assess which changes and adjustments are necessary to enable effective data integration [121]. In addition, they can provide information about their companies' strategic goals and priorities and assess the role that data integration plays in developing DDSs and value creation [252]. Furthermore, existing problems, potentials, and limitations of different approaches and technologies can be uncovered [532]. By comparing and synthesizing the statements, overarching patterns, and success factors can be identified, which are relevant for developing recommendations for action and best practices [421].

3.2.2 Development and Comparison of Algorithms

The comparison of algorithms is an empirical method used in applied computer science and data science [16]. This method aims to evaluate the different approaches and procedures for solving a specific problem in terms of performance, efficiency, and scalability [102].

In the context of developing DDSs such as recommender systems, the comparison of different algorithms plays an essential role [9, 575]. The selection and optimization of relevant algorithms determine the quality and effectiveness of the services and require a systematic approach and empirical testing [11].

This thesis implements and compares algorithms to analyze companies' challenges when developing a recommender system as a concrete example of a

DDS in section 6. Recommender systems are complex applications that can promise personalized value-added services based on analyzing large amounts of data and applying advanced AI algorithms [515]. The development of suitable recommender systems requires the selection and optimization of algorithms that can generate relevant recommendations and take the needs and preferences of users into account [9].

The comparison of algorithms also makes it possible to evaluate their scalability, robustness, and suitability for use in real application contexts based on real data [9]. This can reveal the limits of the algorithms, analyze their adaptability, and observe problems in their development [73]. The knowledge gained can be used to develop algorithms further and improve the recommender system's quality.

Therefore, section 6 uses a real data set of a German food retailer in the food domain to model repeat purchase decisions. The Aggregate Time Distribution (ATD) model aggregates the repeat purchase behavior per product based on the average purchase intervals of customers to predict the probability of a repeat purchase. In contrast, the Modified Poisson Gamma model uses a personalized analysis of purchase probabilities based on Bayes' theorem. Considering individual purchase patterns improves the model's precision and can also be used with sparse data.

For this purpose, the transaction data is first pre-processed to identify recurring purchase patterns and generate suitable features for both models. The models are then evaluated on training and test data sets [14, 311].

The models are compared by evaluating the ability to predict repeat purchases using the evaluation metrics precision@k, recall@k, and f1score@k, which are frequently used in recommender systems. The quality of the recommendations is analyzed utilizing a top-k-list, where k represents the number of recommendations considered and must be specified based on the use case [14].

Precision@k thus reflects the proportion of actually relevant elements in the top-k-recommendations as:

$$\text{Precision@}k = \frac{\text{Number of relevant elements in the top-}k}{k}$$

A high value of precision means that the majority of the recommended products are relevant to the consumer.

Recall@k also measures the percentage of relevant elements found by the recommender system under the top-*k*-recommendations as:

$$\text{Recall@}k = \frac{\text{Number of relevant elements in the top-}k}{\text{Number of all relevant elements}}$$

This indicates how complete the recommendations are, i.e. how many of the total relevant elements are contained in the top-*k*-list.

The f1score@k represents the harmonic mean of Precision@k and Recall@k as:

$$\text{f1score@k} = 2 \times \frac{\text{Precision@k} \times \text{Recall@k}}{\text{Precision@k} + \text{Recall@k}}$$

In this way, the f1score ensures that both precision and recall are considered in the evaluation.

3.2.3 Focus Groups & Semi-structured Interviews

Focus groups are moderated group discussions on a specific topic. They are used to gain deeper insights into the motivations and attitudes of the participants through an exchange within the group [331, 315].

As a form of qualitative interview, the interviewer uses a guideline with open questions as part of semi-structured interviews and, at the same time, leaves room for spontaneous follow-up questions and in-depth discussions [91]. The open character of the interviews allows individual perspectives, experiences, and concerns of the participants to be gathered [406].

Focus groups are used in this thesis in combination with semi-structured interviews in section 10 to investigate the interests of renters in sensor data of vehicles as well as the willingness of potential rentees in their sharing in the

field of peer-to-peer carsharing, as this data serves as a basis for the development of algorithm-based reputation systems [602].

The triangulation of both methods thus makes it possible to gain a holistic understanding of the user perspective and to shed light on the requirements for the algorithm-based reputation system from different perspectives, which, among other things, reflect their needs for the required information but also about potential privacy concerns [146]. The insights gained form a basis for the subsequent design case study, in which the system is developed and evaluated iteratively [691].

Integrating user needs and requirements into the design process aims to ensure that the developed reputation system meets the expectations and requirements of users and is accepted and used in practice [452]. The early involvement of users through interviews and focus groups thus contributes significantly to the success of the subsequent design case study for the user-centered development of the algorithm-based reputation system.

3.2.4 Design Case Study

Following a three-phase approach, the design case study methodology is applied to ensure a practice-oriented perspective. The design case study combines design research methods with qualitative and ethnographic user research methods, which aim to explore and understand practices before and after an intervention [691].

1. Empirical Pre-Study: The initial stage, known as the Empirical Pre-Study, is dedicated to illuminating the potential for crafting an intervention and gaining a comprehensive understanding of user practices. These practices encompass the observable actions of users and the underlying infrastructures, technologies, tools, meanings, and competencies that shape their behavior [186]. Examining user practices through this lens, the challenges that need to be addressed and the areas where support is required to facilitate the transformation from the current state to the desired state become evident. This perspective identifies the problems to be solved and the necessary steps to enable users to adapt their

practices successfully [690].

2. (Participatory) Technology Design: During this next stage, the artifact is developed by transforming the (tentative) requirements identified in the empirical pre-study into tangible features. Rather than solely relying on researchers or designers, this phase may also involve end-user participation through collaborative activities such as workshops or shared data analysis, ensuring a more inclusive design process [690].

3. Evaluation: The design case study cycle concludes with an iterative design artifact evaluation. This final phase aims to assess the product from a user-centric perspective within an authentic context. The primary focus is determining the artifact's feasibility and suitability for its intended purpose. Insights gained from this analysis can subsequently inform further requirements and guide design refinements, mainly because users can often only articulate specific requirements once they have a tangible prototype to interact with [690, 691].

In principle, the development and comparison of the recommender system in section 6 and the illumination of the willingness to provide data to enable algorithm-based reputation systems in section 10 can be understood as an empirical pre-study. At the same time, independent design case studies in food and mobility are carried out in the following papers, including another empirical pre-study.

In part III, two case studies are conducted that are oriented on the framework of a design case study, each in the consumer domains of food and mobility. In the food domain, a recommender system for critical consumers is developed in section 9, which is an extension of the previous study on recommender systems from section 6 based on user-centered design perspectives. In the mobility domain, the case study incorporates the knowledge gained into the user-centered development of an algorithm-based reputation system in section 11.

Nevertheless, the case studies in this thesis do not represent complete design case studies, as there is no long-term evaluation of appropriation.

3.2.5 Design Science Research

Design science research (DSR) aims to develop innovative artifacts such as models, methods, systems, or processes that solve practical problems and simultaneously provide scientific findings [266, 479]. As a design-oriented approach, DSR focuses on designing and evaluating novel solutions for relevant problems in practice, in contrast to behavior-oriented research, which primarily focuses on explaining and understanding phenomena [232].

The DSR approach characterizes an iterative procedure in which artifact development and evaluation occur in close collaboration with practitioners [566]. By involving experts and potential users, requirements for the artifacts can be specified, their usefulness and applicability continuously reviewed, and any necessary adjustments made [660]. This ensures that the solutions developed to the problems are scientifically sound, practically relevant, and usable.

DSR is used in this thesis to develop a tool in section 14 based on the findings from parts II and III to support businesses in the development and design of DDSs. The findings from Section II are intended to show which technical challenges companies need to overcome when integrating and analyzing data from different sources to enable effective DDSs [121, 147]. At the same time, the results from part III should make it clear that the needs and perceptions of users play a central role in the acceptance and success of these services and how they must be designed in a user-centered way [414].

This allows for creating an artifact that offers companies sound and practical support in developing DDSs. The tool can help overcome the technical and business challenges of data integration and analysis while focusing on the needs and expectations of users. In this way, companies can develop services that offer real added value for consumers [365]. At the same time, the companies themselves benefit from more efficient processes, a better basis for decision-making, and the possibility of a shared understanding of the DDSs to be developed [153].

3.2.6 Literature Review

A literature review is a systematic and critical analysis of the existing research literature on a specific topic [198]. It aims to analyze the current state of knowledge, identify research gaps, and derive new findings and recommendations for action [676]. Literature reviews follow a structured and reproducible procedure that includes selecting, evaluating, and synthesizing relevant literature [642].

Literature reviews are particularly suitable for the consolidation of existing theories, models and empirical findings on a topic and the subsequent development of new concepts and frameworks [46]. By systematically processing the literature, patterns, trends, and best practices can be identified, which serve as a basis for deriving practical implications [93].

This thesis uses the literature review methodology to derive a data storytelling process based on the findings of the previous sections and the current state of the art on data communication in section 15, which supports companies in communicating DDSs.

To successfully implement DDSs and utilize their potential, companies must communicate these services' functionality, benefits, and implications to various stakeholders understandably and convincingly [152]. Data storytelling can present complex correlations and findings from data clearly and narratively, thereby promoting understanding and acceptance among the target groups [319].

By analyzing case studies, empirical studies, and conceptual work, data storytelling's central elements and success factors can be worked out [319]. On this basis, a holistic data storytelling process is derived that offers companies a structured guide to communicating DDSs. This process should cover the various phases from data preparation to story development and presentation and provide specific techniques and tools for each phase [177]. By incorporating insights from practice, for example, by analyzing best practice examples, it can be ensured that the process is practical and feasible.

Part II

Uncovering Problems in Developing Data-Driven Services to Support Business-Data Alignment

4 Introduction

To support businesses in Business-Data Alignment and developing value-creating DDSs for consumers, an understanding of the challenges companies face when developing them must be gained [70]. This part, therefore, serves to answer RQ1 and sheds light on the problems businesses face when sourcing data in section 5 on the one hand and developing and generating value from DDSs [274] in section 6 on the other.

As pointed out in section 2.1.2, two socio-technical characteristics make data valuable: Interconnectivity and portability. While interconnectivity defines that data can be connected to other data to gain new insights, requiring advanced technical methods and human collaboration, portable data can be used in different application contexts, which requires a technical infrastructure enabling this usage and impacts working practices and organizational models on a social level [235].

Interconnectivity requires business information system architectures to link data from different applications and sources. Digitalization and the associated increase in cloud services mean that more and more data from external sources must be integrated to be used in DDSs (see section 2.2.1) [698]. Therefore, section 5 deals with the consideration of business practices of linking different data sources, which should be simplified by data integration platforms (Integration Platforms as a Service, IPaaS) [178]. Nevertheless, current IPaaS solutions are predominantly provided by U.S. vendors, which poses a challenge to businesses regarding trust, security, and sovereignty of data [337]. Furthermore, prior research mainly investigated technical or legal aspects of IPaaS, so there is a research gap regarding a nuanced understanding of the factors that influence IPaaS adoption. The empirical study based on expert interviews in section 5 aims to derive challenges and success factors that support platform designers' and companies' integration practices by providing guidelines and recommendations for action.

As the consumer informatics application field is increasingly penetrated by DDSs, recommender systems in the food domain offer the possibility of reducing information overload and supporting consumers in purchasing deci-

sions (see section 1.2). Value-added services like personalised recommender systems can be developed using an integrated database (see section 2.2.2) [626]. Therefore, in section 6, a recommender system for food retail is being developed that uses integrated transaction, product, and master data to offer consumers individual added value through a recommender system that suggests when they will next buy a product [62]. This is intended to investigate the portability of data, which, in addition to classic data analyses such as sales forecasts, pattern, and trend recognition, also enables the application context of personalized support as added value for consumers. Furthermore, this shows how such an innovation can transform a traditionally low-data BM into a data-enhanced BM (see Fig 5).

The insights gained are summarized in section 7 and analyzed to motivate the following part III on Consumer-Data Alignment.

5 Decoding IPaaS: Investigation of User Requirements for Integration Platforms as a Service

Abstract

Due to ongoing digitalization, more and more cloud services are finding their way into companies. In this context, data integration from the various software solutions, which are provided both on-premise (local use or licensing for local use of software) and as a service, is of great importance. In this regard, Integration Platform as a Service (IPaaS) models aim to support companies as well as software providers in the context of data integration by providing connectors to enable data flow between different applications and systems and other integration services. Since previous research has mostly focused on technical or legal aspects of IPaaS, this article focuses on deriving integration practices and design-related barriers and drivers regarding the adoption of IPaaS. Therefore, we conducted 10 interviews with experts from different software as a services vendors. Our results show that the main factors regarding the adoption of IPaaS are the standardization of data models, the usability and variety of connectors provided, and the issues regarding data privacy, security, and transparency.

5.1 Introduction

In the course of the ongoing digitalization, more and more cloud services find their way into organizations and companies. These services are usually used as software-as-a-service (SaaS) that address many different application domains (e.g. Customer Relationship Management or Enterprise Resource Planning) and promise possible financial advantages [248]. However, the constantly growing range of new SaaS solutions adopted by organizations requires a steady development of new Application Programming Interfaces (API) to exchange data between the different services and prevent data and application silos [370]. The software vendors of those organizations face the task to synchronize evolving multitenant SaaS architectures [668] with the

customer system and the other integrated applications [708]. Moreover, new costs arise from the integration as well as risks regarding security and legal compliance (e.g. privacy) [313].

In this context, Integration Platform as a Service (IPaaS) models have emerged, which are defined as "suite[s] of cloud services enabling development, execution and governance of integration flows connecting any combination of on-premises and cloud-based processes, services, applications and data within individual, or across multiple, organizations" [482]. As they bring SaaS providers and businesses together, IPaaS platforms can be described as a multi-sided market, which aims to reduce the integration effort for enterprises, focusing on the development, management and execution of integration processes [401]. Nevertheless, they are facing different challenges regarding their role as intermediaries like the critical mass problem [190]. Furthermore, as the majority of existing IPaaS platforms are provided by U.S. vendors, this poses a challenge to European SaaS providers and enterprise customers due to differences in trust, security, and sovereignty of data [337]. To promote transparency and set an open European cloud standard, initiatives such as Gaia-X have already been formed [115]. In the context of IPaaS platforms, for example, the Open Integration Hub could be mentioned as a European open-source data integration framework [85].[1]

Research Gap: While prior research mainly focused on technical or legal aspects of the IPaaS, we lack a nuanced understanding of the factors that influence adoption. Ebert et al. claim that "critical success factors for using IPaaS [...] have not been investigated" [178] to a large extent although understanding drivers and barriers to adoption from both a technical and a business perspective might contribute to a better design of those platforms. For this, it is also important to understand the current integration practices of software vendors and why they are perceived as beneficial from the vendors' perspective. Against this background, we aim to answer two research questions:

- RQ1: What are current integration practices and issues with existing IPaaS solutions of software vendors?

[1] https://www.openintegrationhub.org/?lang=en

- RQ2: What are perceived design-related barriers and drivers to the adoption of IPaaS?

Addressing this research gap, this article reports on a qualitative empirical study with 10 experts from different SaaS software vendors. We conducted semi-structured expert interviews with an average duration of 56 minutes. The focus was on current integration practices leading to drivers and barriers to the adoption of IPaaS as well as desired functions and opportunities regarding IPaaS platforms of market participants in the form of SaaS providers.

Since IPaaS platforms are multi-sided platforms [49, 401], the next chapter describes multi-sided platform markets and the role of IPaaS in the platform economy. The methodology of this work is described in chapter 3. Chapter 4 presents the results of our study in relation to our research questions. The results are discussed in section 5 and we draw conclusions in section 6.

5.2 IPaaS Platform Economy

Traditional linear business models create value through consumers buying and using their products or services. Platform-based business models, on the other hand, create value through other market participants using their platform. Here, the platform does not own and process the traditional factors of production, but provides an infrastructure that enables a transaction by using a matching algorithm to mediate the sides of the market. The strength of the platform economy lies in its ability to remove barriers to trade by using the increased exchange of information between different actors and the dissemination of data to its advantage. This creates a more open economic system with much greater user participation. To benefit from the effects between users and providers prevailing in networks, the platforms mostly rely on multi-sided market forms [135].

5.2.1 Characteristics of Multi-Sided Platforms

When two or more participants in a market come together to transact with each other, they are referred to as two-sided or multi-sided markets. The dif-

ferent market participants pursue different, complementary interests. They are networked with each other via the digital platform [135]. That is why digital platforms also belong to the category of network goods. The basic rule on digital platforms is: The more users there are on the platform, the greater the benefit for all participants involved. These so-called network effects work on digital platforms primarily in an indirect way [678]. For example, customers on a trading platform have no direct advantage if there are many other customers there. Rather, it is the different, complementary market sides that influence each other. Thus, the benefit from one side of the market depends on how many participants are active on the other side of the market. This type of interdependence is appropriately referred to as an indirect network effect [411]. A platform becomes more interesting the more providers there are in the marketplace. This is because customers have a greater choice and can expect advantages in terms of price and possibly also the quality of the products and services. At the same time, the value of the platform also increases for the providers if more potential buyers are on the marketplace. The intermediary, i.e., the platform operator, is the entity that brings the different market sides together, develops and coordinates the market and designs the platform in such a way that it is attractive for the respective market sides [1].

When developing a new platform, the platform operator faces the challenge of convincing not just one but two or more market sides of the benefits of the platform. The question arises for the operator as to which market side it should develop first and attract to its platform. To do this, the intermediary sometimes has to make great efforts to establish a flourishing exchange between the market sides. And even if the platform is initially accepted by users and develops positively, this does not automatically mean that success will last. Not only do users need to be convinced of the platform's benefits in the early period of market development, but they also need to recognize a benefit in the long term. So the intermediary has to deal with different coordination problems.To do so, he needs a deep understanding of his market participants [135].

5.2.2 IPaaS as an Intermediary

IPaaS platforms provide a cloud-based multi-tenant system that helps enterprises develop, manage and govern their integration flows between the wide range of required applications and data sources [178]. As an intermediary, the IPaaS provider thus connects the market side of the SaaS providers and the businesses.

The main task here is to design the data synchronization and mapping between the various applications. In particular, the connectivity of the various SaaS and on-prem systems and the data preparation, transformation and migration into the existing databases play a key role [708]. A central advantage of the IPaaS solution for companies and SaaS providers is a reduction in the number of connectors required (cf. Fig. 10), which is made possible by the platform structure. Businesses and SaaS applications no longer need to develop or provide many connectors to various other solutions and systems; instead, only one interface to all integrations is usually required. This can reduce errors and increase data security and integrity by supporting governance, management and monitoring [493].

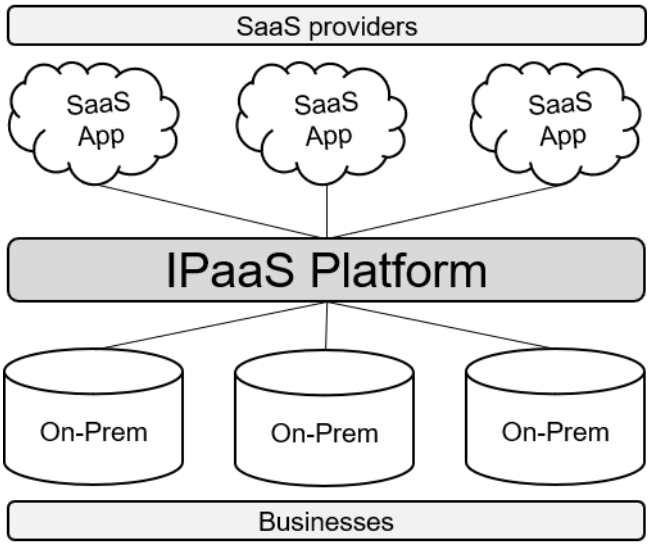


Figure 10: Relationship structure of an IPaaS market

In addition to the main focus, further functionalities become feasible in the form of orchestration of the various applications into business processes and

support for the development of individual connectors, data mappings, and the execution of automated processes [178]. Due to the integrated and cross-process databases, intelligent process automation (IPA) using machine learning is also gaining in importance here [117].

The motivation of SaaS providers and companies to participate in the IPaaS platform is mainly explained by indirect network effects in addition to the simplification of integration, whereby companies can benefit from an increasing number of SaaS applications through a more diverse selection. SaaS providers benefit indirectly on the one hand through increasing business users via additional sales channels. On the other hand, direct network effects can also be seen here through the possibility of integration between further applications and the resulting increases in efficiency on both sides [135, 178].

However, platform providers face the challenge of adapting to the fast-moving cloud market and evolving in line with user needs. It is therefore of central relevance to know the requirements, barriers and wishes of their users [568].

5.3 Methodology

From a user-centered design perspective, it is of central relevance to implement user requirements in the development of artifacts [400, 377], also to promote their adoption [126]. To understand software vendors' perspectives on requirements for IPaaS platform adoption, we conducted 10 semi-structured expert interviews with SaaS providers from Germany (see Tab. 5) within an interpretive research stance [54, 138], which has already been applied in other studies dealing with the derivation of user requirements [396, 634]. Our sample was selected based on contacts to software vendors from previous research projects. The focus of the selection was on an inclusion of differently sized providers of various SaaS applications to gather a broad range of perspectives and possible customer and application scenarios.

The interviews lasted 56 minutes on average. After obtaining participants' consent to collect, record, and anonymously process their interview data, the interviews followed a semi-structured guideline that covered the following topics:

- Current Integration Practices and previous Experience with IPaaS to derive Drivers and Barriers to the Adoption of IPaaS
- Desired Functions and Requirements of IPaaS platforms
- Opportunities regarding IPaaS platforms

Table 5: Overview of the Interview Participants

	ID	Business Area	Job / Area
SaaS providers	S01	CRM	CEO
	S02	Scheduling	CEO
	S03	CMS	Product Manager
	S04	ERP	Software Engineer
	S05	ERP	Product Director
	S06	BI	CEO
	S07	BI	Integration Manager
	S08	Billing	CPO
	S09	DMS	CEO
	S10	Marketing	Managing Director
where:	CRM: Customer Relationship Management		
	CMS: Content-Management-System		
	ERP: Enterprise Resource Planning		
	BI: Business Intelligence		
	DMS: Document Management System		
	CEO: Chief Executive Officer		
	CPO: Chief Process Officer		

After transcribing the interviews, the analysis was done using the inductive approach of thematic analysis [86] with a focus on 1) current integration practices and issues with existing IPaaS solutions, 2) required functionalities and Services, and 3) possible opportunities for integration platforms to improve the actual situation. After coding the interviews, the authors developed the themes collaboratively to grant a common understanding of the material.

5.4 Results

Based on previous integration practices and problems with existing IPaaS platforms, the following section derives drivers and barriers regarding adoption as well as the functionalities and services required by SaaS providers. Additionally, ways to improve existing IPaaS approaches are discussed.

Table 6: Overview of Results

Issues with Existing IPaaS Solutions	Required Functions and Services	Opportunities
Standardization	Variety of Connectors	Open-Source / Community
Cost-Benefit Ratio	Customizable Connectors	Ecosystem / Cloud Initiatives
Reliability	Intelligent Mapping	Niche Services
Usability	Detailed Documentation	User Experience
Transparency	Test Options	Intelligent Process Automation
Data Protection and Security	Data Hub	Process Optimization
Good Support	Data Enrichment / Open Data	
	Reporting (traceable data transfer)	

5.4.1 Current Integration Practices and Issues with Existing IPaaS Solutions

Especially among the smaller Saas providers interviewed, integration tools are rarely used. The required connectors are generally still implemented by the companies themselves. If IPaaS solutions are used, the large providers are mostly chosen. However, experiences with these platforms reveal a number of problems.

Although many connectors are offered for individual needs, this occasionally results in high costs. The **cost-benefit ratio** is not given in some cases, especially for small and medium-sized enterprises with specific solutions, since in most cases such an excessive offering is not needed and specific connectors are not available. Some participants therefore report a **lack of appropriateness** in the range of existing data integration tools and a **lack of added value** for their solution.

"[...] in the past we have looked at a few of these, let's call it connector suites, but honestly there's nothing there that's at least any good for our specific purpose. We've been operating on and off for a year and a half, but hardly [...] had any value add, apart from the fact that it's relatively expensive." –

[S01]

The lack of added value is often explained by rudimentary suites whose connectors do not meet the requirements of the solutions. For example, various authentication mechanisms demand individual logics in the background from the solutions. Furthermore, non-uniform data structures of the connectors and the resulting need for adaptation in the form of data preparation and transformation lead to additional effort.

"[...], these solutions are too rudimentary and actually don't know where the problem lies with connectors. Most connectors have different authentication mechanisms, different data structures, and with many APIs a call is not enough, instead many calculations, logics have to take place in the background of our solution, e.g. generate synthetic tables, so that it makes sense for the customers at all. [...], that's a nagging ongoing issue for us." – [S06]

For most respondents, therefore, the issue of **standardization** is the essential factor for adopting an IPaaS solution.

"[...] the topic of standardization. If one already makes the whole effort to build a connector. Then there should also be a standard for the individual data modules. – [S04]

They are currently missing suitable definitions for data models and data exchange. Merging the data from the connectors therefore confronts the solution providers interviewed with challenges. Although the data can be analyzed in isolation for each connector, it is extremely difficult to analyze and process cross-system data. This is due to different granularities within the respective data and different data formats.

"[...] if I want to connect data from Facebook, Twitter and Google Analytics within one chart, it sounds easy from a customer perspective. But from an IT perspective, it's complex because the different sources have different granularities. [...], here I get it on an hourly basis, there I get it on a minute-by-minute basis." – [S01]

"We've been looking for years for a format for explicating documents, including our meta-information, in a meaningful way." – [S09]

In general, the issue of **data protection and security** is also highly relevant

for the respondents in the context of existing data integration solutions and the connectors provided.

"If an American [connector] is involved, that can also be a criterion for exclusion [for one of our customers]. [...] and then, of course, you simply have to weight up how great the benefit is for what we want to do [...]. So data protection plays an important role in the consideration, so we developed our connectors ourselves." – [S02]

The **probability of failure** plays a significant role in the considerations of solution providers for migration to a data integration ecosystem. Some providers are also hesitant about **support** expectations in the event of an outage and are weighing up between developing their own solutions or connecting to a platform.

"So if that's an integrated part of your own system, you're less likely to have a failure. So that was our consideration. [...] And good support, that's also part of it." – [S03]

In terms of usability, a **quick and simple on-boarding** is of essential relevance. This also ensures continuous use of such a platform.

"I think the key is to get a moment of success within the first few minutes. After 5 minutes, I have my first functioning small solution. This moment of success has a lot of added value and ensures that people continue to use the platform." – [S01]

"From my point of view, the use and administration should be intuitive. If you do it really well, an administrator should actually be able to do it himself as a customer." – [S04]

5.4.2 Required Functionality and Services

With regard to expectations for the functions and services of an IPaaS platform, the respondents agree that it must predominantly provide a **wide variety of connectors**. If connectors are not available, they should be available on request within a very short time.

"[Such a platform] should offer 1000+ connectors with the ability to add new

ones super fast. That's also critical. [...] a customer comes, he would like to have a specific connector, so that you can then offer it directly." – [S09]

To address the issue of a lack of standardization, **customizable templates** for the APIs are discussed. Although many different connector templates are already available for the larger providers, the smaller providers in particular do not yet offer enough of these.

"[...], templates for API interfaces also play an important role for us. We would like to adapt these to our individual needs. There are already many different templates for Google, etc., but the smaller integration providers in particular do not (yet) offer this on a large scale because this results in higher costs." – [S07]

In addition, respondents would like to see further customization of the connectors to best fit their system.

"[...], there has to be an option like: ok, you want to connect to Facebook, you can update the data synchronization in one, six or twenty-four hour intervals, so I can trigger that myself." – [S06]

If the provided data schema does not meet the requirements of the SaaS provider, an **intelligent mapping** function is desired, which merges the (standard) data model with the new data fields.

"[...] or some kind of intelligent algorithm or end-user interface where I say, this is my data schema and whatever you get back, bring that into this data schema and store it in the following database." – [S06]

Before connecting to the platform, the providers would also like to see **detailed documentation** of the interfaces and a **test option** for various requests to evaluate how the data schema is designed and whether it fits their own solution.

"[...] so one wish is that you get a detailed documentation for the interface and maybe just a test possibility. That can make things just very simple. So such a negative experience with different providers. We would also just be able to test different requests, [...] whether that suits us." – [S02]

In addition, the platform should also provide whitelabel support in the event

of a SaaS app extension through the connectors and services provided.

"[...] we would find it exciting to use such services without a front end to the platform. The frontend would have to be left to us, because that's where our USP is." – [S08]

Some participants also mentioned various added values that such a platform could provide based on the integrated data situation. They described a central **data hub** that temporarily stores the integrated data to process it for machine learning applications and enrich it with additional data. In this regard, Intelligent Process Automation was also discussed as an opportunity.

"So the big advantage would be that you have an API that you can address and could then use many data sources, for example also open data. [...] I could definitely imagine that something like this would be extremely interesting for many people regarding data enrichment." – [S07]

"Keyword intelligent processes. Data is collected everywhere in a data lake, right down to the last [process] step, and through intelligent analyses I can perhaps see automation and optimization potential at the end. That I can also say, I don't know, an alarm system that predicts that my stock will soon be empty and then directly triggers consequences. But that is still up in the air." – [S06]

However, the participants were not unanimous on this. Some spoke of added value, others were critical with regard to data protection. One participant sees a compromise in the possibility of offering such a data hub as an option. In this case, it must be ensured with the highest priority that others are excluded from access.

"I think this [data hub] lacks acceptance due to data protection, at least in German-speaking countries." – [S03]

"I wouldn't want that as a customer. Except when I can simply decide for myself: The connector is active between app 1 and app 2 and I can say where the data goes or with which systems I then want to work and it is not then automatically assumed. One fear is certainly that anyone could have access to it." – [S01]

With regard to data protection, SaaS providers would also like to have a means

of **proof to their customers** in the context of the correct and error-free transfer of data between the various systems and applications.

"If you look at the issue of data protection, for example, the DMS app sent data to the ERP system. And the ERP system got into mischief with it. Some kind of data breach or something. And then, if necessary, you could also use the platform to say: Look, on that day, we transferred this data to the ERP system on the basis of the order processing contract. You could just say that we acted within the scope of the order processing, but from then on we were out of it." – [S09]

5.4.3 Opportunities

One opportunity to overcome the problem concerning the lack of standardization, transparency, data protection, and security is seen in an **open-source approach** with a broad **ecosystem** of connected companies and a large **community** dedicated to the common definition of open standard formats.

"As many large companies as possible would have to get together and say "we're doing this". If this is done as part of an open-source-based approach, for example, everyone could participate and benefit from it. If you could export a standard format from such a solution, that would be a perfect side effect, also for the participating community." – [S09]

"[...] and especially open-source. There are very few people who have problems trusting it in terms of data protection, because it is transparent in that sense." – [S08]

Cloud Initiatives should also be actively involved to promote such an open-source approach and create awareness of its relevance among SaaS providers and businesses alike.

"There are such cloud initiatives in Germany and I would be very happy if they address as many [SaaS] providers and companies as possible in the direction of adaptation, because the basic idea of an open-source platform would be really great." – [S09]

Moreover, **niche services** should be integrated, as the existing platforms do

not pay enough attention to them and usually do not offer them.

"I would just like such a platform to explicitly take care of relevant services, but also niche services. Because no one else offers them." – [S02]

Another point addressed is the creation of a **positive user experience**. According to the experts surveyed, some platforms lack a suitable usability and user experience concept. European platforms, in particular, have a lot of catching up to do compared to the competitors.

"The question is, for example, what does Zapier have that Flowground does not? Zapier is an American company with a high level of investment, while Flowground has a German team. And American companies have understood that they have to put user experience first. This is not done in many German and European teams due to higher costs. The urgent recommendation is therefore to bring an employee into the team who does nothing but user experience." – [S05]

Regarding a central data hub, the issue of **Intelligent Process Automation & Optimization** became also apparent as this could support small and medium-sized companies within the digital transformation. The experts agree that this is an important topic for the future, which could sustainably increase the use and acceptance of the platforms by companies.

5.5 Discussion

Reflecting on our research and the presented research questions, we aim to discuss the empirical results towards design implications for IPaaS that more broadly consider the business requirements. Thereby, the focus is on the main factors identified and those specific to IPaaS, against the background of identified opportunities. Therefore, generic factors such as support are less considered in this discussion.

5.5.1 Main Factors for Adopting IPaaS

5.5.1.1 Standardization In order to promote the adoption of IPaaS platforms, the standardization of data models is required in particular, according

to the interviews. This is also confirmed by [415] to reduce complexity of point-to-point integration.

To be able to offer a European alternative in the long term, a joint initiative is therefore needed to address the problem of standardizing data models and creating an adequate integration platform. Nevertheless, the experts surveyed assume that Europe can only act as an imitator here and that this will ensure a lack of economic interest with regard to the resistance to migration to the cloud, which is particularly prevalent among SMEs. This also slows down the consideration and further development of trends in the direction of intelligent process automation.

5.5.1.2 Usability and Variety of Connectors SaaS providers and their customers expect a user-friendly and simply designed platform that provides a wide variety of connectors and constantly updates this offering. Furthermore, the expectation of fast and demand-oriented integration of specific connectors presents an IPaaS platform with the challenge of having to respond to individual customer requests [568]. For SaaS providers, however, this represents a critical point in order to be able to profit from the direct network effects on the platform as part of their own solution.

5.5.1.3 Data Protection, Security and Transparency In Europe, in particular, the debate about data protection and data security is also being given a strong emphasis, which has so far kept many SaaS providers from adopting the established and predominantly American providers. Therefore, SaaS providers are demanding a European data integration solution that meets the providers' and their customers' requirements. However, the existing European approaches predominantly exhibit problems in usability and user experience compared to their American competitors. This is, according to our participants, caused by the too low willingness to invest in this direction.

5.5.2 Towards an Open-Source-Based Approach

If there is no common strategy at the economic level with regard to a general standard, an open-source IPaaS solution could be a possibility to ensure a European standard through a common goal and an ecosystem of software providers, stakeholders, and a broad community, and to carry this standard into the SME sector. Through the open system, which can be seen by all, the demands for transparency and neutrality towards the existing providers could be taken into account.

Open source solutions can be ”based on around solutions that could be preconfigured solutions, with a focus on implementation methodologies and diverse hosting options” [230].

The suitability of open-source solutions for standardization has already proven to be a successful tool in several other domains. For example, it proved to be successful in the context of the definition of Decentralized Identity (DID) for blockchain solutions [36].

However, the open-source approach does not offer an adequate solution for all reported issues. Due to the non-monetary structure of the open-source approach, a trade-off will have to take place between transparency and neutrality and the desired services. This applies in particular to the continuous development and individual provision of connectors. Therefore, in the sense of a multi-sided market [1], consideration should be given to the integration of an additional market participant in the form of connector developers. Those, in return for monetary compensation, can keep the connectors up to date and develop individual connectors on request. This could provide additional indirect network effects on the platform and increase its attractiveness and acceptance.

5.6 Conclusion

In this paper, we have dealt with the current integration practices of software vendors and the perceived design-related barriers as well as drivers for the adoption of IPaaS.

In terms of practical implications, our results show that there is a significant

need to create standardization of connectors and data models. According to our expert interviews, this could only be realized by a joint initiative of IPaaS and SaaS providers. In this context, the barriers of medium-sized companies in particular to migration to the cloud must also be overcome. To this end, the experts cite an open source approach as a possible solution, which, however, encounters problems in the field of tension between the expected individual and timely response to user requirements and the non-monetary, transparent structure. Therefore, such possible open source approaches as well as existing IPaaS platforms need to think about expanding to other market sides that support on-demand and customized connector provisioning. In addition, European IPaaS providers in particular should focus more on the usability of their solution in order to offer added value compared to SaaS providers and to be able to compete.

Our results are limited by the fact that the interview participants are exclusively German SaaS providers and not all industries in the SaaS sector are covered. Furthermore, further research should address the quantitative verification of the success factors as well as investigate the other market sides.

ACKNOWLEDGEMENTS

This research was funded by Federal Ministry for Economic Affairs and Energy; grant number 01MT20003D. The authors declare no conflict of interest. The authors would like to thank the project and Veerle Türling, Rosa Maren Steinprinz and Anna Zens in particular for their support. Furthermore, we also would like to thank the reviewers whose comments and suggestions helped to improve this work.

6 Recommender Systems in Food Retail: Modeling Repeat Purchase Decisions on Transaction Data of a Stationary Food Retailer

Abstract

In growing online retailing, recommendation systems have become established that derive recommendations from customers' purchase histories. Recommending suitable food products can represent a lucrative added value for food retailers, but at the same time challenges them to make good predictions for repeated food purchases. Repeat purchase recommendations have been little explored in the literature. These predict when a product will be purchased again by a customer. This is especially important for food recommendations since it is not the frequency of the same item in the shopping basket that is relevant for determining repeat purchase intervals, but rather their difference over time. In addition to critically reflecting classical recommendation systems on the underlying repeat purchase context, this paper derives two models for online product recommendations from the literature, validated, and discussed for the food context using real transaction data of a German stationary food retailer.

6.1 Introduction

In times of digital transformation, enormous importance is attached to data. This manifests itself not least in current trend topics in literature such as Big Data and Data Science [381]. Food retailers have also long since discovered such opportunities for themselves: A well-known example is the insight into the connection between the purchase of beer and diapers on the weekend. The information obtained about consumer behavior has been used to optimize advertising and pricing mechanisms [205]. However, this requires a broad database [120], which has led to asymmetric business models such as "Payback", which collect user data through discounts or loyalty programs and make it available to cooperation partners in anonymized form [273, 605].

But food retailing is not only experiencing change from a digital perspective: consumers' lifestyles are also currently transforming strongly. Health, ecological, ethical, social and culinary issues are gaining importance. Nutrition and eating habits should no longer merely satisfy hunger, but be an expression of the consumer's individual personality. A growing health consciousness among consumers is also having an impact on the demand for fresh and healthy foods. In the course of this, the regionality and origin of products in terms of quality, environmental awareness and ethical aspects are coming into the focus of buyers in order to eat healthy and climate-conscious [283, 352, 605]. This leads to growing demands on food and an increasing need for product variety in food retail [283].

Especially in online retail, which is predestined for the collection of user data [291], transaction, behavioral, and rating data is used to ensure a personalized experience for customers by providing them with relevant content through recommendation systems [619]. In online food retailing, Amazon Prime Pantry, for example, also relies on the use of recommendation systems to design customer-centric marketing activities [170].

However, recommendation systems are not yet omnipresent in food and online retail. While click pattern and user preference analyses are still relatively easy to integrate [490, 693], the more complex modeling of customers' repeat purchase behavior is particularly important in the food sector. This is due to the fact that purchase decisions there are often habitualized and therefore the question is not what to buy but when to buy it [296, 182]. While established studies and models in the context of marketing research mostly deal with the consideration of specific products or brands [192, 431], there are, however, first approaches that discuss recommendation systems for repeat purchase behavior using online tracked data [62, 165].

This paper critically derives the problems of existing recommender systems for food retail, and builds on them to validate current models for repeat purchase recommendations for this domain. However, in contrast to prior research, our work examines real-world transaction data from stationary supermarket terminals of a large German food retailer, as purchases in food retail mainly take place offline [489]. This represents a major difference to online

tracked data, which can usually be collected in a more structured and traceable form [291]. The paper concludes with a critical discussion of the results and identification of possible improvements to make the models more suitable for both online and offline food retailing.

6.2 Recommender Systems

6.2.1 General

Recommender systems are designed to support users in their (future) decisions based on their previous usage history and that of other users. In principle, they can be differentiated into non-individual, collaborative, content-based, knowledge-based, demographic and hybrid filters [13, 73]. While non-individual recommendations are the same for all users and thus lack personalization of the respective products (e.g., the most clicked products) [73], collaborative filtering (CF) examines the preferences of different users based on their consumption and usage patterns to identify similar individuals or items (e.g., movies on Netflix). Recommendations are made either based on the similarity of two items in terms of their user ratings (”item-based”) or on their ratings of similar users (”user-based”) [549, 707, 376]. In contrast to collaborative approaches, content-based filters generate their recommendations based on the characteristics and content of the items already consumed. For example, items similar to a user’s previous ones (e.g., articles about science and technology) are recommended [656, 73, 425]. Knowledge-based filters are based on satisfying customer needs through specifically defined product features. Here, explicit rules are used to generate recommendations (e.g., specifications such as size, min. and max. price, or zip code when buying a house) [13]. Demographic filtering makes recommendations based on a socio-demographic profile (e.g., age, gender, nationality) of a user [636]. Hybrid approaches represent a combination of different filters. These are often used to circumvent problems of individual methods as well as to increase the accuracy and efficiency of the filtering [13, 636].

In recent years and mainly driven by online retailing, collaborative and content-based recommendation methods have become established [89, 376]. How-

ever, these systems have fundamental problems that are problematic for recommending food products for repeated purchase.

6.2.2 Issues for Food Retail

6.2.2.1 Data Distribution A fundamental problem of recommender systems is the cold start problem, which leads to inaccurate preference capture for new users or products due to a sparsely populated customer-product matrix [636]. This is particularly problematic in food retailing given that the ”long tail” of items (e.g., niche items) has only sporadic ratings and therefore will be difficult to predict [134]. This is especially affecting the collaborative approaches, as they are based on historical customer preferences. Since user-based collaborative filtering is built on comparing item scores from different users, many neighborhood scores are needed for a specific item [95]. Content-based methods are not affected as much, since they determine similarities based on item descriptions and thus recommend products for which there are no reviews yet [636]. Due to dimension reduction, the matrix factorization (MF) approach can lead to better results [169, 336]. In the context of probabilistic model-based methods, such as the Bayes classifier, missing values are ignored in probability estimation [285]. However, in content-based systems, the so-called portfolio problem occurs, which ensures that only items are recommended due to overspecialization, which are very similar to already recommended products [637].

For the transaction data in food retail, this means that a recommendation system would need to have many purchases of the same product in order to form a neighborhood of similar customers. This is especially relevant regarding the cold start problem, so that new customers can be quickly assigned to a suitable customer group to be able to make appropriate recommendations. However, when it comes to sparse purchasing data with only a few comparable purchases, this leads to problems in generating neighborhoods and the model does not work accurately.

6.2.2.2 Scalability Another problem arises from the resource intensity of the algorithms for computing the optimal neighborhood. The demand for time

and memory increases linearly with the number of users and scores [707, 95]. For a food retailer with a broad product and customer base that wants to make recommendations in a split second, such an algorithm would create time and cost pressures. This can be countered by means of subsampling and model-based methods. With subsampling, only a subset of users is selected at a time, which is intended to relieve the storage capacity. However, the computation of the neighborhood remains fixed [362]. In the context of model-based clustering methods, users are grouped into clusters based on common properties. The active user is now compared to a group of users rather than individuals, so that the closest neighborhood can be quickly identified. However, problems with missing data also arise here when the distance functions lead to non-intuitive and unstable clusters [340, 295, 376]. Bayesian classification also provides an advantage here, as it is a probabilistic model that represents the historical data and therefore can perform classifications without having to retrieve the entire customer-product matrix [13]. Furthermore, MF approaches also provide acceleration of recommendations by reducing dimensionality [550].

As many transactions are needed to calculate neighborhoods, this leads to a high computational effort and accordingly raises resource issues. Memory-based methods therefore seem rather unsuitable for this type of recommendation, which is why probabilistic (Bayesian) models in particular can offer added value here.

6.2.2.3 Inherent Meaning Collaborative filtering is based on the assumption that users have common preferences. The more homogeneous the preferences of different user groups, the more functional the model will be. Furthermore, collaborative filtering is particularly suitable for subjective characteristics (e.g., musical taste) that influence a selection decision. If, on the other hand, there are predominantly objective quality criteria (such as price changes), which do not have to be weighed against each other, other models should be used. Accordingly, a homogeneity of the items is also desirable. Thus, they are similar with regard to their objective criteria and differentiate only by subjective characteristics (e.g., music albums usually have a similar price, a similar length and similar sale channels) [95, 376, 707].

In the case of purchasing data, this homogeneity is usually not given, since it is precisely the objective characteristics that cause behavioral changes in customers and are therefore often addressed by marketing (e.g., weekly offers). Due to the resulting trade-off between subjective and objective criteria, collaborative filters can be seen as problematic for the generation of recommendations in food retailing.

6.2.2.4 Data Persistence The temporal validity and relevance of the data should also be assessed. From the requirements of data distribution, the problem arises that items that are only relevant for a short period of time (e.g. daily news) are rather less suitable for collaborative filtering, since in principle few ratings can be expected. Historical ratings are also less helpful when users' tastes change quickly (e.g., preferences for clothing items that have since gone out of style) [707, 376, 362, 95].

In the food sector, the temporal validity of the data could play a subordinate role, since the tastes of the customers are mostly habitualized or develop there [296].

6.2.2.5 Synonomy Many recommender systems have problems when similar and closely related products have different names (e.g. clothes and dresses). Collaborative filters are often not able to find a match between such items and therefore do not calculate their similarity. This problem is solved, for example, by the automatic term expansion [378, 21], as well as the singular value decomposition (SVD) in the context of MF (esp. Latent Semantic Indexing) [549].

Again, food transaction data cause a problem, because there are often many very similar names for an almost identical product. For example, different names for the same mineral water (”Still”, ”Sparkling”) could lead to model inaccuracies. Here, extended methods of term expansion or machine learning methods for word embeddings, i.e. the reconstruction of linguistic contexts based on words, should be used [227, 347].

6.3 Integration of the Time Factor

Most recommendation systems are based on static principles, as they only take into account information about whether a user buys a product or not. However, the time factor is also already integrated in some recommendation systems. In principle, two types of time-based data are distinguished in the literature: The product introduction time and the time of purchase (implicit) or evaluation (explicit) [474]. Tang, Winoto, and Chan integrate temporal characteristics of items to reduce relevant candidate sets in the context of movie recommendations (production year), improving the accuracy of recommendations [621]. Ding, Li, and Orlowska consider user rating time to optimize an item-based collaborative filtering system. Here, weights are calculated based on the rating times of different items [168]. Lee, Park, and Park developed a time-based collaborative filtering system using implicit data (transactions). It is based on a time-based pseudo-rating matrix that takes into account product introduction time and purchase time under the assumption that a user's current preferences are disproportionately influenced by more recent purchases and that more recent items exhibit higher user interest [363].

Lathia, Hailes, and Capra (2009) focus on the effect of weekly retraining within a CF algorithm as a time-dependent predictive model. As an adaptive temporal CF method, it adjusts the neighborhood size of a k-Nearest Neighbor approach based on performance measured up to the current time [342]. Koren sees an inevitable need to incorporate temporal changes into a recommender system. He defines an MF model that analyzes temporal change behavior over the entire data history and validity [326]. Cho, Cho, and Kim consider customer purchase sequences over time to optimize recommendation quality. They use an extended customer-product matrix, which shows the purchase times in addition to the products. Furthermore, they cluster the transactions into homogeneous subclusters using the self-organizing map (SOM) technique. A change of the cluster membership by each individual transaction of a customer defines its purchase sequence. The time-dependent change in a customer's cluster membership also enables prediction of a customer's future purchases. A recommendation is generated by the system following the identification of the most similar purchase sequences compared to the current

customer, generating a set of products that the active customer has is most likely to purchase based on the N most frequently purchased products in the cluster [127].

The ”Eigentaste” algorithm of Nathanson et al. analyzes the time-related changes in customer preferences when selecting a product to recommend based on the last evaluation [438]. Chu and Park describe a machine learning algorithm that can improve the recommendations of new products by continuously updating time-based features (e.g., popularity, freshness) in relevant content profiles personalized [129]. The time factor is also being integrated into initial Deep Learning approaches for recommender systems. Most companies, which do not have access to long usage histories, have so far had to resort to item-to-item recommender systems. Hidasi et al. use recurrent neural Networks (RNN) to address the problem of only short session-based datasets in a memory-based approach [269].

6.4 Repeat Purchase Recommendations

The aspect of repeat purchases addressed in this paper is dealt with only sparsely in the literature on recommender systems. Repeat purchases describe any situation in which a customer buys more than one unit of a product [62]. A key publication on the topic of repeat purchase decisions for brands is Ehrenberg’s Repeat Buying Theory. Ehrenberg describes that most aspects of brand buying behavior can be explained in terms of just two variables. These are the market penetration and the average purchase frequency of a product, whereby even these two variables correlate with each other. Furthermore, a product purchase decision essentially depends on the timing of the purchase of a specific product class as well as the brand choice. Accordingly, almost all influences can be adequately explained if purchase frequency processes are specified per brand (see Fig. 11) [182, 579].

Ehrenberg’s theory looks at the purchase histories of individual users from consumer panels, i.e. the chronological sequence of purchases in all of a customer’s shopping baskets over a specific period of time and a specific point of sale, to study the purchase frequencies of a particular brand. It turned out

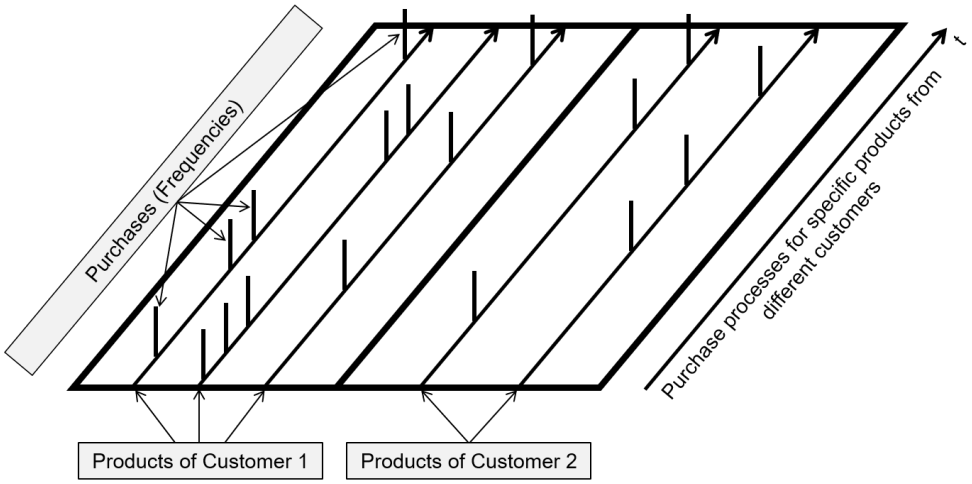


Figure 11: Purchase Events as Independent Stochastic Processes according to [182].

that the analysis should focus in particular on purchase occasions, i.e. the frequency of purchase of one or more items of a product at a specific time in a store, and not on quantity or price. Repeat purchases can therefore be described for a specific item by its market penetration and purchase frequency, where penetration represents the proportion of people who buy a particular product in the first place and purchase frequency represents the average number of these customers who buy at least one product in the period under consideration. The average purchase frequency here represents the basic measure of repeat purchases [182].

Ehrenberg distinguishes between three types of repeat purchases: 1) A customer may buy an item in more than one purchase in a given period. Different customers can be characterized by the number of their respective repeat purchases of the product. 2) A customer can buy an item in more than one period. 3) A customer can buy several units in the same purchase. For the problem definition of this paper, the first point in particular comes into consideration. The frequency distribution of (repeat) purchases and thus the number of consumers who have made 0 or 1 or, in the repeated case, 2, 3, 4, etc., purchases, can be described by the Negative Binomial Distribution (NBD) or the Logarithmic Series Distribution (LSD) [62, 221]. Most products in everyday life

have a certain time interval between the purchase of one product and the next (e.g. weekly purchase of mineral water). However, certain branded products in particular are rarely bought again, even over a longer period of time. Many customers of a brand buy a product only once, some twice, etc. Due to a thus relatively small share of buyers of a branded product in the population, the highest frequency is observed among non-buyers [182]. This results in a very skewed frequency distribution [199], which can theoretically be described by a mathematical function. The underlying function to fit the above problem is the NBD, with the LSD being a simplifying approximation to the NBD [182].

Bhagat et al. (2018) further discuss different approaches for repeat purchase decisions. In contrast to the models above, they aim to generate individual recommendations for repeat purchases. In doing so, they shed light on the temporal dependencies between repeat purchases of products, as these depend on when a product was last purchased and how quickly the customer runs out of it [62].

Therefore, they assume that a customer who has bought a product frequently in the past will buy it again (Repeat Customer Probability Model, RCP). The recommendations are ranked in descending order according to the number of repeat purchases. However, the problem arises that frequently purchased items are not relevant for a certain recommendation period, but are still considered due to their high number of repeat purchases. Even the integration of a certain time decay, which models repeated purchases based on a specific half-life, would be problematic due to the assignment of the highest rank directly after a repeat purchase, as this would increase the rank in the recommendation list. It can be assumed that the attractiveness of a product for repeated purchase immediately following the purchase of this product is rather low. In order to include the temporal relevance of products, it is also discussed that the purchase of items represents a periodic phenomenon and is therefore subject to a certain time interval (Aggregate Time Distribution model, ATD). Other approaches based on this idea are the Poisson-Gamma model (PG) and the Modified Poisson-Gamma model (MPG), which perform better due to the personalization of purchase rates by a Bayesian Hierarchical Model [62, 129, 158, 229].

6.5 Methodology

6.5.1 Data

The data set includes eleven features: CustomerId, ShoppingCartId, MarketId, Date, ShoppingCartValue, ItemId, ItemName, ItemQuantity, ItemPrice, CategoryId, and CategoryGroupLevel. Some products with negative item price and negative item quantity are included. They were mainly marked as NaN (Not a Number) and describe items such as deposits and empties, which are refunded to the customer. Negative item prices and quantities as well as NaNs are already removed from the data set before data preparation, as they would distort the data analysis due to their negative values. In addition, other items with the designation ”deposit” or ”delivery” existed, which were also deleted.

The reduced and cleaned dataset includes 49,920,981 transaction records. The transactions took place between December 28, 2018, and April 29, 2019. Overall, while the data set has a wide range of analysis capabilities with approximately 50 million transactions and 200 thousand products, these already reveal likely weaknesses with respect to modeling repeat purchase decisions. For example, the average number of purchases per customer across all products is 5.1 with a standard deviation of 5.6. This indicates a high number of customers who also purchase fewer products in the complete period under consideration. The average number of products in a shopping cart also confirms this impression with only 1.27 items. The maximum number of items in a shopping cart is only 22 for 200 thousand different products. Accordingly, although there are many individual purchases of specific products, there are no frequent repeat purchases by a specific customer. The number of purchases per customer is in the range of one to two purchases in the considered period. The number of one-time repeat buyers (loc) per product is less than five in 50% of cases. The number of repeat buyers (moc) is even less than three. These key figures result in the Repeat Customer Probability – i.e. the probability of a product being purchased by a repeat buyer.

Since this is a time-dependent classification problem, the data is divided on the basis of a reference date. The test data thus represent (actual) future trans-

actions of the customers, which are to be predicted by the model and used for validation within the framework of a confusion matrix with their specific key figures. Here, the due date was set at March 31, 2019, and the last month of the data set thus serves as test data. Accordingly, the training data set comprises 37,289,860 rows (74.7%), and the test data set in turn 12,631,121 rows (25.3%). In addition, it is assumed that a customer's purchase will only occur in the test period if that customer purchased a product in both training and test data [62].

Due to sparsely purchases at customer level (data distribution), two models are to be evaluated for the present data. These are the ATD model and the MPG model. While the ATD model addresses this weakness by considering the aggregated repeat purchases per product, the MPG model analyzes the personalized purchase rates of a product using the Bayes approach as a probabilistic model-based method.

6.5.2 Modeling

The problem of repeat purchase recommendations is described as estimating the probability of a repeat purchase as a function of time since his last purchase of the item under consideration, given the customer's previous purchase history. Accordingly, the associated purchase probability density P_{A_i} is to be estimated for a future time interval t_{k+1} assuming that a customer C_j has purchased an item A_i k times in the past with time intervals $t_1, t_2, ..., t_k$. Thus:

$$P_{A_i}(t_{k+1} \mid t_1, t_2, ..., t_k) \tag{6.1}$$

It is supposed that the customer's purchase times for different products are independent [62].

Furthermore, it is assumed that the above purchase probability density is composed of two components. Q_{A_i} represents the probability of a repeat purchase of a customer who buys a product for $(k+1)^t h$ times with k previous purchases. R_{A_i} defines the probability distribution of t_{k+1}, which depends on the repeated purchase of the item by the customer ($A_i = 1$).

$$P_{A_i}(t_{k+1} \mid t_1, ..., t_k) \approx R_{A_i}(t_{k+1} \mid t_1, ..., t_k) \cdot Q_{A_i} \tag{6.2}$$

Moreover, the time distribution $R_{A_i}(t_{k+1} \mid t_1, t_2, ..., t_k)$ is supposed to be asymptotic to $R_{A_i}(t \mid t_1, t_2, ..., t_k)$ [62, 158, 649].

$$\begin{aligned} &R_{A_i}(t_{k+1} \mid t_1, t_2, ..., t_k) \approx R_{A_i}(t \mid t_1, t_2, ..., t_k) \\ &where: \int_0^\infty R_{A_i}(t)dt = 1; \int_0^\infty P_{A_i}(t)dt \leq 1 \end{aligned} \tag{6.3}$$

The above mentioned RCP model serves as a basis for the following models in order to consider only products that are suitable for repeat purchase. It is defined by analyzing aggregate repeat purchase behavior as the ratio of the number of customers who have purchased a product A_i more than once (*moc*) to the number of customers who have purchased a product A_i at least once (*loc*). The derived repeat customer rate RCP_{A_i} approximates $Q(A_i)$ and without considering the time intervals between purchases also P_{A_i} according to the following formula [62, 191]:

$$\begin{aligned} &RCP_{A_i} = \frac{\text{moc}}{\text{loc}}, \\ &P_{A_i}(t_{k+1} = t \mid t_1, ..., t_k) \approx Q(A_i) \approx RCP_{A_i} \end{aligned} \tag{6.4}$$

Only products with an $RCP_{A_i} > r_{threshold}$ are taken into account further.

6.5.2.1 Aggregate Time Distribution Model If there are only a few repeat purchases at the customer level, but a large number of customers at the product level who have bought the product repeatedly, a model is suitable which analyzes the aggregated and time-based repeat purchase behavior across all repeat purchasers of a product. This models the determination of the probability distribution of the time intervals (t) of repeat purchase of a specific product across all repeat purchase customers. For this purpose, Baghat et al. examined various distributions in the context of determining mean time intervals for each customer in a sample of repeat-purchased items, and the log normal distribution achieved the best fit [268, 62].

$$R_{A_i}(t) = \frac{1}{\sqrt{2\pi}t\bar{\sigma}_i} \exp\left[-\frac{(\ln t - \bar{\mu}_i)^2}{2\bar{\sigma}_i^2}\right], t > 0 \tag{6.5}$$

Accordingly, the ATD model estimates the parameters of the log-normal distribution for each suitable repeat purchase product by fitting them to the different repeat purchase time intervals t of all repeat customers. Here, $Q(A_i)$ represents a fixed constant q for all products A_i of a given time t. Recommendations are made based on the descending order of probability density P_{A_i}(t) at a given time t using $P_{A_i}(t_{k+1} \mid t_1, ..., t_k) \approx R_{A_i}(t_{k+1} \mid t_1, ..., t_k) \cdot Q(A_i)$ [62, 268].

6.5.2.2 Modified Poisson-Gamma Model A Bayesian model is assumed whose evidence is Poisson distributed and the prior on λ is a gamma prior (PG model). It is subject, on the one hand, to the assumption that successive repeat purchases are uncorrelated and that repeat purchases follow a homogeneous Poisson process with repeat purchase rate λ. On the other hand, λ across all customers follows a gamma distribution of the form α with an inverse scale parameter β. The parameters of the product-specific gamma distributions are estimated by fitting them to the maximum likelihood estimators of the purchase rates of repeat purchase customers. This is followed by a Bayesian estimate of a customer's repeat purchase rate based on the combination of the prior distribution and the individual's past purchase history [62, 649]:

$$\lambda_{A_i,C_j} = \frac{k + \alpha_{A_i}}{t + \beta_{A_i}}, t > 0 \tag{6.6}$$

In addition to the shape and scale parameters α_{A_i} and β_{A_i} of the gamma prior of product A_i, k describes the number of purchases of the specific product A_i by customer C_j. The elapsed time between the initial purchase of A_i by C_j and the current time, is expressed by t. Regarding recommendations, R_{A_i} is assumed to be Poisson distributed, with the inverse scale parameter estimated using λ_{A_i,C_j} and the likelihood function estimated using the following equation R_{A_i,C_j} [229, 158].

$$R_{A_i,C_j}(t) = \sum_{m=1}^{\infty} \frac{\lambda_{A_i,C_j}^m \exp(\lambda_{A_i}, C_j)}{m!}, t > 0 \tag{6.7}$$

Here m represents the number of expected future purchases and Q_{A_i} is considered as a fixed constant for all products A_i. Recommendations are made by classifying all items previously purchased repeatedly by the customer based on their estimated probability density P_{A_i} in descending order at a given time t using the equation 6.2 [62].

Unlike the classical Bayesian methods, the Bayes estimation of the a priori distribution here is done empirically using the underlying data, rather than defining it fixedly in advance without taking the data into account. Therefore, this method is also titled Empirical Bayesian Method. The PG model is thus a parametric Empirical Bayesian Model whose likelihood and a priori distribution take simple parametric forms. It approximates a Hierarchical Bayesian Model. The Bayes theorem allows the individual purchase decisions of a customer to be combined with the aggregate purchase behavior of a product, thereby personalizing this model. The PG model has already been used in the past, but not yet in the context of stand-alone repeat purchase recommendations [62, 158, 192, 431, 578].

If λ is the purchase rate to be estimated at $Y \in \{0,1,2,...\} \in \mathbb{N}_0$ purchases occurred in period N, then $E(Y) = N\lambda$ describes the expected number of purchases. Because λ follows a gamma distribution at the product level, a maximum likelihood estimator (MLE) of the gamma parameters $\Gamma(\alpha,\beta)$ for each product of the respective repeat purchase customers is established as $\hat{\lambda} = \frac{Y}{N}$. Because Y represents the number of purchases with the above expected value $E(Y) = N\lambda$, the likelihood assumption is $Y \mid \lambda \sim \Gamma(\alpha+Y, \beta+N)$. The Bayes estimator (a posteriori mean) can be expressed as [62, 142, 578]:

$$\hat{\lambda}_{posterior} = \frac{\alpha+Y}{\beta+N} \tag{6.8}$$

Due to the estimation of the posterior distribution of the purchase rate λ under the a priori assumption of a gamma distribution, whose parameters are approximated via the MLE using the actual purchases for each product, this is a parametric empirical Bayes model, which is an approximation of a Bayesian Hierarchical Model. The difference is that scale and shape parameters (α,β) of the gamma prior are estimated from the actual data rather than from additional hyperprior parameters (see Fig 12).

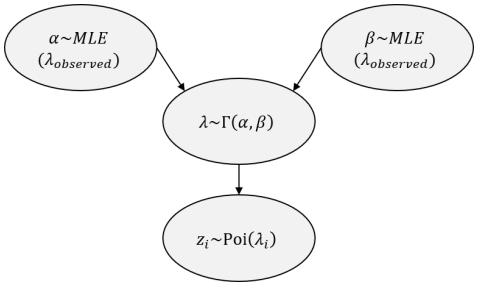


Figure 12: Empirical Bayesian Model.

However, the assumption of a homogeneous Poisson distribution is not valid for any type of product since the purchase events can theoretically represent a time-independent constant. Therefore, the PG model needs to be modified (MPG model). This is due to the fact that a Poisson process represents a limiting case of the sequence of Bernoulli processes in the boundary between a large sample and a small constant probability and is memoryless. However, this is not expected for purchase behavior for a number of products, since a customer's need to repurchase an item after purchasing it is initially small but variable as time progresses [62, 468, 158].

The MPG model assumes that a customer's purchases are correlated and repeat purchases follow a modified Poisson process, which uses a single parameter λ as the repeat purchase rate. λ depends on the last purchase of a product by the customer under consideration. Thus, it differs by the homogeneous Poisson process in the PG model. Furthermore, a gamma prior is assumed on λ. Thus, λ follows a gamma distribution of the form α across all customers with an inverse scale parameter β [158].

In analogy to the PG model, the estimates of the parameters are made via a parametric empirical Bayes model that fits them to the MLEs of customers' repeat purchase rates. The estimation is optimized per customer by estimating the mean time interval for repeated purchases of a specific item based on the first and last purchases. The model assigns the observed mean value to the associated highest repeat purchase rate. This is achieved by making modifications to the PG model: t_{buy} denotes the elapsed time interval between the first and last purchase of product A_i by customer C_j, t represents the elapsed time interval between the last purchase of A_i by C_j and the current time. t_{mean}

represents the estimated mean time interval of repeat purchases of A_i by C_j. For $t < 2 \cdot t_{mean}$, the estimation of the repeat purchase rate of the MPG model is done according to [62]:

$$\lambda_{A_i,C_j} = \frac{k + \alpha_{A_i}}{t_{buy} + 2 \cdot |t_{mean} - t| + \beta_{A_i}} \tag{6.9}$$

α_{A_i} and β_{A_i} are shape and inverse scale parameters of the gamma prior of A_i. k represents the number of purchases of A_i by C_j. If $t \geq 2 \cdot t_{mean}$, λ_{A_i,C_j} is determined via the separate formula of the MPG model. This ensures that λ_{A_i,C_j} increases from $t = 0$ to $t = t_{mean}$ and then decreases until $t = 2 \cdot t_{mean}$. At this point, the MPG model is equivalent to the PG model. A Poisson distribution is further assumed for R_{A_i}, whose likelihood function is estimated using $R_{A_i,C_j}(t)$ of the PG model. Moreover, $Q(A_i)$ can be determined via RCP_{A_i} using equation 6.4. Recommendations are generated by ranking all products in descending order based on their estimated probability density $P_{A_i}(t)$ at a given time t using equation 6.2 [62, 142, 229].

6.6 Results

The average RCP is 6.7% with a standard deviation of 12.7%. The median is 3.0%, which means that 50% of the products have a higher and 50% a lower RCP. However, products with a repeat customer probability of 100% also occur and are excluded in the following. These are mostly products that have a very low but equally high loc and moc frequency. Figure 13 provides information about the distribution of repeat customer rates for all products. It can be seen that the most frequent values are concentrated in the range of a $RCP < 0.2$. Outliers are considered for all products with a $RCP > 0.13$.

Determining the threshold for products to be considered suitable for repeat purchase is therefore difficult. One study outlines that products are to be classified as suitable for repeat purchase if they have a $RCP > 0.27$ [410]. However, different thresholds are tested to show the impact of RCP filtering, as this is expected to have a large impact on the model results.

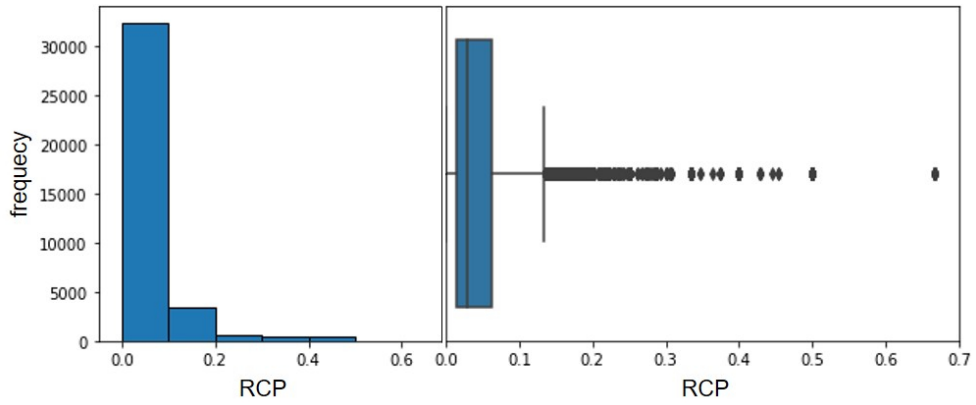


Figure 13: Distribution of Repeat Customer Probability.

6.6.1 ATD Model

Since many customers in the data set make only a few repeat purchases, the ATD model draws on the aggregate and time-dependent repeat purchase behavior of all repeat purchasers of a product. Figure 14 shows that the aggregated mean repeat purchase periods of e.g. mineral water are approximately log-normally distributed. Therefore, R_{A_i} is determined over the interval of the log-normal distribution.

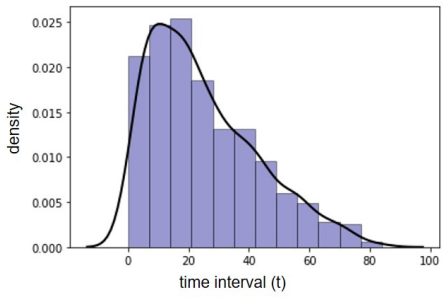


Figure 14: Distribution of Time Intervals (t) for Repeated Purchases of Mineral Water over all Customers.

Since the purchase rates at the individual level are often unrealistic due to sparse data, looking at the aggregated purchase rates of products across all repeat purchasers ensures considerable results (see Table 7). Increasing the RCP further improves these. It can be seen that an RCP of 25% provides the

best results, a further increase would lead to a model degradation.

6.6.2 MPG Model

At the beginning, the individual purchase rates of the products are determined for each customer. An exemplary distribution of the purchase rates of mineral water can be seen in the following Fig 15a). It shows that the normalized a posteriori distribution of λ achieves a good fit with respect to the data. The normalization is stronger the closer the last purchase (t) is to the current date (cf. Fig. 15b).

A conversion of the Bayesian estimators $\hat{\lambda}_{posterior}$ to the actual observed purchase rates is still evident (see Fig. 15 a, b). Even for long periods since the last purchase, the purchase rates are well reproduced. However, there is a general trend in that the closer the first purchase of a product is to the cutoff date, the more the purchase rates diverge. Figure 15 b) reveals that the estimate is more in line with the true rate when the first purchase of a product was made recently and a customer is simultaneously is a first-time repeat buyer. Figure 15 c) shows the distributions of the probabilities $P_\lambda(Y = 1)$ of all repeat buyers.

Table 7 shows that the MPG model outperforms the ATD model. This is due to the fact that Bayesian model-based methods can perform well with sparse data [285] and personalized purchase rates are considered here. Also for the MPG model, the best results are obtained with an RCP of 25%. Compared to the ATD model, the MPG model leads on average to an improvement in precision@k of 26.3%, recall@k of 21.1% and f1@k of 23.9%. These are particularly driven by the increases under a relatively small RCP of 7%.

6.7 Discussion

6.7.1 Model Results

With regard to the problems of traditional recommender systems, the model-based methods considered can reduce many weaknesses. They are suitable for sparse data, which is achieved in the ATD model by aggregations of the mean

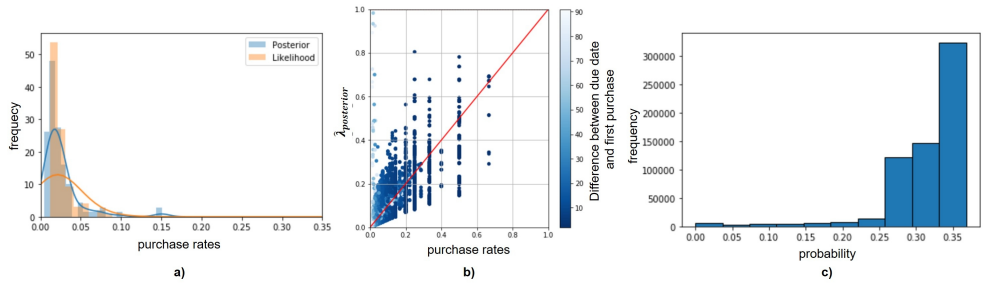


Figure 15: MPG-Model – a) Likelihood and a Posteriori Distribution of Mineral Water, b) Residuals between True Purchase Rate and the Bayes Estimator and c) Distribution of Probabilities of all Customers for a Purchase of Mineral Water within 30 Days.

Table 7: Model Evaluation.

		0.07			0.20			0.25			0.30		
model	k	p@k	r@k	f1@k	p@k	r@k	f1@k	p@k	r@k	f1@k	p@k	r@k	f1@k
ATD	1	0.14	0.16	0.15	0.22	0.36	0.27	0.23	0.37	0.28	0.09	0.14	0.11
	5	0.14	0.22	0.17	0.23	0.38	0.29	0.24	0.38	0.29	0.10	0.15	0.12
	10	0.15	0.22	0.18	0.23	0.38	0.29	0.24	0.38	0.29	0.10	0.15	0.12
MPG	1	0.17	0.24	0.20	0.24	0.40	0.30	0.27	0.42	0.33	0.12	0.16	0.14
	5	0.22	0.27	0.24	0.24	0.42	0.31	0.28	0.44	0.34	0.14	0.20	0.16
	10	0.22	0.27	0.24	0.25	0.42	0.31	0.28	0.44	0.34	0.14	0.20	0.16
where: p@k: precision@k, r@k: recall@k, f1@k: f1score@k													

repeat purchase intervals per product and complemented in the MPG model by Bayesian personalization of purchase rates. The high computational effort regarding memory-based methods is also improved here by using probabilistic models. The results show that the MPG model can outperform the ATD model. In particular, when a low RCP is applied, the potential of the MPG model becomes apparent (see Table 7).

Furthermore, in the food context, the model assumptions show better suitability than similar models discussed for repeat purchases in the marketing literature. These focus on repeat purchases of brands, where the assumption is that an increasing time difference between the current date and last purchase results in a decreasing probability for a repeated purchase [423]. Especially for food (e.g. staple foods), the assumption is contradictory. Here, customers are more likely to buy a product if this time difference increases. This is, in par-

ticular, true as brand loyalty is rather low, but products need to be purchased frequently.

Nevertheless, MPG model errors often occured, when a customer was rarely a repeat buyer and usually just before the set cutoff date, so a small number of purchases were combined with a small time gap from the first purchase of the product. This is precisely the case if the customer does not appear again in the test data set and not regularly as a repeat buyer of the product. This can be explained by the shifting behavior of consumers, who frequently change the point of purchase (e.g. supermarket and discount store) or adopt different consumption behaviors (e.g. going vegan) [352, 605]. Further research should therefore address the integration of changing preferences and shopping habits.

6.7.2 Influence of the RCP

Since there is no explicit information on RCP in food retailing in the literature, different thresholds were tested here. The results show that the RCP has a strong influence on the model quality. With increasing filtering of the products regarding their probability that a customer buys the product repeatedly, the model quality increases significantly. As a study suggests [410], in this case an RCP of 25% provides the best results. However, this is related to a trade-off between model goodness and the product variety considered, which – especially with sparsely populated data – also filters products that normally follow a habitualized and regular purchase process.

If an RCP of 0.3 is reached, there is a decline of the model, because now mostly only products are considered, which combine low loc and moc values, resulting in a tendentially high RCP.

6.8 Conclusion

This paper has focused on validating a time-based recommender system for repeat purchase decisions in food retailing. For this purpose, an introduction to recommender systems was given at the beginning and classical problems for transaction data of food retail were derived. The consideration of state-of-

the-art solutions for the integration of a temporal component into recommendations created a transition to current models in the marketing literature. Two specific models were identified that deal specifically with repeat purchase decisions. These were applied to a real data set of a stationary German food retailer and discussed with respect to the previously derived problems of recommender systems. It is shown that the Modified Poisson-Gamma model is well suited for the sparse data situation at hand, but that model inaccuracies occur due to different consumption and shopping habits. Therefore, future research should focus on the integration of these consumption patterns.

7 Comparative Summary of Findings

The previous sections shed light on research question RQ1 and consider companies' challenges in Business-Data Alignment. In particular, integration platforms to support data sourcing, which forms the basis for developing DDSs, were examined. On the other hand, a DDS was developed as a recommender system in food retail, enhancing the business model with data (data-enhanced business model, see section 2.1.2). The findings of the two studies are summarized comparatively to highlight abstract implications for supporting businesses in Business-Data Alignment. These implications are also included in part IV to develop suitable tools to help companies create DDBMs and value-added DDSs.

7.1 Technical and Organizational Challenges in Data-Driven Service Development

The increasing importance of data as a strategic asset for the creation of DDBMs and as a basis for the development of DDSs requires their systematic integration into the organizational structures to make decisions based on the data stocks in the area of tension between the alignment of data and business strategy (see section 2.2).

Developing the recommender system for repeat purchase decisions contributes to creating a data-enhanced business model in food retail that offers added value through personalized recommendations [122]. However, it became apparent that the heterogeneity of data sources is a crucial problem here. Transaction data from different channels and touchpoints must be linked by businesses to enable a good performance [106, 663]. For repeat purchases, there is the specific challenge of seamlessly integrating all purchase processes and transactions, as otherwise, the purchase rates are modeled incorrectly (see section 6). Although an aggregated probabilistic model can provide a remedy by using mean purchase rates, changing consumption practices and changes in the point of purchase result in a missing consideration of necessary data due to a lack of integration options across all consumer practices. Businesses

must align their data strategy with changing consumption behaviors to develop more adequate DDSs. Using the example of food retailing, this requires the strategic business coordination of all branches of a food retail chain to reduce the obstacles caused by data fragmentation due to a lack of transparency and control over the data [529, 148].

The integration of data thus presents companies with the challenge of harmonizing large amounts of heterogeneous data from different sources and integrating it efficiently into their processes and DDSs [248, 529, 148]. Section 5 shows that IPaaS platforms are a suitable tool to facilitate data integration and, thus, data sourcing for developing DDSs. However, selecting a suitable IPaaS solution requires carefully assessing the technical and business requirements.

The experts in the study show that the standardization of data models is a decisive factor in adopting IPaaS platforms. Companies often work with data sets from various sources, such as SaaS applications or on-premise systems. An adequate IPaaS platform should, therefore, offer the flexibility and adaptability to convert data into standardized formats so that it can be seamlessly integrated into different systems and made usable. Without this standardization, it can lead to fragmented and inefficient solutions, making it almost impossible to implement a coherent data strategy.

According to the experts, customizable and diverse connectors are another critical factor in supporting companies in data integration. A cost-benefit analysis occurs in this context, setting high costs concerning the required functions. Small and medium-sized companies often find it difficult to justify the investment in IPaaS if the solution is oversized and does not offer the expected added value. The criticism is that IPaaS platforms often offer too many unnecessary connectors, while specific connectors for particular business requirements are unavailable. Accordingly, the experts point out that companies should choose IPaaS solutions that provide a wide range of connectors. However, these connectors must also be adaptable to individual preferences and needs to support specific processes [148, 529].

Compliance and data security are additional factors that the experts are focusing on. It is essential to ensure that the IPaaS platforms offer efficient integration and that the regulatory requirements for data protection and consumer

privacy are ensured. For example, it was shown that if several applications are linked in a process, the data is exchanged between these applications and, accordingly, even with a European provider that is concerned about data protection, it can end up with another provider that does not meet the requirements of consumers for the protection of their data. For example, even providers of European applications demand the possibility that, in the event of data integration and the subsequent linking of their solution with other non-European applications, it can be legally proven who is responsible for data misuse, for example.

Finally, it is emphasized that IPaaS solutions offer flexibility and scalability to meet the growing requirements of companies and can cope with increasing data volumes due to an increasingly data-driven business model and the increasing integration of data into business processes and DDSs [324, 44, 505]. The need for scalability is also evident in the case study-based development of a recommender system in food retailing, as the processing of large amounts of data is often very resource-intensive [529, 284, 14]. For this purpose, probabilistic model-based methods like the MPG model were highlighted, which can provide a remedy here [62].

7.2 Personalization of Data-Driven Services Beyond the Average Consumer

When evaluating the results of the different models tested in section 6, it becomes evident that if a customer rarely made a repeat purchase shortly before the cut-off date that separates training and test data, resulting in a small number of purchases with a small time gap to the first purchase of the product, this leads to an increased error rate.

These problems can be attributed to the case where the customer does not reappear in the test data set and is, therefore, not regularly classified as a repeat shopper. This lack of reappearance can be explained by the changing behavior of consumers who frequently change their place of purchase (e.g., between supermarket and discount store) and therefore always buy their food from a different place and therefore do not provide consistent transaction data

or by changing their consumption behavior (e.g., veganism) [352, 605].

It becomes evident that the consideration of individual consumer behavior through personalization is necessary to develop adequate DDSs. However, the existing literature shows that existing recommender systems are often primarily based on popularity or similarity [644], and thus, individual consumption patterns outside the norm are insufficiently integrated (see section 2.3.3). This is underpinned by the recommender system evaluation in section 6, which is used in online retail and represents the current state of the art regarding repeat purchase recommendations. The approximation of a Hierarchical Bayesian Model (MPG model) makes it possible to personalize a consumer's purchase decisions, combined with the aggregated purchase rates of a product among all buyers. This means that individual purchase rates are taken into account to some extent. Still, these can be distorted by considering the aggregated rates on the one hand, and on the other hand, a change in individual purchase intervals is considered, but not changing consumer preferences.

As the case studies in part II on Business-Data Alignment focus on technological issues in the development of DDS rather than on user preferences, the need to consider changing preferences and shopping habits and to develop models that also address different consumer types becomes clear. To this end, a user-centered design of such DDSs should be at the forefront to understand how different consumer preferences can be appropriately considered and subsequently integrated into the solution.

This aspect emphasizes that purely technology-driven developments for offering DDSs [5] are insufficient to create added value (see section 2.4.1), which must also be appropriate for those consumers who follow niche practices and thus tend to be lost in overall aggregated data. Accordingly, the following part III deals with the consumer perspective on DDSs to gain insights into their perceptions and needs in the context of DDSs and achieve a business-consumer Fit.

Part III

Exploring Consumer Practices and Consumer Data Strategies for Consumer-Data Alignment

8 Introduction

DDBMs are central to companies' value creation in an increasingly digitalized world. The focus here is on aligning the business strategy with the data strategy in the form of Business-Data Alignment to offer innovative services. Business-Data Alignment contributes to efficiently using data assets to optimize business operations and support strategic goals.

However, the internal alignment of data and business goals alone is insufficient in a world where DDSs increasingly shape consumers' everyday lives. Users act as consumers and suppliers of DDSs, meaning that added value must be created to best support consumers in their needs and practices [692].

The previous section 6 has shown that the use of current models in the development of DDSs is often insufficient in terms of supporting consumption practices, as the metrics are aimed at the average consumer and are not oriented towards consumption practices and user needs [644, 5].

User-centered service design is dedicated to needs-, preferences- and goal-oriented design for and with consumers. It closes the gap between technical possibilities and user-centered DDSs (see section 2.4.1). The participatory design focuses on democratizing the design process and balancing the business interests of companies and consumers' individual goals and practices through their active involvement. Without adequate consideration of the user perspective, there is a risk that technologically sophisticated but ultimately useless or unhelpful solutions will be created [5].

Joint negotiation within the participatory design framework can create added value on both sides. However, the problem here is that due to the complexity of dealing with data, there is usually a lack of clarity about the consumer's data strategy (see section 2.3.2). Data literacy is needed to understand a data strategy and assess the value of data, both for consumers [488] and businesses [501]. However, there is a lack of practical tools to support a typical design process that considers the complexity of data and addresses a common language and understanding [114].

To support a joint design with data and consumers in the negotiation processes with businesses (see section 2.4), this part aims to gain insights into the prac-

tices of consumers, their needs, and goals in the user-centered design of DDSs and their handling of data in Consumer-Data Alignment. To this end, three case studies were conducted in the consumer sectors of food and mobility, which are currently being heavily penetrated by digitalization (see section 1.1 and 1.2).

As already introduced in section 1.2, consumers face various problems in the respective fields of application that are to be addressed with DDSs (see Fig. 3). For example, recommender systems in the food application field are dedicated to reducing information overload due to an overwhelming abundance of data by personalizing the information presented, primarily based on popularity or similarity-based metrics. However, section 6 confirms existing research (see section 2.3.3) that these recommendations are not suitable for all consumption types and practices but are predominantly tailored to the average consumer to stimulate consumption [586]. To find out how critical consumers like vegans can take a stronger position in DDSs, the following case study in section 9 follows from the previous section 6. To this end, a participatory design approach is being pursued to develop a personalized food recommender system to support critical consumers. An initial empirical preliminary study was conducted with 24 consumers to create an understanding of current practices and value negotiations. In the following, we developed the recommender system and evaluated it with ten consumers. It shows how the complexity of searching for products and shops in the context of ethical values and sustainable consumption practices can be reduced by DDSs and what implications this has for the design to support a wide range of coordination processes in the context of Consumer-Data Alignment.

The mobility application area deals with peer-to-peer carsharing as part of the sharing economy, the mutual trust mechanisms involved, and the implications for intermediary platforms as digital business models (see section 1.2). P2P carsharing is based on the idea of communal use and promotes the sharing of private vehicles with other people. The shared use of vehicles contributes to reduced individual car ownership and, thus, a lower environmental impact. Peer-to-peer transactions often occur on digital platforms, which are economically part of the platform economy. P2P carsharing platforms are business models that enable renters and renters of private vehicles to contact

each other. In addition to providing the technical infrastructure for processing transactions, P2P carsharing platforms act as intermediaries between the parties and orchestrate the transactions.

As already described in section 2.3.4, trust plays a central role in the sharing economy. Renting out a private car to a stranger requires a high level of interpersonal trust [526] in the renter and their responsible and reliable behavior, which is problematic due to the anonymity of digital platforms. Due to insufficient information about trustworthiness, there is also talk of information asymmetry about the true intentions of the renter, which contributes to the unwillingness of landlords to offer their private vehicles. In the platform economy, trust is often built up through reputation systems, which are based on the experiences of previous users in the form of ratings or reviews. However, these systems are subjective and prone to bias and manipulation. To overcome these challenges, algorithm-based reputation systems based on objective telematics data, such as driving behavior and vehicle usage, are under discussion. Two consecutive case studies were conducted to determine whether DDSs, such as algorithm-based reputation systems, can overcome these challenges and how consumers align their data strategy with the consumer strategy.

The first study in section 10 looks at the tension between trust and privacy. It gathers insights into what telematics data renters are interested in and to what extent renters would share this data. This data forms the basis for an algorithm-based reputation system and thus shows how an alignment between rentees as providers and renters as consumers can occur. Therefore, three focus groups and eight individual interviews were conducted, identifying an area of negotiability in which consumers are willing to disclose critical data for appropriate added value. Furthermore, insights into the data strategies of consumers revealed that a perspective adoption of the provider often characterizes these. This adoption of perspectives occurs when consumers recognize a legitimate interest in the shared data and the corresponding purpose.

Based on the identified principles for an algorithm-based reputation system, the following case study in section 11 is built upon, and an algorithm-based reputation system is co-designed. The aim is to uncover the extent to which

such a DDS can address the problems identified above in P2P carsharing and support trust building in P2P carsharing. Therefore, we initially conducted 16 problem-centered interviews to investigate consumers' understanding of algorithm-based ratings. A prototype app was then co-designed and finally evaluated with 12 interview participants. The results show that such a DDS can contribute to building trust and minimize the problems of traditional user ratings. Furthermore, insights on the user-centered design of such DDSs are collected, especially regarding integrating individual preferences and norms.

In addition to deriving factors contributing to the support of Consumer-Data alignment, contributions to the expansion of recommender and reputation systems literature from the consumer perspective are also obtained (see Fig. 3).

9 Buying the 'Right' Thing: Designing Food Recommender Systems with Critical Consumers

Abstract

Critical consumerism is complex as ethical values are difficult to negotiate, appropriate products are hard to find, and product information is overwhelming. Although recommender systems offer solutions to reduce such complexity, current designs are not appropriate for niche practices and use non-personalized intransparent ethics. To support critical consumption, we conducted a design case study on a personalized food recommender system. Therefore, we first conducted an empirical pre-study with 24 consumers to understand value negotiations and current practices, co-designed the recommender system, and finally evaluated it in a real-world trial with ten consumers. Our findings show how recommender systems can support the negotiation of ethical values within the context of consumption practices, reduce the complexity of finding products and stores, and strengthen consumers. In addition to providing implications for the design to support critical consumption practices, we critically reflect on the scope of such recommender systems and its appropriation.

9.1 Introduction

> "The notions of 'critical/ political/ ethical consumer' and 'critical/ political/ ethical consumerism' play an increasingly important role in discussions on how to bring about a sustainable development in a market driven economy [234, 647, 613]."[224]

From an ecological and social perspective, critical consumerism[2] has taken on a new significance as the consequences of food consumption choices have become more apparent. Critical consumers, as diverse as this group is, address various environmental or social issues of the food system, e.g. a reduction of

[2] We refer to critical/political/ethical consumers as critical consumers.

meat, fair trade, or organic food [224]. Addressing these issues, prior HCI research suggests to "support planning and shopping by making information accessible and visible" and considering "costs, sustainable, ethical, and health" values in design [212]. Thereby, according to Håkansson et al. [240], design needs to support the personalization of what are the 'right' values, rather than enforce a normative understanding.

Such design, however, needs to support consumers within the complexity of practices that are entangled with cultural aspects and personal situations, e.g. budget, allergies and surrounding infrastructures with different stores and ten thousands of products [475, 110, 224]. This burdens consumers with making decisions with conflicting values and under information overload [224].

Reducing information overload and finding matching products is usually addressed by recommender systems (RecSys). Current systems are, however, designed from a vendor perspective to motivate and nudge more consumption [586, 226], rather than to engage and support critical consumption. And even 'independent' RecSys that are presumed to support consumers, turn out to be driven by market interests [150, 303]. Also in academia, food RecSys neglect the complexity and diversity of diets and practices, e.g. they do not work well for vegan or vegetarian diets [644], as they focus on popular food rather than niches and do not take the implications, e.g. allergies or values, into account. Moreover, they reinforce current behavior, instead of allowing for exploration and setting new goals [599], and rely on intransparent ethics [303]. Resulting from this, there is a **research gap** in the question of how personalized food RecSys can be designed and used to support critical consumers in realizing their complex values, ethics, and restrictions [303].

In light of this research gap, this design case study [691] aims to combine both, the discourse of sustainable food practices and personalized food RecSys, by examining how critical consumers can be supported by a personalized RecSys and how such a system should be designed to appropriately to consider their values, restrictions and situations. Based on the literature review to frame the research focus (Section 2), we conducted a pre-study with 24 consumers to understand their current consumption practices and present first considerations for the design (Section 3). The challenges and require-

ments identified in this study were addressed in a co-creation workshop that supported and inspired the development of F-RS4CC, our knowledge-based **F**ood **R**ecommender **S**ystem for **C**ritical **C**onsumers, which is presented in Section 4. Finally, in Section 5, we present the evaluation of F-RS4CC with ten consumers. The focus of the study was not on technical accuracy, but practical requirements, appropriation, and the scope of the approach.

The contribution of our research is twofold. First, we contribute to food RecSys by designing and evaluating our prototype and showing how to bring in personalized values and context information. Thereby, the evaluation shows how future design needs to go beyond ethical values and incorporate settings for the personal situation and allow for different complexities of profiles. From a sustainable food practice perspective, our results contribute by showing how food RecSys can indeed reduce complexity but are limited in regard to everyday usage as consumers' shopping planning turned out to be informed by an established set of products rather than new explorations. Still, in new situations, consumers found the prototype helpful to support new situations.

9.2 Related Work

9.2.1 Critical/Political/Ethical Consumerism

The nexus around critical consumerism is often defined as ethical-value-driven and more conscious consumption practices that consider the broader implications of own consumption. While this nexus of critical, ethical, and political consumption is often used without any exclusive definition, it mostly refers "to the conscious choices implied in attempting to connect values and actions in the market-place"[224] and "may be seen as a form of political, or extra-parliamentary participation"[224] in the food system.

Critical consumerism typically addresses personal concerns, such as health, and broader altruistic values "such as human rights or work rights, [...] use of scarce resources, environmental impact, climate change, animal welfare, and politics, such as perceived social irresponsibility"[224]. However, bringing values into action is quite difficult. Consumers face the challenge of balancing their often numerous ethical values and restrictions against the offering

of their surrounding markets and the available budget, gain practical knowledge about products, and resist unplanned shopping desires [110]. From a practice-theoretical perspective, the transition from unconscious consumption practices to critical consumption practices can be seen as a learning of new competencies, exploration of new products and stores, and negotiation and appropriation of values [652, 350].

Although one individual consumer might not change the whole food system, consumers aim to diffuse change into society by sharing practices and the meanings of alternative consumption [300, 547]. This participation and its expected influence on the food system are criticized by some researchers as the political dimension is considered an expression of legislative weakness in neo-liberalism [45] and consumers might be disenchanted about the effect of their consumption [224]. Others [422, 612], however, see it as complementing legislation by low-threshold and non-bureaucratic market participation, which allows even groups with little legislative power to shape the food system.

Supporting the planning of shopping is seen as a major opportunity to bring critical consumer values into action [110, 212]. Thereby, previous research found that consumers use various information resources to inform their consumption decisions [459, 350]. However, this is challenging as products are not properly labeled, information is not given, and appropriate products are invisible within the nearly infinite number of products and stores [350].

9.2.2 HCI to Buy the 'Right' Things

Closely related to critical consumerism, Prost et al.[495] introduced the concept of food democracy. This concept entails going a step beyond the neoliberal roles of consumers and producers and start building on alternative food systems, e.g. urban gardens or consumer co-operatives, that compete with the current system. Still, as not everybody can participate in alternative food systems and systems may co-exist [495, 244], "we also need to acknowledge and design with (or around) already existing food infrastructures with competing interests and values"[496]. Supermarkets, as "the visible end of the corporate food infrastructure and their hegemony [are] prevalent through their efficiency, pervasiveness, and in many cases, aggressively cheap pricing"[496].

Addressing those infrastructures, Ganglbauer et al. [212] suggest to "support planning and shopping by making information accessible and visible" [212]. In doing so, design should consider "costs, sustainable, ethical, and health" [212] values in a personalized rather than normative way [240]. This entails not only the values but also the cultural background of the consumers or their personal situation, such as budget, available stores, and products [475, 224]. Moreover, Parker and Grinter [475] conclude that future design should "provide insight into how users can eat healthfully given their cultural, local and temporal reality". Towards this end, HCI has been researching the support of buying the 'right' food from different perspectives.

For healthy food, research [298] focused on reducing the complexity of nutritional features by introducing more simple indicators. To support in-store decision making and searching for healthy alternatives, Ahn et al. [17] introduced an augmented reality prototype that successfully reduced the search time during shopping. Further research [509] introduced systems to reflect on consumption patterns and track health-related goals of past purchase decisions during shopping, but also at home.

Similarly, for sustainable purchases, research focused on in-store decision making and reflection on consumption patterns to bridge intention-behavior gap [573]. Some research [299, 64] engaged in decision support during shopping. Their research used the opportunities of scanning products with the shopping cart and visualizing simple information on regionality or organic quality with LEDs. The display allowed for comparison with other consumers. Likewise, Herbig et al. [265] designed and evaluated a mobile app to in-store, scan and assess products. They highlight the importance of trustworthy data and customization towards their understandings of sustainability. Still, scanning products in the store is assessed as a tedious task that needs to be reduced. Besides, research[304] used digital receipts to allow participants to reflect on their organic consumption patterns. Reflection here proved to increase organic food consumption.

While the presented research focused on motivational aspects and in-store assessment, research on critical consumption highlights the importance of planning [110] as well as the exploration of products and retailers that sell

the appropriate products [350]. Regarding this, Clear et al. [133] suppose to support purchase planning by showing sustainability information right behind the selected item in the shopping list. A broader evaluation of the approach is missing.

9.2.3 Food Recommender Systems

To reduce information overload and filter for appropriate products and stores, RecSys offer a well-known solution, as they are already adopted by millions of users and allow for real-time information access at any time [303, 185].

According to Tran et al. [641] there are four types of food RecSys. The first type gives recommendations of recipes or products based on similarity to consumer preferences [203, 255, 183, 641]. Type two integrates the recommendations of (health) professionals and their assessments [653, 3]. While type three represents a link between the food preferences of individual consumers and nutritional recommendations [185], type four generates recommendations for consumer groups, e.g. a family or group of peers [184, 641].

Within these attempts, mostly content-based, collaborative filtering or hybrid approaches were used for food recommendations [641, 648]. Thus, recommendations are mostly based on popularity or similarity of food. Based on this dominant approach, food RecSys do not work well for critical consumer practices, e.g. recommending vegan food [644]. While this approach ensures broad acceptance of the recommendation, the recommended foods are rather unhealthy and unsustainable [599, 645, 646].

Due to the outlined issues, current approaches start to reexplore the opportunities of knowledge-based approaches where the aim is to recommend food based on a holistic user profile that includes individual health conditions [436]. However, there is still a gap in research on food RecSys that consider the preferences of individual consumers in a profile [641] and bring personalized critical consumer values and ethics into action [303].

Regarding this, Karpati et al. [303] found that there are already some food RecSys available that use ethical concepts. However, those are not personalized, which means that consumers are only able to find the 'right' product

if they share the value set of the RecSys provider. For those services, ethical guidelines, algorithms, and the underlying business model are often intransparent [150, 303]. Therefore, they [303] argue for research on RecSys that allow for personalization of ethical values and consideration of consumers' needs for transparency and trust.

9.3 Empirical Pre-Study: Understanding Critical Consumer Practices

The objective of our empirical pre-study is to understand the diversity of critical consumer practices, current approaches to finding appropriate products, and potentials of a food RecSys to support planning and exploration.

We therefore conducted and analyzed 24 (PS1-24) interviews (avg. 26 min) with critical consumers. The sample has been recruited with a snowball sampling approach, starting from contacts to critical consumers from prior studies. In taking such an approach, we were not aiming for a representative sample, but a broad and diverse sample regarding critical consumer practices, values, and combination of values and other restrictions. For example, our sample covers vegetarians, vegans, allergies, political boycott, religious restrictions, diet (low-carb, low-sugar), sustainable/organic and plastic/packaging reduced consumption practices. The sample is characterized by 9 male and 15 female participants, aged between 21 and 35. The education ranges from apprenticeship to a university degree. Living conditions range from alone to living with a family.

The semi-structured interviews [38] followed an interview guideline that covered the participant's food consumption practices, including their ethical values and restrictions, issues of critical consumerism, and their current efforts to purchase decisions. All interviews were transcribed and the analysis was conducted by two researchers following the inductive approach of thematic analysis [86]. After coding, codes were consolidated and themes were collaboratively developed by the group of authors towards a mutual understanding of the material.

9.3.1 Critical Consumerism

Our participants' perceptions of their consumption were as diverse and ambiguous as the definitions of critical consumption in the literature. Some identified themselves with role models, such as vegetarian, vegan, or political consumer, while others loosely describe their restrictions and considered (ethical) values.

> "I use the word vegan or would describe myself that way because it is easier. But actually I pay attention to consumption. Consumption is a crucial point for me, because for me every purchase is a ballot. I am actively influencing our economy with it." – PS11(26,m)

However, none of our participants was just committed to one restriction or value. In particular, we found various examples, where participants described their consumerism as evolving from one direction towards the negotiation of different goals and ethical stances. For example, veganism to minimalism (PS7(26,f)), religious to political boycott (PS15(29,f)), food to whole lifestyle (PS14(31,m)), or sustainability to political boycott (PS16(35,m)). Thereby, the goals do not necessarily have to be altruistic ones, but complexity can arise from additional personal goals as well, e.g. reducing calories (PS20(30,m)), allergies (PS18(26,f)), keto-diet (PS22(26,m)), reduction of ingredients suspected of being carcinogenic (PS17(25,m)), or health reasons (PS16(35,m)).

> "So for many people it is only vegan eating or clothing, for me it actually means the complete lifestyle [...] I'm really trying to keep the whole lifestyle completely free of suffering and I try to extend it to things like not shopping at H&M or Primark, trying to buy organic, as environmentally friendly and free of suffering, also for humans." – PS9(27,f)

In a broader reflection on the political or economic aspects of their consumption, our participants expressed ambiguous feelings. The perspective, thereby, ranged from having high market power and that every consumption matters,

to little power for individual consumers, but mass matters, and finally to rejection of power, but an emphasis on being at peace with oneself and own values.

> "I don't know if I'm making a mark with this but I believe that the crowd can make a mark by buying certain products less, so a single one probably won't make any difference but in masses you can make a difference" – PS15(29,f)

9.3.2 Complexity & Effort of Critical Consumerism

Participants described their practices as associated with various efforts and complexities that need to be resolved. Although, most participants explained that their consumption practices become more convenient and easy over time, e.g. when they found new 'basic' products that match their needs. Nevertheless, in uncertain situations, e.g. facing new products or moving, new complexity arises.

> "The problem is the everyday life. Everything I just said [about my consumption] sounds very good and idealistic. But in the end you are only human. And that is really exhausting. Personally I find it very exhausting to go through it all time. If something slips in between, I am not so strict. [...] Of course it is easier than before, but if you want to go away or eat something else." – PS11(26,m)

Thereby, the complexity arises from various issues that are entangled with the situation of the consumers. Participants articulated how they try to find products that best match their values, but still, trade-offs have to be made between different values as there is no 'right' thing. In some cases, there is even no product that fully matches their requirements, at least from perception.

> "For a long time I thought, [...] organic is good, but when I get it sold in cardboard and plastic, I start to have doubts and think

> 'hmm, then I'd better buy the normal ones that are on the shelf without packaging, regardless of whether they say 'organic' or not, because that's part of the deal for me. One tries to see the whole thing and not just trust on some label" – PS24(30,m)

Moreover, not all products are available at a single retailer or even retailers nearby. Although most of our participants explained how they usually purchase products at bigger retail stores, even those do not have all the products they need. Therefore, another problem arises from getting an first overview of the offerings of different stores, and second, managing shopping trips that include multiple stores.

> "Sometimes I think my life is a bit too hectic to think about it. I would really need an afternoon, okay I'll go to the zero waste store and buy this and that. The problem is that sometimes some things are used up faster than others, which means I buy it on demand. And so it is hard for me to say okay, I am gonna go to the zero waste store again. Exactly because this convenience factor is essential for me and that I am looking for which store is close." – PS8(26,f)

The field of tension between products, stores, and ethical values is further influenced by the personal situation. On the one side, consumers already struggle to weigh the properties of the food, as the above examples of PS24 shows. On the other side, time in everyday life, financial capabilities or simply the distance to the next store add tension. Accordingly, this tension in everyday life can sometimes not be resolved in favor of the value-matching product, but the given situation. For example, the budget especially concerned the students in our sample, but also unemployed or low-income consumers. They describe organic and unpackaged products as more expensive than regular products and therefore, not in scope of every shopping trip.

> "This week I bought in the organic supermarket for a lot of money. 50 euros. Then I know the next few weeks I should perhaps not necessarily go out to eat and that I should not go to the organic

> supermarket again. Then I go for example, to a discounter and just buy the basics and try to cook more for me." – PS8(26,f)

Furthermore, family, roommates, or partners are a factor that adds effort to food purchase. Besides, friends are mentioned, e.g. when the participants invite for dinner. At this point, the own consumption needs to be aligned with that of other consumers. Thereby, these can also have their ethical values that are not necessarily compatible with the participant's values.

> "I'm not always strict on it, because I can't have so much influence on it. I am just one person and most people, for example, my family, they consume exactly what I partly avoid." – PS19(27,f)

9.3.3 Current Approach to Purchase & Planning

To cope with the complexities, consumers developed certain strategies to buy the 'right' thing. Thereby, the acquisition of food information turned out to be an important issue. This includes certain ingredients, labels, and sources for further checking of the quality, impacts, and values of the products. In this manner, participants describe how they use various labels, e.g. the vegan, organic, or fair-trade label. In addition, for unlabeled products, the ingredient list is checked for allergens and unwanted ingredients.

> "I look at the product and there is often also the vegan label on it and otherwise I look exactly at the ingredients. I think I've probably googled some ingredients, but usually I look at what's inside, because I have to do it due to my allergies." – PS10(29,f)

Sometimes, however, consumers face the challenges of not understanding or recognizing ingredient names, finding no label or certain information is hidden or not provided, e.g. the carbon footprint or ingredients summarized as 'natural flavoring'. For this reason, most of them established strategies to check the product. A tool repeatedly mentioned is CodeCheck, an app to scan the product's barcode and access data about ingredients, qualities, and the

assessment of other consumers. Another approach is using lists of websites, googling the ingredients, or asking peers offline or in social media.

> "I google sometimes, whether something is not marked with a vegan label, whether it is vegan, for example, such a cereal. Or for example Peta, they sometimes have lists. Or CodeCheck, if there are any hints." – PS5(23,f)

Against this background, most consumers explained how they established a set of 'basic' products to reduce the tedious checking-effort in everyday shopping. Despite this set, some participants, however, address the value of pleasure and the joy of exploring new products. They do not see their consumption as a limitation, but as room for new explorations.

> "For me it is not a renunciation, but actually an enrichment, because I have been cooking much more diverse since then. And yes, I do not eat dairy products and eggs anymore, but otherwise I like to cook a lot and very much and I also attach importance to good food. I like to eat." – PS2(21,f)

Exploration, however, is associated with again negotiating values and checking products. Common sources of inspiration, such as leaflets or advertisements are not relevant for critical consumer practices, as they address regular consumers and often do not promote appropriate products or give the necessary information. Here, our participants explained how they join online communities, where other critical consumers post messages about newly explored products, answer questions about products, or just help with finding retailers that offer certain products nearby.

> "I'm in a Facebook group that's called 'vegan action' where people post discounts from different supermarkets or shops and that's sometimes very useful." – PS6(25,f)

9.3.4 Considerations for Designing Recommender Systems for Critical Consumer Practices

To support planning grocery shopping and easy access to relevant information and products, we derived design implications from the literature study and our complementary interviews. The main implications that guided the development of our prototype are outlined in the following.

- **Support Planning and Product Exploration:** In line with prior research [110, 350], the support of planning and exploration activities for critical consumers turned out to be a critical issue to be addressed by our prototype. Consumers can not make use of traditional leaflets and have to cross-validate information about products. Supporting with information and appropriate product recommendations before a purchase can provide means to reduce the effort of critical consumer practices.

- **Personalized Profiles:** Critical consumers are not a homogeneous group nor is it easy for individuals to describe their preferences as they are usually multifaceted [224, 110]. Furthermore, consumer practices underlie a transition, e.g. from vegetarian practices to vegan ones [350, 652]. Therefore, our RecSys approach should support the negotiation and exploration of values and restrictions and allow for personalized changes in the value arrangement [265]. Here, a knowledge-based approach might be appropriate to capture the complexity and flexibility of practices [436] while allowing for the ethical autonomy of the consumers [303].

- **Well-known Role-Models to Reduce Complexity:** Although not all consumers identify as well-known types of consumers or role models (e.g. vegan), the design should use such framing to give guidance for less experienced consumers and reduce the complexity of the system itself. Such role-model approaches to consumption patterns relate to the interview study were consumers explained that their approach is for example, veganism plus additional restrictions and ethical values.

- **Considering the Context of Practices:** A critical point for consump-

tion practices beyond their ethical and altruistic stance is their embedding into the personal situation of the practitioner. Thereby, the approach needs to consider the context as given by the household of the consumer, available retailers, and financial capabilities[34]. Furthermore, personal goals and restrictions, e.g. allergies, need to be considered.

9.4 Development of F-RS4CC

To support the exploration of new products and planning practices, we developed F-RS4CC as a cross-platform application. The design was accompanied by a co-creation workshop (approx. 3h) with ten participants. Again, they covered a wide range of consumer practices and values. To document the workshop we took notes on the discussions and explanations and took pictures of the paper prototypes. The consumers were between the ages of 23 and 57 (six males and four females).

The co-creation workshop focused on understanding the consumers' approach to bring their values into action. In particular, we aimed to understand the requirements for the UI, abstract ideas of how a RecSys should filter and rank the products, and which input they want/need to provide. Therefore, the workshop was organized along (1) a welcoming and problem introduction, (2) creative phase to sketch and discuss the profile UI, (3) discussion of its influence on the ranking, (4) creative phase to sketch and discuss the shopping list UI, and (5) reflection on additional factors, e.g. food sharing, and conclusion. Although the participants were critical consumers themselves, we provided personas and purchase scenarios [449] (inspired from the pre-study), to allow for participation without self-disclosure in the group, e.g. of allergies or diseases. For the creative phases, we provided standard materials, such as paper, pens, and scissors. Moreover, example product data (name, ingredients, labels, and nutrition table) and printed food labels and symbols were available for both the sketches and the discussions.

Inspired by the consumer concepts, we designed the final prototype. In the following, we present the overall workflow (see Figure 16) of F-RS4CC and

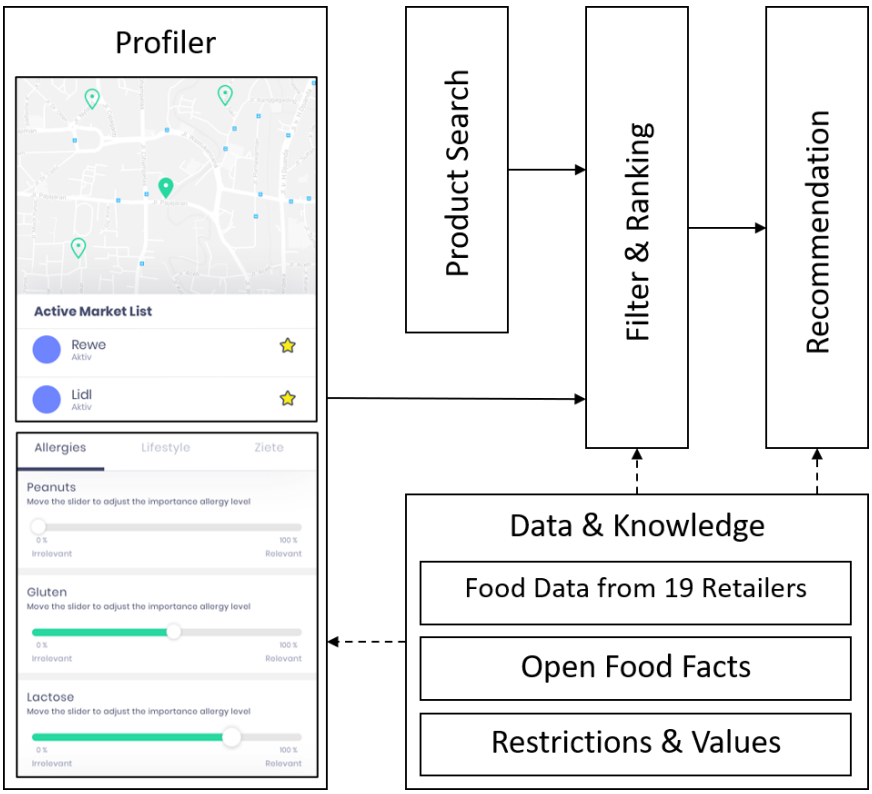


Figure 16: Workflow of F-RS4CC

give an overview of the different functions.

9.4.1 Workflow & Data Sources

The employed recommendation approach can be classified as a knowledge-based RecSys [13], as it exploits knowledge about food products, nutritional values, ingredients, labels, and environmental knowledge to identify the most suitable foods. In line with prior research [436], the recommendation is based on three main components: A profiler, a filter, and a ranking. Furthermore, we added a search and a shopping list to allow for bringing recommendations into action.

Our prototype uses data from 19 participating local retailers including supermarkets, full-organic shops, zero waste shop, and directly selling farmers. The data from these markets consists of their products and the related master data that usually contains the product name, category, ingredients, and nu-

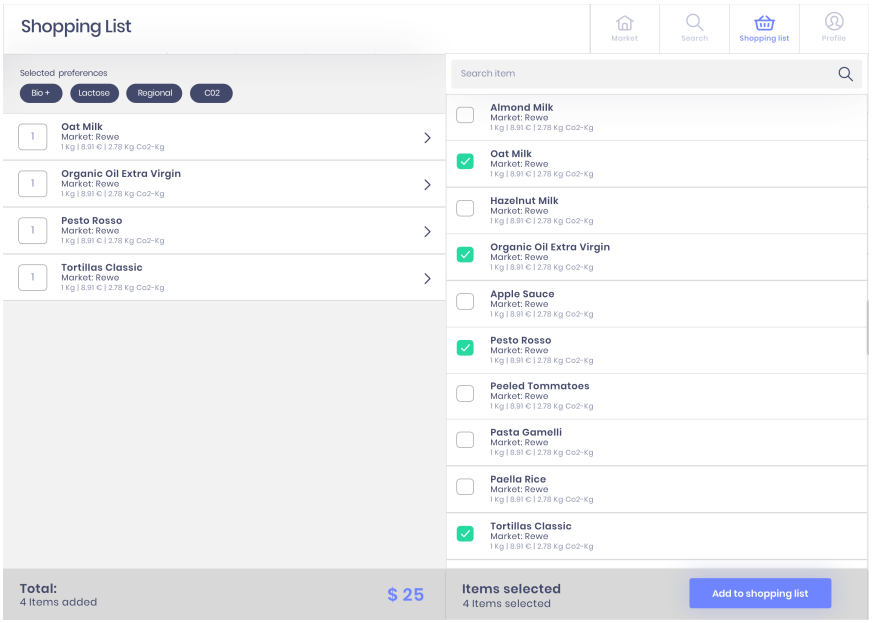


Figure 17: Search, Recommendation & Shopping List

tritional values. Furthermore, we complemented the product data with data from Open Food Facts[3], a collaborative platform that adds carbon footprint, nutritional assessment, labels, and user-based assessment of food characteristics (e.g. vegan). Therefore, we mapped the products based on their Global Trade Item Number. Moreover, food knowledge is provided in the form of rules about restrictions and ethical values that are mirrored in the properties of the food.

9.4.2 Profiler

The consumer initiates the F-RS4CC by setting up a profile that consists of two parts:

Restrictions & Values: Core to the prototype is the profile setup with critical consumer values and restrictions to infer matching food recommendations. As shown in Figure 16 (Profiler, lower screen), we broadly differentiate between allergies, altruistic, and personal goals. The allergies section consists of declarable allergens plus some additional allergens, such as apples, that came

[3]https://world.openfoodfacts.org/

Table 8: Example Rules of F-RS4CC

Rule-Name	Formalization
Vegan	Has a vegan-label or does not contain milk, butter, meat, ...
Boycott Nestle	Is not produced by nestle or one of its subsidiaries
Low Carb	¡= 20% of carbs in the nutrition table

up during the pre-study or workshop. The altruistic section, for example, includes settings such as vegan, organic, or Nestle boycott. Personal goals, for example, consist of cost-saving, low-carb, or reduction of carcinogenic additives. Overall, the system provides more than 20 different settings.

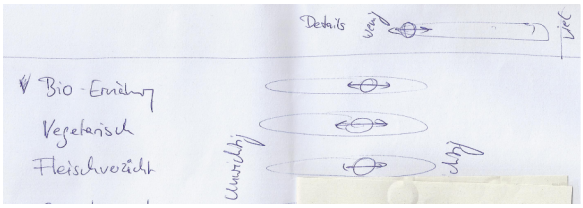


Figure 18: Conceptualization of Sliders during the Workshop

Each of these settings is based on food knowledge that is formalized as a rule to filter and rank the food products (Examples are given in Table 8). From a user perspective, each rule consists of a slider that allows for weighting each rule between Off (0%) and On (100%). Besides the ranking, for some rules, such as allergens the On-Setting has an additional filter function that filters out foods with respective ingredients completely. This is indicated in the UI as a subtext within the rule-UI-element, e.g. 'All non-vegan products are filtered out'.

This design decision is based on a deeper discussion in the co-creation workshop, where the participants tried to sketch a UI for the profiler (e.g. see Figure 18). First attempts focused on binary decisions, e.g. Do you have a soy allergy? - Yes/No. However, it became apparent that many rules can not be seen as binary restrictions but that trade-offs and rankings between different

rules exist. For example, for some products, consumers search for the vegan-as-possible recommendation, but would still eat the vegetarian one as long as it is organic. For allergy sufferers or strict political consumers, allergies or the boycott of certain companies need to be binary options. The current design, therefore, combines both the trade-off and clear restrictions.

Locations & Markets: As shown in Figure 16 (Profiler, upper screen), the consumer is shown all participating markets on a map. Here, the consumer can explore new markets and select the ones s/he is frequently visiting, e.g. in a certain radius around the neighborhood. In this sense, the selection of markets provides an additional filter criterion, as it allows to restrict the markets, e.g. to only buy from fully organic retailers or local farmers.

The selection of markets within the profiler also arose as a topic during the co-creation workshop, as the selection of markets can also support critical consumerism. For example, by supporting small retailers only or to simply account for everyday practices, such as reaching all stores by foot.

9.4.3 Filter & Ranking

The filter & ranking is executed based on the profile setup. This part of the application is based on food knowledge (see Table 8), product data from the retailers combined with Open Food Facts, and the search.

Search: Usually, knowledge-based RecSys are designed for shopping infrequent items, e.g. a new TV [13]. However, for purchasing food, it is necessary to not just get the recommendation based on the current offering of the markets, but search for specific products and get a ranked list of those, e.g. 'burger patties'. Therefore, our design makes use of a search that is matched with the product name and category. This search can be seen as a pre-filter of the items to be recommended.

Filter: Whenever certain rules, such as allergies, are set to 100% they are used as a binary filter to completely remove certain products from the recommendation candidates. As explained earlier, this need for strict filtering arose from the workshop. A further filter is based on the market selection, where only products available at those markets are included.

Ranking: After filtering and searching, the candidate set of food items is still large for most of the products. Furthermore, certain values (with a weight ¿0% and ¡100%) have not been considered by now. Therefore, based on a score for each candidate food, a ranked list of recommendations is created.

$$score(f,P,C) = \sum_{p \in P} i_p * ruleScore(f, r_p, C) \tag{9.1}$$

Thereby, the ranking is based on the $score(f,P,C)$, which denotes the overall ranking score for a food item f, the profile P and a candidate set of food items C. P is a set of all rules with i_p as the weight between 0 and 1 and r_p as the rule.

$ruleScore(f,r_p,C)$ is defined as a function of r_p that assigns a float value between -1 and 1, depending on the food item in relation to the other candidate food items. Some rules, e.g. allergies, allow for binary classification only. For those, the function either assigns 1, does not contain allergen, or -1 contains an allergen. For 'linear' rules such as carbon footprint reduction, all candidate items $c \in C$ are mapped to a value range between 1 and -1. Each item f is assigned a value between 1 and -1 according to its relation to the other candidates. Based on each score(f, P, C), food candidates are sorted in descending order to create the final ranking.

9.4.4 Food Recommendation & Shopping List

Once the list of food items is sorted, it is shown to the user (see Figure 17). Here, the products on top are the most appropriate products according to the profile. Each product is represented by its name, the price, its carbon footprint, various labels depending on the selected rules (e.g. lactose-free or not), and the selling retailer. Labels are used to sensitize consumers if the rule is not set as a binary filter but a weighted factor among others. When clicking on products, they are transferred to the shopping list. Furthermore, users can open a pop-up with more detailed product information.

In addition, the UI contains the search bar and a quick overview of the profile settings and the resulting ranking of the rules according to the assigned weights (see dark blue boxes in Figure 17). Regarding this workshop partic-

ipants demanded certain transparency about the current profile settings when receiving their recommendations.

9.5 Evaluation: Understanding the Design & Appropriation of F-RS4CC in Use

The objective of the evaluation was to discuss the design of the prototype with a focus on the profile setup, the recommendation itself, and further improvements and directions for design. Furthermore, we aimed to understand the usage and appropriation of F-RS4CC in everyday critical consumer practices.

Therefore, we conducted pre-interviews (avg. 37 min) to understand the current practices of the participants and their first assessment of the prototype, followed by a usage phase (avg. 2 months), and post-interviews (avg. 27 min) that focused on their experience, appropriation, and design issues.

Ten participants (E1-10) took part in the evaluation study. The sample was recruited with a snowball sampling approach, from the network of the pre-study. Again, the focus was on recruiting a sample with a variety of consumption practices, values, and restrictions. Moreover, we focused on diverse housing situations ranging from living alone, to flat-sharing to families with children. The sample is characterized by 6 male and 4 female participants, aged between 24 and 45. The education ranges from apprenticeship to university degree.

Both interview studies followed a semi-structured interview guideline. We transcribed the interviews and analyzed them following the inductive approach of thematic analysis [86]. The thematic analysis[86] focused on understanding the appropriation of the prototype, changes in their consumption practices, and design issues with an emphasis on the recommendation including the profile and the food knowledge. After coding, the themes were collaboratively developed by the group of authors towards a mutual understanding of the material.

9.5.1 Exploring Products & Reflection on Action

In general, the prototype was well received by the participants. They especially liked the inclusion of so many local markets and the long list of products. Here, in particular, the filtering and ranking according to the own settings were assessed as a positive concept.

However, what quickly became apparent was that the participants did not appropriate the prototype as part of their planning in the sense that they search for products, transfer them to the shopping list, and execute this planned behavior. Although five participants reported on trying to use the tool as a shopping list, they quickly returned to their established mode of purchase, which was either no shopping list at all, especially for younger consumers, a paper-based list, or the established tool within the household, e.g. a collaborative list that was shared with the partner.

> "I use a shopping list app on Android where my wife and I put in our shopping needs, I sort into several lists sorted by the shop where we buy." – E8(32,m)

Regarding this, most participants explained how product research and informing about new products is something completely different than writing a shopping list. Moreover, consumers explained how they have established a set of 'basic' products for their daily shopping. Those products meet their restrictions and ethical values, such that they do not need to start product research every time.

> "For example I sit at the table, make myself lunch or breakfast and notice "oh, the liverwurst is slowly running out" and then it's just easy to take the mobile out of my pocket, write liverwurst on it and put the mobile back in my pocket. That is also important that I do not want to start a product research at that moment." – E5(45,m)

However, participants used our F-RS4CC prototype as a tool to explore and research new products. Thereby, it was especially valued how the prototype

offers the chance to easily find new products that match the own set of values and restrictions.

> "So still I find it good to look for new vegan burgers, here I'm doing it right now. Organic veggie burger for 2,99. I do not know it yet, why I do not know it yet." – E4(32,m)

In this context, some participants compared the tool to a search engine that reduces the effort of searching for food and checking its characteristics by having a personalized profile.

> "I have to think about it, I mean, you know it a little bit from search engines, filtering can make your work incredibly easier when you know exactly what you want and what you don't want." – E1(24,f)

Reflecting on the influence on their consumption, some participants explained how they stopped buying certain products when they became aware that there are better alternatives or the product does not match their preferences. Others explained how the application rose their awareness for certain filters that they were not able to apply by their current approach. For example, the carbon footprint, besides some other factors, was repeatedly mentioned to be a factor that was added to their practices as the information was made available by F-RS4CC. Here, the reduced effort of gaining information was mentioned as a main driver for the exploration of products and profile settings.

> "I think I planned my shopping more consciously than when I normally just made a list with my roommates, because I think more consciously when you see the CO2." – E7(32,m)

9.5.2 Consolidating Infrastructures

Besides the food, some participants used the prototype to explore and consolidate the local consumption infrastructure, in the sense of finding stores that

offer appropriate products or discover formerly unknown stores. For example, the local farmers market was an option that a lot of consumers did not know before their usage experience.

> "Yes, I think it helps you that it helps you to discover products that you didn't know so much about before, or perhaps to get to know shops, for example, the Birchfarm, which is present in the program, and otherwise you would not know about it." – E10(30,f)

In this context, participants reflected on their experience of becoming critical consumers and changing their practices. Some participants explained how such a tool would have helped them in their early exploration of alternative consumption infrastructures. E9(26,f) further explained how even in new situations, such as moving to a new city, the exploration of shopping infrastructures is valuable, as it allows for a quick and easy overview of appropriate products and markets in the neighborhood.

> "Especially at the beginning or when I move somewhere, it would save me having to search everywhere, who sells this specific product I need right now, that would be cool, that would help me and of course, it would also save me a lot of work, because I wouldn't have to read all the labels in all the supermarkets to find out what it is right now, where it was produced and under what conditions." – E9(26,f)

Still, it was criticized that the current implementation of retailers as binary filters does not account for purchase practices. While participants, in general, appreciated the possibility to select nearby markets in a radius, they articulated the requirement of prioritizing the stores. The main reason mentioned was more efficiency during purchase and a reduction of time and kilometers traveled.

> "Which would also be good, would be a sorting [...] if you know that you go to one store, you can prioritize it in the product search that you don't have to drive to another store." – E7(32,m)

9.5.3 Reliability of Data & Definitions

Speaking about the recommendation itself, participants raised our awareness about the reliability of the data and the underlying food knowledge and algorithms.

A particular issue was wrongly labeled data that resulted in unwanted recommendations. While one might think that this only arises from the data provided by the Open Food Facts Community, it was also caused by retailers' data. One example of such recommendation was given by participant E5(45,m) who found chocolates that were labeled as regional when searching for sweets.

> "I just noticed, there is a lot of regional chocolate. For category regional? The supermarket has indicated this, which means that you have to check the data of the supermarket. Who knows where they get it from, you have to find out more about it." – E5(45,m)

A further issue came along with the rules themselves. Although we added a short explanation text for each profile setting, some rules were intransparent for the consumers. For example, 'organic+'(extended organic condition with better husbandry conditions) was not known by most of them. They wondered about the underlying definitions in contrast to regular organic food. Moreover, some participants recognized some minor bugs that led to wrong recommendations. This was, for example, based on ingredients that were not properly recognized.

> "We had the problem that despite my preference "vegetarian" with a 1 I was still shown some meat products, or convenience products containing meat." – E10(30,f)

Overall, wrong recommendations and wrongly labeled data were perceived as decreasing the trust in the recommendations. Reliability was particularly important in the field of tension between ethics, values, and restrictions, as participants neither want to consume products that cause an allergic reaction

nor do they want to eat, for example, products that contain animal products if they follow a strict vegan diet.

> "The database would have to be really complete, because as soon as I have the feeling that there is a gap somewhere in the database, then it loses some kind of credibility, which I would accept once or twice but then I would say 'oh no, I can't trust that'" – E4(32,m)

9.5.4 From Complex Rules to Customized Rules

Along with wrong recommendations, non-reliable data, and trust, the rules themselves were discussed in more detail. From an overall perspective, they were assessed as comprehensive and interesting to experiment with. Even though, for allergy sufferers, some special allergies could be included. For four consumers, the set up of a profile bears a certain complexity as there is too much to consider and think about when negotiating and weighing the settings.

For some consumers, the provided rules did not offer enough flexibility and room for their own settings beyond common categories and role models. They formulated the requirement to manipulate or develop their own rules. One such example was given by participant E9(26,f) who explained how for allergies some consumers have different levels of sensitivity, e.g. in grams per 100 grams of product. Similarly, this ratio was discussed for low-carb. A further example is veganism, where two participants mentioned the requirement to in detail modify rules, e.g. such that honey is not considered to violate it, as this reflects their unique approach to consumption.

> "That you can configure the filters by yourself, or a separate filter that somehow says it is vegan certified, the other one is vegan but has honey or other things in it, or it is vegan by recipe but not certified." – E6(29,m)

During the interviews, the topic of collaborative food knowledge and the sharing of modified rules were also discussed. While most of the participants

articulate more collaborative access to the profile and the list, some participants go a step further and would be interested in outsourcing and sharing of knowledge. For example, E7(32,m) explained how rules and even profiles, created by certain organizations such as Greenpeace would help with decision making. One such rule was already included in the form of the Utopia list of suspicious additives[4]. Here, two participants criticized that they do not know the organization and their assessment of ingredients, which decreases trust.

> "Utopia is unknown to me, I don't know which list is behind it." – E5(45,m)

While in general participants appreciated our knowledge-based approach, four participants explained how additional information on item similarity or even community-based approaches might contribute to better choices. For example, E4(32,m) explained how he bought newly explored vegan burgers that tasted awful. Additional information about how others liked this burger patty or how it tastes similar to products he likes would have helped him.

> "I recently bought vegan burgers [...], they were terrible. I could have skipped them if someone had commented on the taste." – E4(32,m)

9.6 Discussion

Reflecting on our research, we will discuss the preliminary design implications against the background of our evaluation from different perspectives to more precisely formulate implications for design and research.

9.6.1 Establishing a Set of Products, rather than Planning

Based on the literature research and our pre-study we expected our participants to appropriate the prototype for their planning practices. However, as it turned out, planning is completely different from exploring a new set of

[4] https://utopia.de/ratgeber/lebensmittel-zusatzstoffe-e-nummern-liste/

products or giving the established set of products an update. In this sense, our research confirms prior research [265, 350] that already lines out that consumers do not use in-store information at every shopping trip, but establish a set of frequently bought products.

Still, our results imply the explorative aspects of the design to enable more easy access to appropriate food, retailers that sell them, and the underlying information. Here, especially the combination of stores as purchase infrastructures and the food they offer helps with the arrangement of a new product set during sustainable practice transition. This set itself can be seen as a plan that is repeatedly performed [110]. This supports the assumption [350] to more strongly support the search for niche products in dominant consumption infrastructures. In this sense, design should not just address the 'What to buy?' as approached by prior HCI prototypes, but also the 'Where to buy?' as they are closely entangled. Reflecting on the dominant in-store [265, 64, 299] and motivational approaches [304, 702, 299], our research shows how transparent and personalized food RecSys support practices beyond single retailers and shopping trips by addressing the complexity of finding and establishing appropriate products and filtering information, in short, strengthen and protect critical consumers.

9.6.2 From Ethical Profiles to Practice-based Profiles with Ethics

'Where to buy?' was just one aspect that raised our awareness for the embedding of our approach into the context of practices. Furthermore, the pre-study showed how consumers, besides ethical values, need to consider a variety of factors, such as the nearby stores and their available products, their family, their budget, and allergic restrictions, to find the 'right' product. Although current food RecSys start (re-)exploring the capabilities of knowledge-based [436] and ethical approaches [303] to more holistically consider consumers, supporting critical consumerism is still a major gap in the RecSys literature [641, 303]. Especially as the profiles only focus on narrow aspects such as health [436] or sustainability [303] rather than consumption practices in their diversity and complexity.

Here, our research contributes by presenting the design of a personalized

food RecSys that allows for a flexible setup of a profile that considers personal goals, altruistic values, allergic restrictions, and personal situation e.g. through the selection of nearby markets and the option to search for products that make trade-offs between budget and ethical values. To the best of our knowledge, this is the first food RecSys with a personalized profile including ethics [303] and practice-oriented options. On the one hand, this is in line with research [265, 303] on sustainable in-store assessment that already emphasizes flexible definition and weighting of values, on the other hand, it enhances the implications by more strongly highlighting to account for including practices and their context rather than ethics and sustainability, only.

9.6.3 Different Complexities of Profiles

Moreover, our evaluation showed an emerging design space between personalized profiles that go beyond role models, the easy usage of such role models, and shared profiles and knowledge that is created by organizations and/or other users. This design space directly relates to the work of Herbig et al. [265] that also found a certain need for democratization of profiles. While for some consumers such outsourcing of the profile to trusted institutions and role models might make it easier to set up the profile, others even demand more detailed options for personalization. For those consumers, role models, such as 'vegan' provide a general direction but do not allow for an adjustment towards their own consumption patterns.

Against this background, our results sensitize us for the requirement of a flexible profile that makes use of both institutionalized and shared profiles and rules but also room for in deep personalization that allows for modification or creation of rules and subrules, e.g. incorporating own definitions of what veganism is about. Here, further engaging in end-user development [373] of such RecSys could be an interesting perspective to engage in. This resembles the heterogeneity of the critical consumer group [224] and that they can not be simply subsumed under umbrella terms such as sustainable consumers or even more detailed terms such as political consumers in terms of design.

9.6.4 Towards Open and Reliable Data and Algorithms

Although not in the scope of our prototype, the evaluation showed how reliable and open data and algorithms are important for trust and appropriation of ICT to support critical consumer practices. In particular, in the sensitive field of consumption, consumers demand certain reliability, especially if they have certain restrictions or take their ethical choices seriously. Interestingly, it was not always the crowd-sourced data from Open Food Facts but data from the participating markets themselves which led to wrong recommendations. Thus, we want to highlight the importance of thinking beyond the scope of technology design but demand a policy that forces producers to pass reliable data to consumers such that they can make reliable decisions as already called for by Stevens et al. [605].

From an algorithm perspective, consumer protection organizations already criticized that even supposedly consumer-friendly third parties do not have transparent algorithms and it remains unclear how they recommend items [150, 303]. Also, our results show how consumers want to understand the food knowledge in detail and how the system creates the recommendation. This is important if the profiler makes use of outsourced food knowledge. Therefore, future designs should not only use open and reliable data but also open and transparent algorithms that allow to understand why a certain product matches the profile.

9.6.5 Critical Reflection on the Scope of Recommender Systems

Besides the presented implications for designing RecSys to support critical consumer practices, we want to use the space to reflect on the approach itself. In line with the critique on the neo-liberal agenda that burdens consumers with the transformation of the society towards sustainability [45, 224], our system does not account for the responsibility of other agents such as producers or the government. However, the urgent need to quickly reduce our environmental impact from our perspective needs a multifaceted strategy that exercises power towards policymakers and strengthens consumers that already attempt to make sustainable consumption decisions.

Here, the scope is limited when it comes to motivating consumers. However, it addresses the difficulties of motivated consumers, as outlined by prior research [350], in the sense of supporting the exploration of products and consumption infrastructures. Addressing those issues appropriately contributes to an ongoing practice transition rather than leaving consumers alone with the difficulties of practice change once they are motivated.

Moreover, thinking such systems from a broader perspective that goes along with open data and algorithms might contribute to more consumer protection and power towards the producers that nowadays benefit from informational asymmetries [303]. In particular, the increased awareness about consumption infrastructures could help with the articulation of requirements towards producers and policy-makers [419].

9.7 Conclusion

The literature- and the pre-study sensitized us to various aspects of the design of a RecSys for critical consumers. We developed a prototype that allows for setting up a profile that not just considers the complexity and diversity of critical consumer practices but also tries to incorporate the context of practices by inducing holistic aspects of diet restrictions, personal goals, markets, and budgets. From our evaluation, we found how such an approach addresses typical issues of critical consumers that have difficulty finding appropriate products within the diversity of products and retailers. Still, our implications needed to be revised to some extent after the evaluation as, for example, planning did not turn out to be the practice where we can support consumers, but rather we should focus on exploration and reflection as a separate activity. In addition, future research and design should not only use complex profiles to personalize 'what is the right thing' but allow for considering the personal situation. Furthermore, profiles need to be understood as having different complexities for different consumers while some find it appropriate to user role models such as 'vegan' or use outsource knowledge to institutions, others require deep personalization and modification. Lastly, especially when outsourcing, reliable and open data and algorithms are needed to strengthen the position of consumers.

Overall, the results contribute by designing and evaluating one of the first (see [303]), personalized food RecSys that brings in critical consumer values and by reporting on the requirements, appropriation, and scope of food RecSys beyond stimulating consumption as a means to support consumer protection and grant access to distributed and complex consumption infrastructures and products.

Given our definition of critical consumption and the related participant sampling, our results are clearly limited to a Western context, where, e.g. veganism is not the norm but a critical minority leading the way towards the important reduction of animal production consumption [300]. In other cultural contexts, low consumption of animal products might be the norm based on other cultural values [530] or poverty. Although our prototype accounts for price-value trade-offs, precarity is a factor that needs more consideration, even in Western contexts, as not all people can afford products that are in line with their values or can only afford environmentally sustainable products, which then finds no acknowledgment in the sense of critical sacrifice. In this sense, the theoretical lens itself can be seen as a privileged Western perspective [105].

While the results clearly support the approach of a RecSys to support exploration and reflection, a long-term study with more participants would be necessary to understand the influence on consumption practices. In addition, the focus of this study should be on consumers that have not already established some set of products and explored markets, but consumers that are in a more early phase of their practice transition. Moreover, our research touches various topics such as shared, personalized, or institutionalized food knowledge that was not implemented in our prototype. Here, more research is needed on the collaborative or End-User Development aspects of how one can make its own rules.

10 Trust versus Privacy: Using Connected Car Data in Peer-to-Peer Carsharing

Abstract

Trust is the lubricant of the sharing economy. This is true especially in peer-to-peer carsharing, in which one leaves a highly valuable good to a stranger in the hope of getting it back unscathed. Nowadays, ratings of other users are major mechanisms for establishing trust. To foster uptake of peer-to-peer carsharing, connected car technology opens new possibilities to support trust-building, e.g., by adding driving behavior statistics to users' profiles. However, collecting such data intrudes into rentees' privacy. To explore the tension between the need for trust and privacy demands, we conducted three focus group and eight individual interviews. Our results show that connected car technologies can increase trust for car owners and rentees not only before but also during and after rentals. The design of such systems must allow a differentiation between information in terms of type, the context, and the negotiability of information disclosure.

10.1 Introduction

The sharing economy market has grown strongly in recent years [494, 704]. Peer-to-peer platforms (short: P2P platforms) such as AirBnB, UBER, Getaround or BlaBlaCar enable private individuals to make goods or a service (e.g., a room, a trip, or a car) available to someone else for a certain period of time, generally for a fee. The concept behind P2P carsharing platforms is different from sharing within the family in that the other person is not usually known personally. Accordingly, for P2P platforms to succeed, trust plays a central role [52, 84, 259, 469]. In this vein, trust is also sometimes referred to as the currency or lubricant of the sharing economy as it is an efficient way of reducing transaction costs in social exchanges [630].

There is no consistent definition of trust, but it is generally understood as a multidimensional, socio-psychological construct [259, 630]. For example,

Hawlitschek et al. [259] understand trust as the expectation and obligation that an exchange will take place in the future. In this paper, we follow the definition of Huurne et al.:

> "[Trust is] the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party." [259]

When trust towards an unknown person is being established, reputation systems provide an important mechanism [188, 630]. However, on P2P sharing platforms those people are often not completely unknown as some personal data has often already been disclosed. For example, typically, users (both car owners and rentees) have a profile page with photos and a text-based self-description. In addition, certain details are frequently checked in advance by the platform operator (e.g., an identity verification on AirBnB or driver license verification on carsharing platforms). Moreover, offering services or personal goods is often implicitly connected with disclosing personal, partly intimate information. For example, on AirBnB the owner of an apartment provides an unknown person access to his or her own home – which is often regarded as the epitome of privacy [592]. In carpooling, one sits together with a stranger in a confined space [194].

In P2P carsharing, people give strangers their cars, which might hold not only an economic but also an emotional value. This value can potentially result in concrete ideas about how to handle it [214] and is one of the main issues why people don't want to share their private cars with strangers [40, 56, 258, 379, 466, 631]. Thus trust is particularly important for carsharing as it entails potentially risky and insecure situations where parties are interdependent [412]. Cars are comparatively expensive goods to which many owners build emotional connections. Therefore, the demand for building trust is higher compared to many other goods. At the same time, cars are constantly being affected by new technologies. The emerging possibilities raise novel privacy questions we want to address in our work.

By providing car owners with information that enables them to track e.g., the

car condition and rentees' driving behavior better, connected car technology may bring new potential for supporting trust. This step, however, would also introduce new challenges in terms of privacy. In this respect, a tension arises between car owners and rentees. This study is therefore conducted from these two different perspectives.

Work has been done on building trust in the sharing economy [82, 188, 259, 630, 301, 345, 392, 604, 629], but there seems to be a lack of what specific information people need to gain trust in P2P carsharing and what information rentees do not want to disclose to maintain privacy. Within the scope of this study, we therefore address the following questions:

- What information do car owners need to build trust in order to share their car?
- What information are rentees willing to disclose to get a car and maintain privacy at the same time?
- How do people classify information from the perspectives of car owners and rentees?
- In P2P carsharing, is it possible to build trust without invading privacy?

By investigating these questions, we contribute to a growing body of literature regarding the relevance of building trust in the sharing economy, taking the different views of both parties, car owners and car rentees, into account.

10.2 Related Work

A relatively new development in carsharing is the growth of P2P carsharing platforms in Europe such as Turo, SnappCar, or Getaround, which allow users to rent cars to each other. In contrast to the local context of traditional carsharing, the peers usually do not know each other personally, giving rise to a challenge known in the contexts of e-commerce and online auctions: Building trust among strangers via the Internet [123, 145, 541]. Over the last decades, reputation systems have been adopted as a prevalent mechanism where, by

leaving comments and rating others, users share their experiences on a previous exchange with unknown people [188, 500, 360]. Also, as research in HCI has shown, additional information such as name, images [188], and personal preferences or number of friends and profile views [345, 392] provided by users themselves to build a profile contribute significantly to signaling their trustworthiness to others. Furthermore, both types of information, ratings from others and self-generated information, serve as surrogates to build trust online [278, 301].

Trust systems are linked to the topics of privacy and trust in two ways. Users hope to create trust by disclosing personal data [604, 632]. Disclosing such personal information is in the interest of users as it fosters building a good reputation and may increase perceived trustworthiness and thus future transaction opportunities [323]. At the same time both car owner and rentee need to trust each other and the P2P platform that information provided will not be misused by any stakeholder [601]. Structurally, in the sense of multilateral security [485, 484], P2P platforms constitute a phenomenon of multilateral trust in which the interests of all actors must be taken into account [606].

According to Resnick and Zeckhauser [513] trust systems must 1) provide information that makes it possible to distinguish between trustworthy and untrustworthy actors, 2) promote integrity and trustworthiness and 3) deter untrustworthy people. Nowadays, data from connected car technology could serve to build trust by e.g., monitoring driving behavior and including it in calculating trust scores. Such new sources of information raise new questions regarding the potential for supporting the willingness to share goods. In addition, they potentially touch on privacy issues as users must balance the perceived costs of disclosing the privacy with the perceived benefits of a better reputation.

In the literature, a connected car is described as a vehicle that is connected to the Internet; can communicate with other devices, cars, and road infrastructures; and can collect real-time data from multiple sources [144, 238, 357, 424, 603, 618]. It is a part of the Internet of Things (IoT) discussion and will play a central role in the automotive industry [144]. Swan describes this development as a movement from quantified self to quantified car. She explored

how quantified self sensors can be connected with automotive sensors to create new applications to e.g., quantify driving behavior, real-time assistance for parking and accidents, or enable stress-reduction functions [618]. Major vehicle manufacturers or auto parts supplier are already storing connected car data. With the eCall, e.g., in Europe, new cars have to be connected by default, pushing the technology into the market [111, 290].

Carsharing providers do not currently use connected car data to build trust in communities. In the insurance industry, however, telematic tariffs are the first approaches towards connected car data-based reputation systems. There, data is collected automatically while the car is in use in order to differentiate between trustworthy, safe drivers and non-safe drivers. This data can then be used to promote defensive driving, by offering the "good" drivers cost reductions [238, 432, 677, 701]. Health insurance companies are also using digital behavioral data – so far on a voluntary basis – to develop individualized tariffs [369, 418, 681].

Currently it remains unclear to what extent such approaches can be transferred to promote trust in the sharing economy. Of importance is what information is suited to building trust and what is perceived as privacy sensitive [162, 357, 671]. While privacy may be particularly of concern for rentees, some information might be demanded by car owners. Therefore, openly exploring potentials calls for looking at both sides.

There are several approaches in literature dealing with the privacy problem in the connected car. However, studies typically take the perspective of car owners managing data disclosure to manufacturers or third-party companies. For example, Walter et al. developed a system based on the privacy calculus model [694, 167], increasing transparency and helping car passengers to control their privacy towards the car owner in a self-determined manner [671]. Derikx et al. studied whether and how passengers' privacy concerns towards the car rental can be compensated by offering monetary benefits [162]. Arguing that the breadth and depth of personal data from connected cars goes beyond already available data from mobile devices, Lawson et al. go a step further by discussing general principles and data protection standards to be enforced via regulations [357]. Currently, the specific potentials and needs

for protection of connected car data in a P2P carsharing context are not fully understood.

10.3 Methodology

This study aimed to investigate how new possibilities of computational trust mechanisms were perceived and discussed by study participants, what sensors or information might be relevant, what information they would be interested in if they were a car owner, and where they had privacy concerns from a perspective as a car rentee. To answer these questions, we conducted three focus group and eight individual interviews. During the focus group interviews, participants tended to adapt their position to the group accordingly. We therefore decided to combine the strengths of both methods to compensate the weaknesses [502, 589, 625, 635].

The focus groups were therefore used to obtain a broad spectrum of ideas and views on the use of connected car technology for carsharing and to obtain initial information on social norms for using data. The focus groups concentrated in particular on individual information and protection needs of the participants, as well as the principles that could be derived for the design of IoT-based trust systems. To improve the analyzability, the interviews were recorded. Following the transcription process, a qualitative content analysis in line with Mayring was performed [408].

Initially, a focus group interview was organized with researchers and professionals from the mobility area. To increase the diversity of the results, we also invited a student focus group and voluntary participants who had contacted us as a result of a local mailer. With the help of the individual interviews, additional insights were gained into individual points of view while group opinions dominated in the focus groups.

Each focus group participant was asked to imagine data sources and connected car information and to write these on sticky notes. Subsequently, their task was to comment on each note by assessing whether the information was important from the car owner perspective and critical from the car rentee perspective. This process was supervised by asking for more detailed information

	Focusgroup Interviews															Individual Interviews							
ID	F11	F12	F13	F14	F15	F16	F21	F22	F23	F24	F25	F26	F31	F32	F33	P1	P2	P3	P4	P5	P6	P7	P8
Gender	M	F	M	F	F	M	M	M	M	M	M	F	M	M	M	F	M	M	M	M	M	F	M
Age	35	35	48	32	34	30	32	20	21	20	23	23	53	56	28	26	20	30	35	41	41	62	65
Category	Experts						Students						Volunteers			Potential Car Rentees				Potential Car Owners			

Table 9: Participants

and promoting group discussions. The results were then used as inspiration cards in the individual interviews.

Because privacy and information disclosure may conflict in IoT-based trust systems, this study took perspectives of both rentees and car owners into account. As a first step, the new possibilities for data collection in cars were explained. Each participant then divulged from the perspective of a car owner what information about the car rentee or the journey s/he would potentially be interested in. In the second round, each participant then took the perspective of a rentee and explained what information s/he would be ready to disclose. In both rounds, the participants were stimulated to think out loud and explain the respective motives more exactly.

Twenty-three people participated in the study, (average age 35 years), of which 26% were female and 74% male. The high male share in the study is justified due to the higher interest of men in carsharing and the sample corresponds to the average carsharing user in Germany [306, 557]. Table 9 shows the participants of the focus groups and the individual interviews. The first focus group consisted of 6 experts (3 women, 3 men, 20-48 years), researching digitization and new mobility concepts. In contrast, the second focus group consisted of 6 students (1 woman, 5 men, 20-32 years) who participated in this study because of their interest in innovative sharing concepts. The third focus group consisted of 3 volunteers (3 men, 28-47 years) who were highly interested in P2P carsharing. Each focus group interview lasted about 90 minutes and was conducted in Germany, in German. To allow a better assignment of the results to the specific participants, we used an abbreviation logic that can be traced back to a unique person.

For example, F32 means that the statement is from participant no. 2 from focus group no. 3.

One weakness of focus groups arises from the possible dominance of individual participants and the fact that only socially accepted opinions might be expressed [589]; rentees or car owners would possibly not express their real opinions to avoid conflicts. To compensate this weakness, we conducted additional eight problem-centered individual interviews with car owners or potential rentees (see Table 9). The transcripts from the focus groups were initially coded and analyzed to identify first categories that could be used to set up the interview guideline for the individual interviews. The coding categories were, therefore, inductively derived from the focus groups, forming the basis for the individual interviews.

The participants in the individual interviews only took one perspective to avoid adopting another perspective and thus distorting the result. First, the participants' attitudes and usage behaviors towards the car and carsharing were asked. They were also asked whether they had already lent or borrowed a car once and what the context was. These questions enabled us to gain insights into participant experience with carsharing. Cards were prepared with information examples from the focus groups, and these cards were then discussed individually with the participants, who were asked to comment on them. Each individual interview lasted about 60 minutes and was also conducted in Germany, in German. Here, too, all interviews were transcribed and analyzed in line with Mayring [407] [34]. All authors coded the interview transcriptions independent of each other and then combined their analyses to form a joint final set of categories.

10.4 Findings

Through the transcript coding, patterns and argumentation approaches were identified and used as categories for the findings. These include the different *types of information*, the *classification of information,* the *situatedness of information*, and the *negotiability of information*. In this section, we explain how these categories arose and demonstrate the findings using anchor examples from the conducted interviews.

Type	Examples
People	Driver identification, Driving experience, Ratings, Creditworthiness, Number of passengers, Alcohol measurement, Microphone
State of the Car	Oil level, Coolant, Brake discs, Tank level, Shock absorbers, Order and hygiene, On-board electronic messages, Loading weight, Trunk camera, Engine noise, Mileage, Media use
Driving Behavior	Driving style, Speed, Brakeing behavior, Acceleration, Distance, Consumption, Steering behavior, Lane change, Clutch, Parking behavior, Dashcam, Kilometres travelled
Special Events (point of time)	Accident, Border crossing, Speeding, Parking damage, Fueled the right fuel?, Clocked?, From whom is my car opened?, Crash record, Rain sensor
Location (time period)	GPS sensor, GPS route, Does the rentee stick to the agreed area / purpose?
Misc	Aspects of contract law (individual regulations between car owner and rentee), copies of relevant documents, legitimacy of contract, claims for damages, declaration of consent /contract

Table 10: Examples of Information Types

10.4.1 Types of Information

First, the interviews showed that the pervasiveness of sensor technology in connected cars was not clear to participants. Second, participants typically interpreted collected data in contexts meaningful to them. Only then did data become meaningful information. Accordingly, the data most relevant for the interviewees were usually those that they were able to use for gaining relevant information for sharing a car and its context in assessing trustworthiness (e.g., driving behavior for evaluation and decision of re-sharing).

In this regard, we identified six different types of information that were most important to participants: *People, State of the Car, Driving Behavior, Special*

Events, Location and *Misc* (see Table 10 for related examples). The category *People* contains information about the driver or the passengers (e.g., identification of the driver). Analogously, *State of the Car* contains information about the condition of the car, e.g., whether it is in an operational condition (oil, brakes, shock absorbers, onboard electronics, etc.). The third category *Driving Behavior* combines People and Car by including information about the use of the car from which conclusions can be drawn about both the driver (e.g., safe or risky driving style) and the car (e.g., heavy wear and tear due to bad driving behavior). This category also includes the recording of driving data, which should help to reconstruct the process in the event of a dispute (e.g., recording of abrupt braking).

The fourth category *Special Events* represents situation-relevant information on the occurrence of certain events (e.g., notification in the event of an accident). While the special events are mainly related to a point in time, the fifth category *Location* is to be regarded as a period of time. It provides information on the location of the vehicle – in particular with regard to compliance with agreed limits (e.g., border crossing, accuracy of tracking) and for other purposes. *Misc* includes various information on e.g., the contractual design (e.g., individual rules among each other) as well as other information not attributable to the previous categories and less often stated by participants.

The information types were derived from information needs of car owners and can be applied to both user groups. While car owners can determine the degree of importance of specific information, rentees express the willingness to disclose it. These categories formed the foundation on which the participants classified when reflecting on the relevance or sensitivity of the data or information from the car owner's or rentee's point of view.

10.4.2 Classification of Information

Various expressions in our interviews were used by the participants to classify the relevance of the data from the point of view of the car owner or the willingness to expose it from the rentee's point of view. The classification was usually also justified with a purpose, which shows the motivation behind the information.

Car owner		Car rentee	
Level	**Quotes for classification**	**Level**	**Quotes for classification**
important	"That would be always important for me, I'd like to follow that permanently, yes." (F22) "This is a basic requirement […]" (P8)	rejection	"[...] that is out of the question." (P1) "[...] no, not at all, not even [...] that's not appropriate, that's one step too far." (F32)
conditional interest	"As long as it moves in certain parameters, this is not interesting for me, but if it moves beyond that, it becomes relevant for me or can become relevant for me." (F11)	conditional disclosure	"[...] in my opinion has no added value [for the car owner], as long as [...]. Unless [...]..." (P4)
unimportant	"I don't want to know what's being said or anything, I don't care." (F14) "[...] I don't really need it at all. [...] it's uninteresting to me." (P5)	disclosure	"Yes, I would say he can have access to it. Would also be interesting to know [...]. I find it useful and understandable if he [the car owner] has access to it." (P2)

Table 11: Examples for the Classification of Information from the Perspective of Car Owners and Rentees

10.4.2.1 The car owners' view From the car owner's perspective, the following three categories can be derived from the chosen formulations: *unimportant, conditional interest* and *important*. Sample excerpts from the interviews to define the respective level are shown in the left part of Table 11.

The category *unimportant* was assigned to data or sensors that were not relevant for the car owners, offer them no added value and thus, in most cases, did not serve as a source of useful information, as P5 argues:

> "Trunk camera, what do you need something like that for? [...] For me that's not interesting information and I can't see any purpose in it either." (P5)

Data was unimportant for the car owners if they could check it afterwards without difficulty (e.g., smoking in the car), if they could not classify it (e.g., trunk camera) or if it represented a too strong interference in the privacy of the rentee (e.g., microphone):

> "I don't want to hear what they're saying. That is already surveillance. I wouldn't like that myself either." (P6)

Noticeably, when classifying information, the car owner also often took the perspective of the rentees to anticipate potential privacy violations. Information of the category *important* was necessary for the car owners or at least aroused a general interest, as part of the process of assessing trustworthiness,

since it could satisfy specific needs and would support the decision process in many cases. Such information especially included data leading to perceived safety-relevant aspects and creating potential personal and legal advantages. Above all, the GPS position was mentioned in this regard:

> ”I find GPS position very important that you know where the car is at all times. Simply for safety.” (P5)

Similarly, the identification and authentication of the driver were often associated with a sense of security regarding theft and handling of the car:

> ”Important for me is also regarding the contract and legal data or theft, a kind of face recognition: Is the rentee really the driver?” (F21)

Furthermore, many respondents also found the driving behavior to be very important, mainly for assessing the rentee and thus to avoid possible disadvantages in the future.

> ”[Driving behavior] would be exciting for me, [...] to perhaps be able to say afterwards that I no longer want to share my car with this user. Because he brakes too abruptly or something like that.” (F11)

With *conditional interest*, the participants often argued for including data that were not always and inevitably relevant but had an informative character, mainly in certain situations, or were just interesting and not of great necessity to the car owner. The negotiable sources of information mainly consisted of information that simplifies the carsharing process by making it no longer necessary to carry out checks (e.g., state of the car when the car is returned) and information that is available in the case of certain events (e.g., accident, parking damage, etc.) and contributes to solving those issues. This information can preserve evidence and thus aid in obtaining legal advantages.

> ”Finally, I just want a report when [the rentee] returns the car to me. Clutch, engine, all things concerning the car that are abnormal, I want to be told. To check after the rental, so to speak.” (P7)

> ”[...] if an accident occurs, to know [by an accident sensor] who caused the accident and thus to be able to settle questions of compensation.” (F25)

10.4.2.2 The car rentees' view From the rentee's perspective, three levels can also be distinguished: *disclosure, conditional disclosure* and *rejection*. Examples of allocation into these categories can be seen on the right-hand side of Table 11.

The *disclosure* category includes data that rentees consider non-critical and that could not be used to generate personally identifiable information. Often the disclosure was justified here, usually by taking on the car owner's view. Rentees generally released data in which they could not identify any personal reference (e.g., weather) or through which they could obtain advantages, either individual (e.g., better driver service) or legal (e.g., legal security). However, even in some specific cases, they showed an understanding for and were willing to release certain data with justified interest on the part of the car owner.

Information in the *rejection* category is not disclosed under any circumstances. There was a strict refusal and no understanding of the necessity or relevance of disclosure. We mainly experienced rejection for the disclosure of personal information when the participants felt monitored by the car owner (e.g., microphone, indoor camera) without meaning or perceived added value for themselves and/or the car owner.

> ”[Microphone] is a no-go, also for legal reasons, especially in a car when you think that you can have a conversation in a private room, and someone listens to it.” (F32)

> ”You can find out a lot about the individual user [...] if you have

> the route. I don't see any added value in having a detailed route, so I wouldn't provide this information either." (P2)

Conditional disclosure refers to data in which the respondents saw a sense or comprehensible reason in the use. The disclosure mostly takes places under special conditions (e.g., occurrence of an accident) and often when adopting the car owner's view, understanding why the information was needed. Some participants generally replied that they are only willing to share information once they had added value from it, such as added value to comfort by receiving a report on their own driving behavior and getting advice on how to improve their driving skills.

> "Whether the vehicle and the windows are locked, I find [also from rentee's point of view] very interesting for the protection for both sides." (F15)

In the case of personal data, conditional disclosure only took place for special events (e.g., driver identification via camera in the event of an accident) and with an underlying meaning or an individually comprehensible request from the car owner.

> "In the event of a dispute, data can be used to create clarity. [...], such as driver identification, in the case of personal rental and, for example, in the event of an accident." (F16)

Also, when disclosing information was advantageous for the rentee (e.g., economic advantages) or mutually advantageous (e.g., legal advantages), rentees often considered accepting disclosure as it was in their own interest, too:

> "Assuming I'm not guilty, you should consider making [the recorded crash] directly available to the car owner." (P2)

However, as P2 states, willingness to disclose information in this case depended on the desire of being free of guilt, which, if disclosure was to be made part of a contract, could obviously not be ensured as setting up the contract would be prior to any possible incident and thus to any question of guilt.

10.4.3 Situatedness of Information Needs

When rentees argued for or against a piece of information, they often added conditions under which they were either willing to disclose or not disclose that data. Our analysis revealed several dimensions in the context of conditional interest on the car owner side and conditional disclosure on the car rentee side. Whether data was interesting for the car owner or not often depended strongly on the situation. Many pieces of information (e.g., driving behavior) played a role before the contract to be able to assess with whom the car should be shared in the first place. During the rental, most of the data would only be of importance to the car owners were an exceedance or emergency to occur (e.g., rain sensor, loading weight, crash sensor), as F11 states:

> ”As long as it remains within certain parameters, this is not interesting for me, but if it moves beyond that, it becomes relevant for me or can become relevant for me, depending on which data it is.” (F11)

Setting corridors similar to geo-fencing was a major theme for such data. On a similar note, some data were only important in the case of extraordinary events, e.g., to recreate context or driving behavior in the case of an accident:

> ”[The rain sensor is] a one-time piece of information that only helps me when an accident or something has happened, as evidence.” (P8)

From the rentees’ perspective, certain kinds of data, such as door lock status, were similarly perceived as a way to prove conditions or behavior.

> ”[Checking if windows and doors are closed:] So, if I can free myself from an accusation, it would be ok, but just to satisfy his curiosity it would be a no-go.” (P1)

The quote highlights the relevance of the purpose for which data are used to assess the appropriateness of disclosure. This background information helped users weigh the interests of transparency versus benefits for both perspectives.

Purpose	Example
Trust-building	Evaluations, Creditworthiness, Driver identification, Driving experience
Prevention	Alcohol measurement, Driving style, Lane change
Detection of abnormal issues and events beyond the norm	Braking behavior, Speed, Acceleration, Rpm, Loading weight, On-board electronic messages, engine noise, Border crossing, Accident, Parking damage
Legal	Individual regulations between car owner and rentee, Clocked?, Dashcam
Monitoring	GPS sensor, GPS route, Does the rentee stick to the agreed area / purpose?

Table 12: Examples of Information Purposes

> ”Everything that is used to check the rental conditions is okay from my point of view. As, for example, the driver identification, [...]. Likewise, GPS tracking to keep agreed distances. Real-time tracking is undesirable for me. I’m only willing to provide the sensor data if I feel that the added value for the car owner is higher than the restriction of my privacy.” (F16)

Based on our participants’ statements, the use of information can be classified into three temporal levels: *before, during* and *after* rental. In addition, the situational statements can be used to determine certain information purposes, which are taken up within the following temporal gradations and are exemplarily shown in Table 12.

10.4.3.1 Before rental Before renting, car owners were mainly interested in information for decision-making regarding the rentee (e.g., driv-ing style). Financial interests also came into play (e.g., solvency). In this phase, information was primarily used for prevention and trust-building (e.g., evaluations) as well as for legal purposes (e.g., individual regulations).

> ”As a car owner, I wouldn’t have to think for myself whether this

> is a good driver; a value or a grade for the respective rentee might be interesting [to then decide whether to share my car]." (F12)

The rentees were able to identify themselves with the car owners in fundamental questions regarding the security needs and thus were often willing to disclose related data. However, there was also a fine line to rejection when participants were unable to comprehend reasons for data transfer or whether it represented a strong perceived violation of privacy (e.g., checking creditworthiness, microphone).

10.4.3.2 During rental During the rental period, information about certain events was of particular relevance to the car owners. In addition to legal and preventive purposes, monitoring and detection of abnormal issues and events beyond a perceived norm also played a role. For example, it was important for a car owner to be informed about speeding, disctance keeping, tailgating, or accidents as a preventive measure. A potential rentee commented on information before the start of driving on the controversial subject of facial recognition. For him, it was legitimate for a one-time authentication to address the car owners' needs for security:

> "The owner may want to use this against abuse. That, whatever, once at the start of the trip or if one starts the engine, that the face is identified." (P2)

In case of an event, the car owners wanted to collect evidence for legal purposes (e.g., in the event of an accident or parking damage via a dashcam). With some car owners, this even turned into a constant desire for monitoring.

> "[The recording of the accident on a dashcam is important] for me to have proof that the driver is to blame or not. Well, that's easier to clarify then." (P6)

After the occurrence of an event, the car owners were also concerned with clarifying the facts as well as preserving evidence and ad-hoc notifications so that they could react directly in the case of events beyond the norm.

> "[...], so for me it would also be important whether [the rentee] was caught by speed trap, [...]. So that I don't find out later." (F26)

Most rentees reacted with understanding and offered to provide information under certain circumstances.

> "[A parking damage sensor] I'd say for insurance, to prove afterwards the innocence, why not?" (P1)

With regard to a detection of abnormal issues and events beyond the norm, some car owners wanted to be informed about the load weight to check whether the permissible limit value was exceeded.

> "Yes, that's important, it goes on the shock absorbers and wears the car out. I would like to have the information if rentees load my car too heavily." (P8)

The rentees made only a few comments about disclosure during a situation; overall, "live data" was perceived as monitoring and mostly rejected. The participants' stances became most clear in the case of video monitoring but could generally be extended to live data overall:

> "[interior camera:] I believe it's not necessary for [the car owner] to have such a strong intervention in my privacy, which a live feed from the car would be." (P2)

Furthermore, spatial-temporal information (GPS) also played an important role for car owners during the rental period and was often associated with a sense of security regarding theft. While some owners wanted a live view on their cars' GPS, they also were able to anticipate the rentee's view, which was more differentiated. Their acceptance depended on manifold factors such as the purpose the data would be used for. Here, rentees often differentiated between disclosing a position or a route. Whereas disclosing the position (especially in the case of an event, e.g., an accident) was sometimes accepted, providing the route was largely rejected for privacy reasons:

> ”Basically, I would have a problem with someone knowing exactly where I am at the moment. But if only the region was known, for example CITY1, CITY2, [...] then I wouldn’t have a problem with that.” (F11)

10.4.3.3 After rental After the rental period, the car owners were particularly interested in information to simplify further processing (e.g., a vehicle condition report), which aimed to detect abnormal issues, prevent abuse and build trust.

> ”I just want to have a report: This and that happened. In the end, when [the rentee] gives me back the car.” (P7)

The rentees were mostly willing to disclose the required information as they thought it would also give them an individual advantage if the car no longer had to be handed over in person.

> ”I think the fuel level sensor is okay. Also handy for the rentee to prove whether the car was full up or not.” (F12)

Overall, car owners had the greatest information demands during the rental. Before and after rental, trust-building and prevention aspects (for further rentals) were the main focus.

10.4.4 Negotiability of Information

As already mentioned, when classifying information, participants often responded with “that would be interesting if...” or “that would be okay if...” and, depending on the perspective, they stated conditions for the interest in or the disclosure of information. This intersection between the classification of car owners and rentees we called the area of negotiability, which is shown in Figure 19.

Analyzing the interviews, we found various conditions with different reasons behind the conditional information disclosure and conditional interest. Often

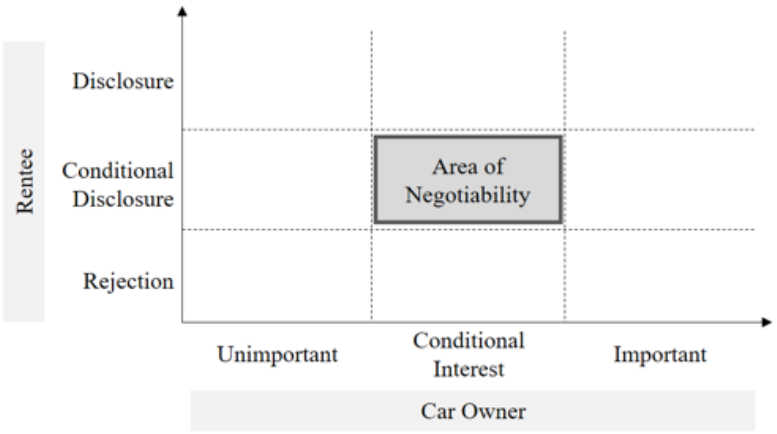


Figure 19: Area of Negotiability

participants argued that they would be willing to share certain information if it was not to be used for observation, which rentees perceive as a disadvantage. Some participants also argued that they would also be willing to share sensitive information as long as they had an advantage in doing so. This advantage could be either financial, legal, or at an individual level.

One participant from the expert focus group explicitly argued from the rentee's perspective that the data made available, e.g., in the event of a dispute, should only be available for 14 days and that all other data should be deleted after the car has been returned.

> "Time-based availability – delete route recording after the car has been returned. The relevant information should in case of dispute only be available for 14 days." (F13)

Both sides often argued that specific information should be tracked to use it as legal evidence. Some participants answered in general that they would only be willing to share information once they could see an added value for themselves such as, e.g., receiving reports about their own driving behavior and advice on how to improve it. One participant e.g., argued, that he was willing to share his location-based information for a discount:

> "The car owner can track my driving behavior, if I get a discount for it when I drive properly." (F21)

Altogether we can state that every car owner and rentee had an individual Area of Negotiability and specific conditions. For some it was enough to have transparency about what happens to their data and when their data will be deleted. Others do not care so much about their privacy if they are financially rewarded for disclosing private information.

10.5 Discussion

This qualitative study provides important information on how the increasing data collection in connected cars is viewed from the perspective of car owners and rentees, what kind of data collection is perceived as helpful or acceptable and why. In the following paragraphs, we discuss the main conspicuities from the interviews and the methodology.

The car owners usually said they would not lend their car to strangers. However, if people feel that their car is unscathed, if they get it back in good condition and if something has happened to ensure that they are not stuck with the costs, then the willingness on the part of the car owner to share the car increases. A comparison between car owners and rentees showed that car owners had a strong interest in data prior to the rental in order to generate legal advantages and avoid disadvantages. This interest is a typical example of information asymmetry prior to contract conclusion, where the car owner (principal) cannot assess the properties of the rentee (agent) in advance (hidden characteristics) [135, 595, 674].

It was also conspicuous that the group of experts, who were familiar with the technical aspects of P2P carsharing, did not consider information on driving behavior to be as relevant as those who were less familiar with those aspects. Monitoring the route were also perceived as important only in cases of infringements. Above all, rentees often expressed the condition that they did not want to be disadvantaged or, quite the opposite, they wanted to benefit from sharing their data.

Regarding the situatedness of information, it became apparent that, in addition to the temporal classification in before, during, and after rental, there were also distinctions in the argumentations between real-time transmission

of information (push notifications) and the manual retrieval of information before, during, or after rental (pull information). Some information needs real-time notifications, e.g., if the rentee travels beyond an agreed zone, while other information is sufficient simply on request, e.g., the tank level.

Another interesting argument with regard to privacy from the rentee perspective was that the sensitivity of individual information was often harmless as long as that information was isolated. When, however, different information about a rentee was aggregated across providers, so that very precise profiles could be created about individuals, that aggregated information can, in the wrong hands, certainly lead to problems. Being able to identify where exactly and at which time a person was and what this person did at this place could lead to social scoring and would be a strong invasion of privacy.

Altogether our findings show that privacy concerns depend on the appropriateness of purposes, which confirms previous findings of other privacy studies [290, 289, 357, 457]. However, the results do not apply to all settings in the sharing economy. They can particularly be transferred to valuable goods such as holiday apartments. Our findings show that information needs are strongly dependent on the context. As an example, the risk of a substantial financial loss (theft/accident) is a valid reason to demand information to build trust. In the next sections, we focus on the derived design principles.

10.5.1 Building Trust: Before, During and After Use

Several studies have shown how current reputation systems aim to build trust prior to use [188, 278, 301, 345, 392]. For car-sharing, IoT-based reputation systems have the ability to build trust through monitoring driving behavior. The data-based mechanism to support trust building also during and after the journey was positively received. The interviews showed that participants could quickly express basic needs, especially from the car-owner perspective, to acquire data during use (“Where is my car? Is it alright?”) and after use (“Are there any unusual incidents? Has it been refueled?”).

The traditional approach to building trust before use can thus be extended by information during and after use. Studies analyzing extended access control

[608, 609] came to similar results. However, when trust is defined as it has been in this paper, receiving that type of information (during and after use) means that the situation is no longer a question of creating trust since there is some kind of monitoring and control. Sensored data that has been recorded during and after use could be aggregated and published as a Driving Score on P2P carsharing platforms, as suggested by some of the participants, and thus serve as trust-building assist in later decision-making e.g., when accepting a rental request. Furthermore, a positive spillover effect seems plausible: knowledge of the better possibility of controlling the goods during and after use creates trust in advance and increases the willingness to share a car. However, this positive effect on the sharing economy must be understood and quantified more precisely.

10.5.2 Principle of Data Economy and Proportionality

From the rentee's perspective, it has become clearer that participants tend to prefer a data economical variant. But even from the car owner's point of view, the participants were not very interested in the details of what the rentee was doing in the car or in knowing where exactly someone was.

In this context, foreseeable privacy concerns of potential rentees were already considered so that excessive accuracy in, e.g., location monitoring was obsolete. From the rentee's perspective, the data in particular were harmless as they were already being recorded (e.g., mileage, fuel level, etc.) and support the car owner with the follow-up check through digital information. The other extreme was the camera in the interior of the car, which was classified as a strong invasion of privacy. Car owners showed little interest in too much information and only considered driver identification to be useful.

An important guideline should be the principle of proportionality. All participants acknowledged the car owners' interest in getting the car back in time and undamaged. The findings suggest that user acceptance increases when data collection serves this legitimate interest, which is perceived as proportionate and data economical. An example of this is the use of location information. Most participants perceived this information as a legitimate car owner need for safety, if s/he knows approximately where his or her car is at any mo-

ment, especially as s/he could then better assess whether the return would be delayed, and s/he might therefore have to reschedule. However, the exact location of the car was usually perceived as disproportionate, as it is none of the car owner's business who the rentee is going to visit. Some participants therefore suggested that not the exact, but an approximate location should be provided.

10.5.3 Principle of Reciprocity

Reciprocity is a universal social principle that refers to reciprocity in social actions [413]. Ridesharing or carpooling, e.g., implicitly expects that the gift of taking someone along is reciprocated in some way [420]. Our findings showed that the reciprocity principle is also applied in IoT-based trust systems on two levels: the sociological and the economical.

On the sociological level, reciprocity is about the mutual adoption of perspectives [413]. This mutuality plays an important role with regard to privacy [470] and assessing how one is seen and judged by others. Therefore, many participants wished to see how one is seen by the others through the respective information. Accordingly, the rentee should be able to get the view of the car owner. The possibility of taking on the other's perspective could help to improve the coordination in unforeseen events, e.g., to find the error in the event of a breakdown together with the car owner.

At the economic level, reciprocity of information asymmetries is discussed in how it can minimize the risk of opportunistic behavior and unfair transactions [19]. In consumer and data protection, this aspect is discussed in relation to exploiting information asymmetries in the context of dynamic prices. In our findings, this form of reciprocity was reflected, for example, in the fact that some participants found the recording of driving data acceptable if it could be used in a legal dispute about damages. However, this acceptance was linked to the condition that both parties could access the data equally and only if both parties had agreed to this beforehand. This protective measure was intended to prevent unilateral overreaching.

10.5.4 Principle of Transparency

In an automated evaluation such as an aggregated Trust Score, the adoption of a perspective also extends to non-human actors. One participant, e.g., suggests *"a value or a grade"* (F12) that reflects the driving behavior of the tenants in aggregated form. From privacy literature [671, 694], it is known that such scores need to be transparent for a user to be able to understand it. To understand the algorithm driving a Trust Score, the rentee concerned must put him- or herself in the position of the algorithm and be able to comprehend its analysis or evaluation. Scoring must therefore be transparent [607]. In the current consumer and data protection debate, this requirement is reflected in the demand for transparency for algorithmic decision-making systems. The interviews have shown that car owners and rentees want to understand how a driver evaluation works. Instead of, e.g., indicating a non-transparent driver evaluation, it seems reasonable for comprehensible metrics to be used, such as number of abrupt braking maneuvers or number of speeding maneuvers, and to inform the rentees about what measures can be taken to improve their evaluation.

10.5.5 Principle of Negotiation Support

As our findings show, certain areas of information are negotiable and therefore this area of negotiability could be considered in the design of an IoT-based trust system so that the respective actors can negotiate these areas independently. Because the area of negotiability differs from participant to participant, it is also not possible to integrate this negotiability in the initial design in a P2P carsharing system. As can be seen in the section on situatedness of information, many data disclosures are linked to conditions that must be fulfilled for the rentee to disclose the information. Therefore, it is important to make a system as flexible as possible so that rentees can specify their own preferences as to which data they are willing to disclose and which not. Furthermore, visible information on what data is collected and for what purposes could lead to wider acceptance [650]. It is also important to set individual conditions for the area of negotiability, e.g., that the rentee gets a financial

return if he or she discloses GPS data, or the GPS data are only disclosed if an agreed zone is exceeded.

10.6 Conclusion

Connected car technology brings new potentials but also challenges in P2P carsharing, especially for the question of promoting trust in the sharing economy. Our findings show that there is a privacy-trust-tradeoff in P2P carsharing since rentees are not willing to disclose all information. However, the interviews have shown that both parties can place themselves in each other's position: Car rentees are willing to disclose more information in exceptional cases (cf. F16) and car owners respect the rentees privacy to a certain degree (cf. P6). Altogether four core problems can be identified from the statements of the interviewed car owners:

- Car Condition: Is my car unscathed? Was there an accident? Will I get it back in good condition?
- Rentee's Trustworthiness: How does the rentee handle my car? What is his/her driving style? Does s/he take care of the car?
- Is the car used within the agreed conditions? Does the rentee use it within the agreed area and purpose?
- Do I have all the necessary data when something happens to the car to prove that I did nothing wrong?

Whereas the following three core statements can be identified for the rentees:

- Why does the car owner need this information? It's none of the lenders' business.
- If I have a benefit such as a reduced price, evidence that I am a good driver, or other added values, I am willing to disclose more information about my driving behavior.

- I am willing to disclose the information under agreed conditions that are transparent to me and relate to the security needs of the car owner.

The interviews have provided important insights into how general principles of data protection and multilateral security can be applied to the design of IoT-based trust systems. The overriding maxim should be to support the balancing of the different interests of the stakeholders. Conflicts due to excessive data collection and one-sided data use should be avoided. To this end, the principles of data economy and proportionality, transparency, reciprocity, and negotiation support should be applied. Regarding the limitations, a certain subjectivity already plays a role in the content analysis. The pool of respondents also influences the findings. Nevertheless, the results show that connected car technologies can increase trust not only before but also during and after rentals for both sides. In further studies, we would like to validate our findings by quantitative methods in a larger sample.

11 Trust-Building in Peer-to-Peer Carsharing: Design Case Study for Algorithm-Based Reputation Systems

Abstract

Peer-to-peer sharing platforms become increasingly important in the platform economy. From an HCI-perspective, this development is of high interest, as those platforms mediate between different users. Such mediation entails dealing with various social issues, e.g., building trust between peers online without any physical presence. Peer ratings have proven to be an important mechanism in this regard. At the same time, scoring via car telematics become more common for risk assessment by car insurances. Since user ratings face crucial problems such as fake or biased ratings, we conducted a design case study to determine whether algorithm-based scoring has the potential to improve trust-building in P2P-carsharing. We started with 16 problem-centered interviews to examine how people understand algorithm-based scoring, we co-designed an app with scored profiles, and finally evaluated it with 12 participants. Our findings show that scoring systems can support trust-building in P2P-carsharing and give insights how they should be designed.

11.1 Introduction

Over recent years, more and more consumers gave up individual ownership for demand-based access to goods [222, 317, 358, 455]. Accordingly, business models emerge, where consumers simply pay for on-demand access to goods [684, 466]. Benefiting from this development, carsharing gained public attention for several years now [434], with 15 million registered users worldwide in 2016 [570]. In addition to commercial carsharing, peer-to-peer (P2P) platforms such as Turo or Getaround allow private car owners to share their car with a previously unknown person [684].

On such platforms, transactions are mostly carried out online, which lowers the costs, but increases the anonymity. However, a lack of trust in other

”buyers” or ”sellers” is stated as one of the most frequent reasons for rejecting P2P sharing platforms [83, 469]. As a result, transparency, reputation, and trust are considered essential for the success of the P2P sharing economy [51, 84, 259]. Trust is typically increased by mechanisms, such as showing self-provided information (e.g. pictures or personal information) as well as peer-provided information (e.g. consumer reviews and ratings) [57]. Besides, Internet of Things (IoT) information could serve as a third trust-building mechanism [210, 249, 604, 669, 670, 602]. For instance, the driving behavior could be monitored by car telematics, which consists of onboard communication services and applications that communicate and record sensor information while driving. The information is then algorithmically processed into a score. Such insights into the driving behavior could increase trust into the driving skills of a possible rentee. So far, such information is used and analyzed by car insurance companies for risk assessment for individual insurance tariffs [417].

In the context of P2P carsharing, algorithm-based reputation systems could create an added value compared to peer ratings, since they are based on objective measures [602]. As the current standard among various reputations mechanisms in the sharing economy [259, 633], peer-ratings have also been criticized for various biases [179, 107, 638]. However, it is unclear if users would trust algorithm-based scores and if it would create an added value compared to peer ratings.

To close this research gap, this design case study [691] examines whether algorithm-based reputation systems have the potential to improve trust-building in P2P-carsharing. Therefore, we conducted a pre-study with 16 problem-centered interviews to find out how people understand algorithm-based scoring (section 3), we co-designed a P2P carsharing app prototype with an alorithm-based reputation system implemented (section 4), and finally evaluated it with 12 participants (section 5). Our findings show that algorithm-based reputation systems can support trust-building in P2P-carsharing and give insights how they should be designed. Our work contributes to the discourse of how algorithm-based reputations systems in P2P carsharing can be designed to support trust-building and decision-making for car owners and especially how the interaction with such systems should be designed to support a trustworthy

coordination of the user groups and reduce social biases and discrimination.

Our design case study builds primarily on the work of [83] and [602]. The study by [83] already focused on trust building and the willingness to share data from the rentee perspective and showed that the willingness to share data exists in the form of an area of negotiation. This paper puts a larger focus on the use case of driver selection based on an algorithm-based score, which is based on the data from rentees but focuses on the interaction with car owners. This is intended to address in particular the information assemetry for car owners and to support a better interaction with the algorithm-based scoring.

11.2 Related Work

11.2.1 Trust-building in Peer-to-Peer Carsharing

As a part of the Sharing Economy [516], the basic idea of Peer-to-Peer Carsharing is a joint consumption following the principle ”sharing rather than owning”. Products are not acquired by the consumer, they only get a temporary right of use a service or good — normally for a certain fee, e.g., based on the kilometers driven [51]. For many sharing economy participants, economic motivations — such as cost saving, reduced burden of ownership, or increase access to resources — play an important role [243]. Moreover, against the background of climate change, sharing vehicles becomes increasingly important [247]. According to a study by the Ford Motor Company [139] with over 10,000 participants, 55% of respondents in Europe would share their car for a fee, however, actual adoption is very low [83, 469]. Some studies in the context of P2P carsharing [684, 40, 368, 450] already address findings about user characteristics and user motivations. People participate in P2P carsharing for economic (reducing mobility and vehicle costs) and situational-practical reasons (availability, convenience, and flexibility) are mentioned [40, 450]. But there are also hurdles as sharing a private car via a P2P platform includes the effort of entering the availability of the car, arranging handover dates and the follow-up check to see whether the car has been damaged. In addition, there is the general fear of sharing a car with strangers [684, 571]. In particular, people often have a personal and emotional bond to their cars [214], which

increases fear of loss [53], because others might not treat the rented car with care, cause an accident, or return it late or dirty [83]. P2P carsharing platforms therefore play an important role, because they have a coordinating and trust-building function between the respective user groups and can digitally support cooperation between car owners and rentees.

Trust is generally acknowledged to be a multi-dimensional, socio-psychological construct [259, 630]. Our work relies on the following definition:

"[Trust is] the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party." [405]

Since such situations are typical for the sharing context, especially as most P2P transactions are executed online, trust is of particular importance in potentially risky and uncertain situations where parties are interdependent. Transparency, reputation, and trust are therefore seen as essential requirements [51, 84, 259], that reduce transaction costs efficiently in social exchanges [630]. However, it is difficult to build and sustain trust in online interactions [259, 426]. For this reason, trust mechanisms have been investigated in different areas, e.g., social media [517], online shopping [219], but also the sharing economy [259]. Regarding e-commerce, there are two dimensions of trust: trust in the seller and trust in the offered goods. In particular, buyers must trust in the integrity, quality and competence of the seller [219]. As products or services are usually offered by private individuals in the P2P sharing economy, users have to trust other peers, the platform, and the offered products and services [259].

11.2.2 Algorithm-based Reputation Systems via Car Telematics

To address the above mentioned lack of trust in online transactions, reputation systems emerged as a trust mechanism [630, 188]. The basic idea is that the parties rate each other, e.g., after concluding a transaction, and derive a trust or reputation score from the aggregated ratings. The resulting score can help other users to decide whether to interact with that party in the future or not.

Thus, reputation systems are incentivizing good behavior, and therefore tend to have a positive effect on market quality. Reputation systems are related to collaborative filtering systems [554] as they use the opinions of a community to help individuals more effectively identifying relevant content from a potentially overwhelming set of choices. Resnick and Zeckhauser [513] give a functional definition of a reputation system, as it must: (1) provide information that allows peers to distinguish between trustworthy and non-trustworthy peers, (2) encourage peers to be trustworthy, and (3) discourage participation from those who are not. We can distinguish three different types of information that contributes to trust building:

Self-provided information: Peers can provide information about themselves as well as the goods or services they offer. Repschläger et al. [511] pinpoint that personal attributes such as name, age, address as well as pictures serve as trust factors. However, Hanrahan et al. [250] observed that self-provided information can often lead to the problem of discrimination against peers based on specific characteristics, such as gender or photo. This problem shall be overcome by aggregated ratings [602].

Peer-provided information: As self-provided information is often suspected of being sugarcoated, peer ratings become increasingly common and serve as a substitute for word-of-mouth recommendations [709]. Consumer Ratings and Reviews are usually written to either recommend a product or service or to warn others about it [264, 567]. One common problem is that users tend to write reviews mostly for products that they perceive exceptionally good or the opposite [160]. Another problem of peer-provided information such as user ratings are susceptible to biases [522, 250].

Computational-provided information: This category includes all information that is not provided by oneself or other peers, but collected by electronic devices, e.g., sensors in smart home environments [630]. The benefit of computational-provided information is that they are not collected manually, so there is no additional effort for users. Secondly, they are based on automatically generated measures and therefore are more difficult to manipulate by users than self- and peer-provided information (or at least it requires an additional effort). Additionally, computational-provided information could solve

the problem of only rating those products or services that were perceived very good or very bad.

Telematics can enable new forms of trust building [83] by, e.g., enabling assessments based on quantitative real-time data on driving behavior. Such systems can be described as the use of technical devices for the identification, storage and/or processing of computational-provided information which are interconnected by means of telecommunication systems [670]. The term is often used synonymously with car telematics and thus, Connected Car technologies. Connected Car technology enables numerous driving and environmental data to be captured by sensors and processed into valuable information via the Internet [603]. This possibility offers the potential to protect reputation systems from manipulation and to stabilize user trust in evaluations and transactions [457, 514, 680, 602]. In addition to integrated sensors, telematic solutions can also be implemented in vehicles via telematics boxes, dongles or provided by smartphone apps [238, 357, 424, 594, 277, 249]. In P2P carsharing, such technologies could enable a condensation of the driving characteristics to a score that can lead to an overall rating. The rating represents the potential risk of a rentee using the vehicle and can therefore increase trust towards rentees with a better score [630, 188, 457, 629]. Looking at the specific domain of car insurance, the basic idea is to record acceleration and position with the help of sensors while driving to detect, e.g., speed violations, braking, mileage and traveling direction and to use this information to assess risk of a person's driving behavior [393]. The fee then expresses the insurance company's confidence based on the risk assessment that the policyholder will not suffer an accident [164, 521].

However, an algorithmic bias can also occur here, which penalizes users with a driving behaviors that do not correspond to the "algorithmic imaginations". Due to the fact that it is almost impossible for sensors to gain an overall understanding of specific traffic situations, computational information should be more flexible in analysis to reduce the "practice of bias" [288, 23]. Rouvroy calls this phenomenon the algorithmic governmentality, which operates with infra-individual data and supra-individual patterns without, at any moment, calling the subject to account for himself. Furthermore, Rouvroy has introduced the term "data-behaviourism", which according to him is the widest

possible zone of indistinction between reality and the world. Because of the representation of the world based on data, it becomes difficult to criticize conclusions based on algorithms, because data are usually described as objective and neutral. This bias may disadvantage individuals who, for good reasons, violate or are unaware of the algorithm's rules [527].

11.2.3 System Intelligibility and Accountability

The relevance of comprehensible systems can be seen in the increasing amount of work in the area of system intelligibility and Explainable AI [2, 395, 616, 680]. Due to recent advantages in Machine Learning (ML) and Artificial Intelligence (AI) we increasingly see automated algorithm-based decision-making in digital technologies. Nevertheless, it is a fundamental challenge to design these technologies intelligible and accountable such that people can understand the information and feel empowered [2]. There is already related work in interpretable, fair, accountable, and transparent algorithms [514, 149, 354] in the AI and ML as well as HCI communities. The European Union approved a data protection law [228, 491] that includes a "right to explanation". Users are often not explicitly informed about the information they see and algorithmic decision-making systems typically do not provide visibility into how the technology works [437]. While transparency on the one hand is necessary to improve collaboration between humans and algorithms, missing transparency on the other hand is a risk, if humans have difficulties figuring out how the systems reached their decisions [97]. Therefore, explainability is needed for collaboration and is related to interpretability and justification [63]. Systems are interpretable if a human is able to understand the reasons for the decision, "justification explains why a decision is a good one, but it may or may not do so by explaining how it was made" [63]. An explanation is important to support users to understand, what outputs the system is supposed to produce and recognize mistakes or errors [504]. While empirical studies showed "the importance of explanation to users, in various fields, consistently [...] that explanations significantly increase users' confidence and trust", current AI and ML systems are weak in this area [63, 98]. Positive examples are rule-based models such as decision trees, which are

easy to understand in contrast to most types of complex neural networks. In particular, the reduction of complexity plays a crucial role here [569, 354]. Since rule-based algorithms are more comprehensible in contrast to complex black box models [310, 276, 33], the specification of individual preferences and norms regarding reputation systems can support a Human-Algorithm Interaction [686, 556].

11.3 Empirical Pre-Study: Understanding Algorithm-Based Reputation Systems in P2P Carsharing

The aim of the pre-study was to 1) understand how people imagine the functioning and derivation of an algorithm-based score on driving behavior enabled by telematics systems (Driving Score) in terms of what information is needed and how it is processed to a score, and 2) to derive first design implications for an artifact, integrating the users needs and perceptions in the context of P2P carsharing.

To address our research goal, we choose a qualitative research approach. The semi-structured interviews lasted about 45 minutes on average. An interview guideline was used to narrow down the subject area and to create a structure for the interview process. The interviews started with a description of P2P carsharing and the principle of telematics solutions as an introduction. Altogether, the interview questions can be divided into three categories. The first category consisted of questions on which data should be collected for scoring driving behavior. Furthermore, the interviewees were asked to explain their understanding of a Driving Score and how they think it works. In the second category, design aspects of a algorithm-based score were examined. We presented four different profiles of people who would like to borrow a car (translated from German) (see Figure 20). However, the order in which an interviewee responded to the questions was flexible, which allowed the interviewees to deal with topics that seemed more important to them. This procedure enabled the interviewer to lead the conversation by asking specific questions without interrupting the flow of the conversation [320].

The study was conducted in Germany with a sample of 16 participants (P1-



Figure 20: Different Driver Profiles Presented to the Participants – A Driver Profile Consists of a 1) Picture of the Person, 2) Name, 3) Age, 4) Driving Score, 5) Kilometers driven, 6) User Rating, and 7) Number of User Ratings

P16), that was recruited through personal contacts. Seven participants were female and nine were male. The mean age is comparatively young (33.8 years in a range of 22-64 years), since the young generation is the main target group for carsharing platforms [383]. Moreover, our sample covers a broad variety with regard to Telematics Experience, Tech Affinity, and experience with car-or ridesharing. For the analysis all interview were recorded and transcribed. The analysis itself followed the procedure of thematic analysis [86]. Here, two authors coded the transcriptions independent of each other using MAXQDA[5], and combined the resulting code system collaboratively [58].

11.3.1 Understanding of an Algorithm-Based Reputation System

At the beginning of each interview, we started with the question of how people understand the computational logic of an algorithm-based reputation system. Since many participants found it difficult thinking themselves into an algorithm, we asked them to put themselves in the position of a co-driver and evaluate driving behavior from their perspective. This led to a list of measurement criteria, which were associated with the term driving behavior. This

[5]MAXQDA is the world-leading software package for qualitative and mixed methods research. (https://www.maxqda.com)

discussion also revealed different levels that could be used to classify the measurement criteria (see table 13). While some of the mentioned measurement criteria are on a technical level and more easy to capture and process into a score (e.g. braking), other criteria are more norm-based and thus require an aggregation of several factors to enable an evaluation.

Table 13: List of Categories of Measurement Criteria of a Driving Score

Norm-based Measures	Speeding Behavior	”[...] speed in relation to traffic rules, [...].” –[P8]
	Pull-Out and Overtake Behavior	”[...] or if he drives on the motorway, that he does not swing out quickly and then go back into lane again. Simply a dignified driving style I would say.” –[P11]
	Cornering Behavior	”[...] how fast he might take a turn?” –[P8]
	Driving Indications	”[...], do you put the turn signal on properly?” –[P16]
	Being Attentive	”Do you pay attention to several things at the same time, such as, e.g. cyclist, train and others.” –[P16]
	Uniformly Driving	”If you now take the shortest route, with side roads, then you brake frequently and shift a lot.” –[P14]
	Road Traffic Regulations	”[...] then whether the driver complies with the regulations or observe the signs.” –[P4]
	Holding Distance	”How close does he get to the front car?” –[P5]
Technical Measures	Braking	”[...] abrupt speed reduction, meaning something like a brake.” –[$P15^{a,b,c}$]
	Steering	”[...], probably also jerky movements from the steering wheel.” –[P1]
	Acceleration	”So, there’s no abrupt acceleration and deceleration.” –[P14]

Often the interviewees oriented themselves to the road traffic regulations without thinking about the technical feasibility. This can be seen in quotes such as: "Do you put the turn signal on properly?" –[P16]. The first part of the quote can be answered with yes or no. The second part, however, contains the word "properly" that is known when, where, how long and how often you must indicate a turn. Here the term "properly" has obviously already been in-

terpreted as correct or in conformity with the law. The quote can be related to the following factual context: The flashing light must be set several seconds before the start of changing the direction or lane. The setting of the signal is accompanied by a shoulder check and a look at the mirror in the direction, which is to be taken. These activities are to be regarded as road traffic regulations. Most participants created a fictitious situation in which they gave an example of how they imagine an evaluation while driving. In most cases, the participants chose the criterion speed for their use cases:

> ”Well, I would pay attention to how fast they are and what speed is allowed on the respective road or motorway and then I would pay attention to whether that was adhered to. So $\pm$10 kilometers per hour are okay, but I would make sure that the driver follows the traffic signs.” –[P4]

As previously stated, most situations referred to already interpreted factual contexts. One connection, e.g., was that there are certain rules to which speed must be aligned. If the driver doesn't follow the rules, a consequence can be expected. Additionally, it is interesting that the interviewees also specify tolerance ranges for rule-based reference values. In the above example, a tolerance of $\pm$10 kilometers per hour was specified despite compliance with traffic signs. Consequently, it can be assumed that reference values in conjunction with the categories of measurement criteria mentioned (according to human understanding) serve the algorithm as a basis for evaluation.

The measurement of the driving behavior in terms of speed was mostly described with words such as "above" or "below", "good" or "bad", without explicitly naming the reference values. There are, e.g., no reference values for the ideal number of changes of direction while driving in the traffic regulations. Trying to represent the statement in a score would favor drivers who drive straight a lot in contrast to drivers who must turn more often. Whether this supports trust in carsharing and result in a realistic Driving Score remains to be discussed. In order to be able to calculate the values of the deviation, the factors must have a uniform unit of measurement. Most respondents stated that the resulting score should be an average value of driving behavior [P7]. Further, a weighting of criteria was desired by our participants.

> ”But I would also like the weighting to be shown to me. What is actually weighted the most or is all weighted equally? Or what the most relevant criterion is. Then I could understand what the score tells me at the end.” –[P1]

The statements on the weighting of the criteria for the driver evaluation varied in the interviews. Each participant has own preferences of weighting criteria. Some participants also mentioned a kind of knockout variable, which alone caused the score to drop considerably. Participant 3, e.g., stated "everything that exceeds 10 % above the speed limit is not okay". But only when the composition and the weighting are transparent, the score becomes understandable and able to build trust. In addition to the information which criteria are included in a Driving Score, participants questioned how strongly these criteria affect the final score. The participants wanted to know how each individual value is considered in the overall score.

> ”I would like to break the score down and see the individual parameters.” –[P1]

11.3.2 Perceived Potentials of Algorithm-Based Reputation Systems for Trust-building

In the following, we aim to show what potentials of algorithm-based reputation systems our participants perceive and how these can support trust in P2P carsharing. As already mentioned we presented four different profiles of fictitious car rentees (see Fig. 20) to our participants and asked them to choose a profile and explain their decision. What was striking about this question in the interviews was that the majority based their selection on the Driving Score and only paid secondary attention to the user ratings. In total 8 out of 16 participants stated that they consider both, the Driving Score and the user ratings in their decision, with a tendency towards a higher weighting of the Driving Score. Six participants made their decision only on the Driving Score and only two participants relied on the user ratings. It was possible to get an impression of when a score was considered "positive" or "negative". Basically, we found many different perceptions of "trustworthy" scores. Often,

the values were compared with the ratings of hotels. The participants didn't state an explicit score, from which a positive or negative impression of the rentee arises, but "5.2/10" was called "not decent" –[P10]. When selecting the profiles, the participants often paid attention to how many kilometers the person had already driven. On this basis, respondents determined how reliable the Driving Score appears [P1]. One of the participants found 500 kilometers driven to be a good basis for the validity of the Driving Score [P1]. Whereas the indication of 177 kilometers was often perceived as not much, but sufficient enough [P10].

Various benefits could be found for the use of algorithm-based scoring for the evaluation of driving behavior. For example, algorithm-based evaluations were perceived as less manipulable [P2]. The respondents described them as more accurate and neutral [P10, P15]. According to Participant 5, personal evaluations are difficult to quantify and do not have a fixed evaluation scheme. Some of the respondents argued for a combination of both information:

> "I think a general score should perhaps consist of both things and not be divided generally [...]. Perhaps, the user rating can contribute to a certain extent to completing the Driving Score." –[P3]

In this context, the reference to reality was emphasized. This factor was often mentioned, which is why the respondents demanded personal evaluations in addition to the algorithm-based scoring. An example of this was a situation where a driver is on the acceleration lane and has to accelerate quickly to drive onto the motorway.

> "If the Driving Score encourages you to drive safely, I think it's really cool. But if you had to accelerate because you're on the acceleration lane to drive towards the highway and two trucks are on the lane and you're afraid to accelerate because you just don't want to risk your 10/10 score, that's not cool. So, if you get too involved into keeping good scores, it could have a negative effect." –[P3]

Since fast acceleration would be assessed negatively, the Driving Score could incentivize an unwanted driving behavior in some situations [P3], e.g., the evasion of ambulances, would possibly falsify the assessment [P1]. Almost one third (5/16) of all respondents stated that the Driving Score must function reliably and be extensively tested [P10]. The participants expressed security concerns in the form that the underlying technology must be mature an be tamper-proof [P1]. Under the premise that the technology is tested and found to be safe, a score calculated by this technology was described as "objective", "trustworthy" –[P9], and "fair" –[P15].

> "I personally would set the weighting to 70% Driving Score and 30% user rating. Personal rating says it all. There is a subjective perception behind it. And this Driving Score is actually trustworthy and objective [...] and I can say, okay, these are data-based ratings, which I can accept without doubt." –[P9]

As a part of the selection decision, potential biases arose both explicitly and implicitly. The majority of participants chose the profile with the highest Driving Score and the most kilometers driven (cf. Figure 20, Christina). However, especially for those who relied on the User Rating, the profile picture and age would probably play a major role.

> "So the picture, I think that would be good if that would be in there. So I would know who is driving the car. I think that's not completely unimportant, and also the age, whether the person has had a driver's license for a long time." –[P5]

Overall, it can be said that many of the participants see potential in the Driving Score for their decision-making. However, the score must meet certain criteria in order to build trust. This includes that it is tamper-proof, has enough data based on many kilometers driven, and is comprehensible for the users. Some participants motivated the combination of Driving Score and User Rating. How a Driving Score can be designed more comprehensible is discussed in the following section.

11.3.3 Designing a Driving Score

The results addressed many aspects that have an influence on the design of a Driving Score. One aspect is the visualization of it. During the interviews, the participants mentioned various visualization options. These include 1) scales, 2) diagrams, 3) ratios, and 4) other visualizations. In our qualitative thematic analysis, we counted the mentions of these visualization types and the 10 of our participants liked the idea of a 10-point **Scale** for the Driving Score.

> ”10 is great and 0 is bad. I’d like that, because there’s such a benchmark, where you can say if a driver gets a 7, [...] you know, this person drove not perfect, but well enough.” –[P6]

In addition to a scale from 0 to 10, other intervals such as 1 to 5 or 1 to 6, based on the German school grading system, were mentioned [P7, P8]. As already indicated in the quote, scales offer a good orientation, because they are easier to interpret due to the clear graduation between the best and worst rating.

Diagrams were mentioned a total of 8 times in various forms. One idea was to create radar charts with all dimensions which impact the Driving Score. Each dimension represents a part of the net and fills up in a corresponding direction. The more points a driver reaches, the more this circle fills out [P8]. Another form of representation was a bar chart with a green area, yellow and red areas [P2], respectively for good, acceptable or bad driving behavior. The ideas of bar and line diagrams were very similar. Also, a scatter-diagram was proposed. The scatterplot has a horizontal line on which an average speed could be plotted. Above the line, points can be displayed at which the speed was above average and below the line, braking maneuvers would be recorded. On axis-based diagrams, target and actual lines could be displayed that allow a direct comparison [P14].

Ratios were mentioned 7 times in total. According to a statements, participants seem to prefer percentages over star ratings [P1, P15]. This is considered an advantage for a drill-down function to obtain more detailed results than a star rating.

> ”Of course, you can do a bit more with percentages than with stars. In particular, I would like to tap it again and see where I didn’t have the right speed. So just more drill-down options so I can see more. I would like that with all data.” –[P1]

A general requirement for the Driving Score was that it should be well structured and comprehensible [P5]. All 16 participants explained that it must be possible to break down the score that is finally calculated.

> ”I’d want to know exactly and transparently all the criteria that go into this score.” –[P3]

Participants propose different forms of visualizations for different criteria. For example, when assessing the speed, deviations from the average speed were desired in the form of line diagrams, while for other criteria, such as smartphone use, people mentioned an indication in minutes. The respondents do not want a too complex score as too many factors would reduce the benefit of a score [P12]. Therefore, they assign importance to a comprehensible listing of the measurement criteria, but the listing must, nevertheless, be limited.

11.4 Prototyping

Based on the results of the pre-study, an interactive prototype [555] was developed based on the results of a co-creation workshop. The workshop lasted about 2.5 hours and included 4 male and 3 female participants between 25 and 65, who are interested in car- and ridesharing. The content and developed paper prototypes were documented.

The co-creation workshop aimed to embed the findings of the pre-study into a P2P carsharing app prototype according to the participants’ perceptions. After introducing the participants to the problem of reputation mechanisms in P2P carsharing, they received an in-depth briefing on the results of the pre-study. In the following creative phase, the participants were asked to sketch the app UI and discuss its computational logic from an end-user perspective. This phase was supported by the exemplary profiles of Figure 20 pictorially and

the prepared results in the form of printed power point slides. Furthermore, printed app icons, diagrams and materials such as pencils, paper and scissors were available. Finally, there was critical reflection on additional factors.

The final clickable prototype (see Fig. 21) was then designed by the authors and implemented in Figma. The individual app components and their functions are presented below.

11.4.1 Preference and Individual Norm Settings

Our pre-study participants perceive algorithm-based driving scores as an "interplay" of various measurement criteria [P1, P10]. However, rule-based systems have weaknesses in dealing with the uncertainties of the real world [276], such as rapid acceleration, which in some cases (in reality) are necessary to get on the motorway quickly. Furthermore, according to [527], the objectivity and neutrality of sensor data lead to additional hurdles for an algorithm: If, for example, the sensor used to measure the distance hold contains objective information (”close object”), but this information is generated by dirt on the sensor and not by the actual object to be measured, the evaluation may be objective based on the given data, but this leads to an incorrect evaluation of the driving behavior. In the co-creation workshop, our participants therefore designed a customizable Driving Score to face these problems by giving the opportunity to set a lower weight to such situations and defining individual norms, as already mentioned in the pre-study interviews, to become more aware of the algorithm’s rules.

To allow an adjustment of the Driving Score and therefore a customization, the participants of the co-creation workshop sketched weighting sliders for the measurement criteria in addition to a brief explanation, which represent the preferred weighting (0 - 100%) (see Fig. 21 a). In addition, the workshop participants were in favor of inserting an example from the pre-study for the individual norms to show how such a norm could look like (see Fig. 21 b).

In line with the pre-study, the calculation of the Driving Score was considered to be an average value of driving behavior [P7]. An influence factor is measured between 0 and 10 and weighted according to individual preference

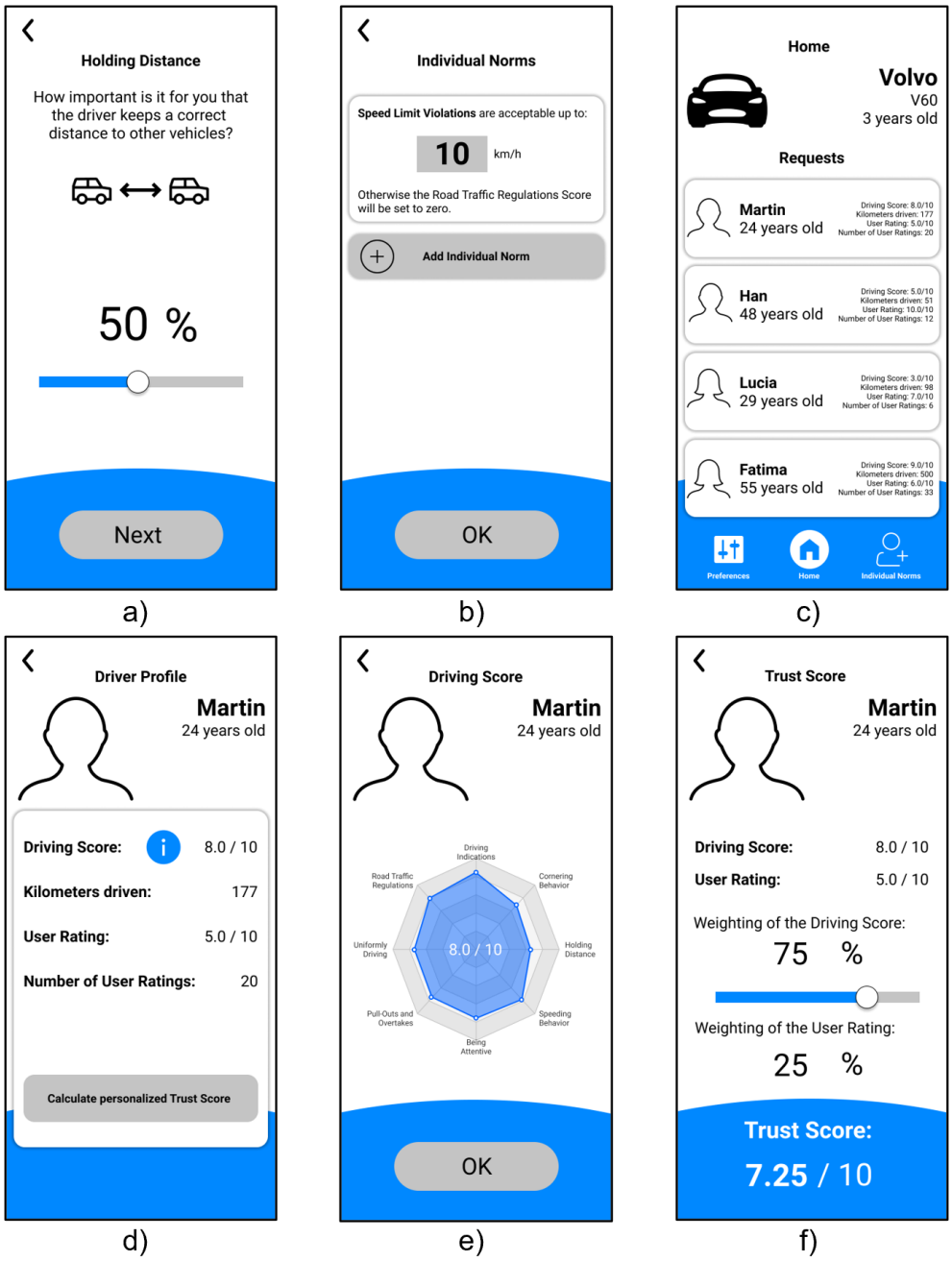


Figure 21: Clickable Prototype – a) Exemplary Preference Weighting, b) Individual Norm Settings, c) Home Screen with Requests, d) Exemplary Driver Profile, e) Exemplary Driving Score Drill-Down, f) Exemplary Trust Score Weighting

settings [P15]. Individual norms are modeled as exclusion criteria, e.g., that every speeding $>$ 10 kilometers per hour leads to an average factor according to road traffic regulations of 0. The Driving Score thus corresponds to the sum of all weighted factors divided by their number.

11.4.2 Home Screen

After participants have set their preferences and individual norms, they are redirected to the home screen (see Fig. 21 c). The workshop participants mainly designed their sketches based on the exemplary driving profiles provided, as these closely resembled a real app. Further, the home screen show four exemplary driver profiles, which clearly display the profile information of the potential rentees. The home screen is designed to simulate a real decision-making situation in which the participants of the evaluation interviews had to choose a profile. To obtain more detailed information about a potential rentee, they should independently outline their motivations and click on the driver profiles.

11.4.3 Driver Profiles and Drill-down

To improve the intelligibility the participants of the pre-study, as well es the participants of the co-creation workshop discussed various types of visualizations for the Score, such as scales, radars or ratios. Radar visualizations were chosen for a deeper understanding of the composition of the algorithm-based Driving Score. With regard to the pre-study, a Driving Score should therefore offer the possibility of obtaining not only the initial information of the Driving Score, but also further visualizations for a deeper analysis. In the sketches of our co-creation workshop, as well as in our prototype this was realized by displaying the score within the Driver Profile (see Fig. 21 d) and visualizing the measurement criteria of the score in the form of a radar chart (see Fig. 21 e). In addition, an overall Score was desired as a composition of the algorithm-based score (Driving Score) and the (average) peer-provided rating [P9]. For reasons of differentiation, we named the combined score *Trust Score*, which is an combination of the Driving Score and the User Rating (see Fig. 21 f).

The participants of the co-creation workshop stated that the drill-down of the Driving Score in the form of the radar chart (see Fig. 21 e) should be optionally accessible in the sense of the evaluation by an info button. This should show whether the participants of the following evaluation show interest in it, or ignore this information from the beginning.

11.5 Evaluation: Impact of an Algorithm-Based Reputation System in P2P Carsharing

The evaluation aimed to discuss the design and suitability of the algorithm-based reputation system in the form of the Driving Score in a scenario-based decision-making situation. The main focus was on the setting of individual preferences and norms for interacting with the algorithm of the Driving Score, the perception of the Driving Score and the user rating for the selection of the vehicle renter, and the relevance of drilling down the Driving Score and Trust Score. Furthermore, it became clear in the pre-study that discrimination due to social bias may occur based on profile information (name, age, gender). In line with prior research [250, 216], racial, gender and age biases were simulated to enable a discussion on whether the Driving Score can contribute to a reduction here. For this purpose, the exemplary profile names (see Fig. 20) were adapted with respect to the names of the potential rentee (see Fig. 21 c).

We conducted interviews with 12 participants (E01-E12) aged between 18 and 56 years, which lasted on average 47 minutes. There were 6 female and 6 male interview participants with different nationalities and carsharing experiences acquired through previous contacts of research projects. They followed a semi-structured interview guide in combination with the use of the clickable prototype. The establishment of preferences and individual norms, the selection of the vehicle rental company, and the role and relevance of User Rating, Driving, and Trust Score were addressed. Finally, with regard to the investigation of potential discrimination, the (expected) influence of profile picture, name, age and gender was addressed. After transcribing the recorded interviews, we analyzed them following the inductive approach of thematic

analysis [86]. The authors undertook the coding of the interview material collaboratively [58].

11.5.1 Preferences and Individual Norms contribute to the Understanding of Algorithm-Based Scores

The settings of preferences for the calculation of the Driving Score is perceived as positive by all participants and contributes to a deeper understanding and transparency. Respondents indicate that they can incorporate their own preferences into the Driving Score, which helps to build trust and acceptance. The setting of high values can essentially be related to the perception of a high risk of accidents and a high level of car wear and tear. The motivation behind this is primarily based on a lack of trust in potential renters as well as their own negative experiences.

> ”I would rate him 100%. If he doesn’t obey the traffic signs, that’s a danger to my car [...]. It is a matter of trust when I rent my car to someone.” – [E05]

However, it becomes clear that some respondents already differentiate here according to rules and norms, which is consistent with the results from our pre-study. These norms result from the reference to their own driving behavior or from the perception that an algorithm-based score should not be too restrictive in some situations. These situations are categorized by some participants as speed-related or weather-related.

> ”He is supposed to stick to the speed limit but he can also speed 10 to 15 km/h too fast but only on the highway. In a 30 zone he should drive 35 at most.” – [E05]

The setting of low values is exclusively based on the orientation to own driving behavior or the assumption that the algorithm-based score does not recognize certain situations fairly. This finding supports the results from our pre-study that most people allow a tolerance range as a norm-based approach.

> ”That’s okay, if it stays within a certain range, e.g. 10 km/h faster.” – [E08]

Based on our prototype, participants state novel ideas for a better understanding of the algorithm’s processing of high or low values. For example, this can be achieved by a clear description of the adjustable levels as feedback on how the algorithm will interpret the respective setting.

> ”I would like a little bit more explanation about the impact when I change between 0% and 100%. E.g. for speed and 0% it could say ’Speeding violations are not taken into account at all’ and for 100% ’From 5 km/h faster there is a penalty, the higher the violation is.’” – [E12]

The participants argued that the definition of individual norms should be as intuitive as possible. Using the example of speeding violations, it is discussed whether a percentage or an absolute speed deviation divided into speed zones should be specified. According to the participants, a percentage figure offers too much room for higher violations.

> ”I would say that 5-8 km/h is still okay, between 8-10 km/h is borderline and above that it is negative in town. There would have to be a certain realism here. At some point, however, it becomes too cumbersome for the user to configure.” – [E10]

In addition to the setting of individual limits for speeding, the specification of a maximum speed, a maximum consumption, and the maximum number of previous accidents are also desired. Overall, two participants mentioned a final summary of the preferences set in form of a persona, which is characterized by comprehensible descriptions and could serve as a filter for preselecting requests.

> ”I would like to have a kind of persona at the end, which represents the characteristics of the driver corresponding to the defined preferences. So as a summary: ’He always keeps to the speed limits’, ’Almost always uses the turn indicators’, etc. You could then use it as a filter or see if there is something to adjust.” – [E12]

11.5.2 Support for Decision-Making and Trust-building

Before participants look at the ratings, some already make a pre-selection based on age or gender. In the second step, they take the Driving Score and the User Rating into account, which they mostly assess in combination with the kilometers driven and the number of ratings as supporting parameters. At this point, some of the participants deviate from their pre-selection by age or gender. This is mainly due to the Driving Score, which distracts the participants from their initial thought and makes them think about the relevance of social factors and driving safety.

> ”Okay, then I would actually rather, [...] funny, I actually didn’t look at the women at all, although Fatima would fit best, at least in terms of kilometers driven and the driving score. First I went through the men and when I look at the women now, I would actually rather choose Fatima at that point because of her Driving Score.” – [E02]

Confirming to the pre-study it becomes apparent that the Driving Score comes to the fore especially among participants who pay attention to driving safety. The User Rating is predominantly preferred by respondents who pay more attention to the condition of their vehicle as well as social factors, which are reflected in cleanliness and reliability.

> ”At the end of the day, it’s business what I do [...] it’s very important to me that he drives safely with my car. How he looks or if he is unsympathetic doesn’t interest me that much.” – [E06]

However, decision-making based on just one score is not enough for most people. For the proponents of the Driving Score, the Trust Score usually has a confirming effect on their decision. A prorated weighting of the user rating is used to ensure that the social component is not completely disregarded. However, the Driving Score is perceived fairer and is therefore considered more relevant by them.

> ”I think that’s cool that you can weight the preference between driving score and user ratings again. It hasn’t had too much of an impact on me now. It had something confirming, because my decision for Fatima was clear quite quickly. First, because of the Driving Score and then the Trust Score actually confirmed it again” – [E01].

Interestingly, the Trust Score contributes to the fact that those who use the user ratings as a key decision-making parameter are sometimes unsure about their decision when setting the Trust Score and, by then looking at the drill-down of the Driving Score, sometimes reconsider and select the rentee with the highest Driving Score. Because of that some respondents also state that they first need to build up trust in the Driving Score to determine whether it matches their individual expectations.

> ”I’d like to see the radar chart on him again, though. Oh, that’s significantly worse. I think it’s good to see the individual points. Now I realize that the Driving Score is more important to me than the user rating. Han scores worse than Martin in the driving score factors that are important to me, even though he is older.” – [E07]

11.5.3 Contribution of Algorithm-Based Reputation Systems to the Reduction of Discrimination

As seen in the previous chapter, when selecting a potential rentee on the home screen, some of the respondents already expressed that information such as age or gender played an important role for them. Here it also became clear that the Driving Score can lead to a change of opinion.

> ”I think at first glance you can already see all the important info. Of course, male, female. I don’t lend my car to a woman then.” – [E02]

When asked explicitly what influence the name had on their decision, most participants stated that it did not play a major role for them. Nevertheless,

three participants said that they were prejudiced against foreign-sounding names. However, two of them finally adjusted their decision based on the Driving Score and chose Fatima.

> "Yes, with Han I already paid attention to the age and Fatima, sounds bad now, the name first struck me negatively. And with Han, too, that's why I was looking at Martin first. But I have now chosen Fatima. For decision making after analyzing the Driving Score, it wasn't so relevant." – [E08]

When asked whether the participants could generally imagine that a (foreign-sounding) name could play a role for other car owners, all referred to its discriminatory effect. Mostly reasons of antipathy, stereotypes and language problems were mentioned. The influence of a profile picture also stood out here.

> "I think with Fatima, many have an image in mind. An older lady, 55 years old with a headscarf and reserved. But if you have a profile picture now of Fatima standing in the middle of life, without a headscarf, who you find likeable. People would certainly not pay so much attention to the name if the picture is appropriate." – [E09]

Opinions are divided on the suitability of the Driving Score for reducing discrimination and social bias. On the one hand, respondents find the Driving Score objective, which they see as an advantage over user ratings. In addition, the Driving Score's assumed objectivity may minimize discrimination against potential rentees by age or gender, as it creates transparency regarding driving ability. This was also reflected in the decisions made, which were predominantly motivated by the Driving Score.

> "[...] with an older woman, where you now assume that she doesn't drive so well anymore, but then you see the Driving Score, which can also be weighted and you can look at the assessments of the specific parameters, that definitely convinces you otherwise." – [E04]

On the other hand, the Driving Score reaches its limits in case of discrimination based on name or profile picture if it is racially motivated. In the case of prejudice, however, looking at Driving Score or Trust Score could possibly reduce it.

Some participants also expressed suggestions to better design the driver profiles to counteract discrimination. It was suggested that the important information, such as the Driving Score and the User Rating, should be visualized larger and that the discriminatory factors, such as name, age and gender, should be made smaller to not draw attention to them. Furthermore, only mentioning the first name is sufficient and a profile picture can be omitted if the renter is verified on the platform by means of depositing the driver's license data.

> "So if you don't show a picture, the Driving Score can definitely reduce discrimination. [...], so if you only have the first name and [...] verify the profile through the driver's license. Then that's perfectly sufficient and draws attention to the Driving Score." – [E12]

11.6 Discussion

Based on our results from the pre-study and evaluation, we want to discuss four main findings in regard to an algorithm-based score. These include how rules and norms can be considered in an algorithm-based reputation systems, why explainability is important for the understanding and trust-building, why such reputation systems could reduce social biases and discrimination, and finally why they are no magic bullet for trust-building.

11.6.1 Algorithmization of Rules and Norms

The findings of our pre-study show that the understanding of a Driving Score and its interpretation is strongly aligned to the Road Traffic Regulations. Good driving behavior seems to be the practical compliance to these regulations. Since every driver always must adhere to it, this standard seems to

automatically serve as a general rule for the evaluation of driving behavior. Within the framework of sociological institutional theory [112], algorithms are perceived as elements of social order processes and considered on four levels. 1) The regulative level serves the implementation and execution of formal rules, while the 2) normative level evaluates alternatives for action regarding their legitimacy. An interpretation and perception of the action alternatives and social contexts takes place on 3) the cognitive level. Furthermore, this concept is extended by 4) a technological level, which ensures that social expectations and rules are reflected, transformed and further embedded in the algorithm [217, 218, 688]. Such a differentiation can be found in our pre-study results, as well as our evaluation. From the user's point of view, the regulative and normative level stands out in particular. For example, many participants judge regulative, so that in the event of rule violations in the form of exceeding or undercutting – which are usually measured according to the road traffic regulations – a negative impulse should be given to the Driving Score. Some participants also mentioned various subjective legitimations for deviations from a regulatory (limit) value, so that a legitimating normative level can also be identified here. This is characterized by the addition of subjective scope to an objective evaluation and, in this sense, alternative actions for the algorithm, e.g., the tolerance in speed by +10 km/h [P3] (see Fig. 21 b).

The evaluation confirm the necessity of differentiating regulative and normative levels. However, they add a cognitive level through the desire to subdivide individual norms into different categories due to the dependence on the (socially) perceived context. Similarly, the technical level is also addressed by providing weighting options and explanations, ensuring that these parameters are consistent with their expectations. Further, the cognitive level is complemented by extending the weighting of preferences by including an accident probability and criticality of the situation within a social context. As driving behavior may be forced in specific situations, like braking abruptly and hard or accelerating too fast, it should be assessed in the context of the situation. For example, the following algorithmic scoring could be performed to assess cornering behavior considering the situation: 1) If someone drives too fast, 2) under bad weather conditions through 3) a tight curve with 4) a high

probability of an accident, he or she should be punished harder than if some parameters are not that critically assessed.

11.6.2 Need for Understanding to Build Trust

Furthermore,the pre-study shows that it is essential to make algorithm-based ratings transparent and explainable, which is confirmed by the results of the evaluation. The principle of transparency and explainability seems to rely on the attitude *"What you don't understand right away may not be understood at all"* for the participants. The clearer and more comprehensible the scoring, the more trustworthy it seems. A clear visualization is of great importance for the comprehensibility and therefore for the trust-building process. It should be possible to break down the Driving Score to support the understanding by increasing transparency of how the Score is composed. According to Aigner et al., the visualization must consider 1) what kind of data need to be visualized for 2) what task, in order to 3) choose a suitable visual representation [18]. Edwards and Vaele [180] discuss, e.g., the right of explanation in the context of automated decision-making. They argue that this right of explanation is helpful but is only a transparency fallacy [180, 627]. Often, users have little use for a precise breakdown of such scores due to a high complexity. Therefore, it is important to design a Driving Score which informs and supports a decision-making process instead of replacing it. According to Harper, it may be necessary to understand the algorithm in its entirety when dealing with HCI, when designing the interaction between a computer system and a user. But the resulting design should be such that an understanding of the computer system is no longer necessary [251]. That means that explanations only matters when they are relevant to the user's purposes. The Driving Score should therefore provide explanations in a degree which make the score interpretable for all parties on P2P sharing platforms, but they should not necessarily explain how the algorithm works.

Here, the evaluation interviews gave some interesting insights about the individual relevance of the specific Driving Score parameters, which should be accompanied by a description. A comprehensible description of the different possible configurations of Driving Score parameters could support a bet-

ter understanding of the algorithm. Furthermore, a clear and understandable summary of the preference settings and individual norms in the form of personas is desired, which should also act as a ”personification” of the algorithm for comparability and indication of possible need for adaptation in case of a too restrictive or too weak specification. This is in line with prior research discussing the generation and incorporation of explanation sentences in the context of recommender systems [706, 119]. At the same time, it can also serve as a filter to pre-select the potential rentees displayed. According to the evaluation interviews, the drill-down of the Driving Score in the form of a radar chart further supports the understanding and in some cases, when the User Rating actually dominate decision-making, even contribute to rethink a decision. It supports a individual reflection of the driving behavior based on specific parameters.

11.6.3 Potentials of Algorithm-based Reputation Systems to Reduce Biases and Discrimination

It becomes apparent that the Driving Score can reduce potential biases against vehicle rentees, but not completely eliminate them. While participants of our pre-study and evaluation consistently value the objective nature of the Driving Score, for some this is not enough to place sufficient trust in the driver. Regarding trust in the Driving Score, individualizability, e.g., in the form of individual norms, can help to make the system more flexible [23] in order to reduce algorithmic bias as well. As trust is composed of interpersonal and technological trust, the combination of Driving Score and User Rating addressed by the participants could have a positive effect on the trust-building process [389, 259, 526, 602]. Furthermore, the Trust Score could also counteract users' bias due to their trust attitude towards both User Ratings and algorithm-based scores [602]. The possibility of individualization and its visualization can increase the comprehensibility of the Driving Score, which is likely to benefit technology-skeptical users in particular [574]. Furthermore, by combining them in the context of the Trust Score, ambivalent users can specify the composition of their preferences over user ratings and algorithm-based scores. This leads to a better decision support of the users and allows a

better cooperation between the different user groups. Especially the findings of the evaluation show that an algorithm-based approach can also contribute to the reduction of discrimination. In several responses, discrimination in terms of age, gender, or nationality could be identified. Complementing to prior literature [104, 179, 107, 638], we could often observe a rethinking of their decision caused by algorithm-based reputation systems, where they first judged based on characteristics like name, age or gender. After a closer look at the Driving Score drill-down, the decisions were adjusted and justified with the good results in the scores. Furthermore, the profile design should be adapted to draw attention to the relevant factors such as the Driving Score. The profile information could be displayed smaller compared to the Driving Score and User Rating so as not to draw attention to them. Highly discriminatory factors such as the profile picture could be removed by verifying the renter through the platform.

Nonetheless, algorithmic bias can occur here as well, which may disadvantage some users or driving behaviors in specific situations [288, 23]. In this regard, shortcomings are also discussed in the following chapter.

11.6.4 Shortcomings of Algorithm-Based Scores

Most of the participants perceive algorithm-based scoring solutions as correct and fair. Nevertheless, they also reach their limits at some point. Often, we have the case that maneuvers, which are necessary, can be evaluated negative. An example of this is the acceleration lane on motorways. Abrupt acceleration is perceived as bad driving behavior, but it is necessary in some situations. Much more serious, however, seems to be the occurrence of unpredictable influences in road traffic. The definition of rule-compliant behavior therefore varies from situation to situation. The algorithm doesn't know if abrupt braking or a strong acceleration is necessary or not. This confirms the findings of Rouvroy and Alkhatib, whose work draws attention to a bias in algorithm-based decisions and seeks to focus on the importance of subjective perception of individuals [527, 23]. This problem can also be seen within the context of autonomous driving [263]. The participants do not want to have a disadvantage by unexpected maneuvers that may have been necessary. An-

other problem can be an unfair algorithm-based scoring. Just because some drivers drive more sporty that does not mean necessarily they drive unsafe. This could lead to a structural tendency of algorithmic systems to specifically misread marginalized groups [23]. Comparable problems can be found in the literature on credit scoring, where the place of residence can have a negative effect on creditworthiness and thus discriminate against people in selected regions [27, 195, 257, 402]. Also, the algorithms can be underminde due to the so-called data-behaviorism [527] by adjusting behavior accordingly to obtain the best possible algorithmic evaluation. In some cases, this can lead to bad behavior for society as a whole. In the example of P2P carsharing, for example, it is conceivable that some rentees avoid certain streets because there are many traffic lights and they would have to accelerate and brake there more frequently. This could lead to congestion on some roads and higher traffic volumes overall.

Further, there is a need of sufficient data to be able to calculate a algorithm-based score. As Bossauer et al. [83] pointed out, there is a trade-off between trust-building and privacy by using car telematics. It is not a self-evident fact that rentees disclose all possible information about their driving behavior from a privacy perspective. Nevertheless, rentees can be incentivized by certain added values and financial incentives to disclose their driving data. Such a score can only be applied if an added value is given, if the utilization of the data is comprehensible, or at least drivers can not be disadvantaged. [35, 83]. It is important here that the users or the data suppliers are involved in the technological processes and that these are also transparent and comprehensible. Here, it is up to the HCI designers to design the systems in such a way that, on the one hand, acceptance for the provision of data increases due to better transparency, but also an incentive is created to interact with the algorithms [251]. This applies to both car owners and rentees.

At the same time, a regular comparison of the algorithm-based scores with the real world is desirable, because this is the only way to counteract the algorithmic biases. For example, an option for the driver to comment on the driving behavior in certain situations would be important to ensure a fair evaluation. This could also increase the acceptance of such scores [527].

11.6.5 Limitations

Regarding the limitations, we have tried our best to exclude subjective influences by, e.g., using multiple coders [58]. In qualitative studies, however, statements are subjected to terminological tests, which means that all statements are subject to a personal interpretation by being assigned to certain categories. The selection of participants also allows only minor conclusions to be drawn about the understanding of algorithm-based reputation systems on the mass. The heterogeneity of the sample makes it possible to give a rough impression of how a small section of society understands scoring algorithms. Nevertheless, the results give a first impression of how people understand the functioning of an algorithm for evaluating driving behavior and how such algorithm-bases reputation systems may support trust-building and decision-making in the context of P2P carsharing.

Further, only one form of presentation of a Score (a number between 1 and 10) was used in the interviews for orientation purposes. This representation was chosen because ratings are the current standard for reputation systems in many areas of the sharing economy [602, 188, 697]. Regarding our prototype, our work is limited due to the fact that a prototype evaluation is just a study of 'what might be' [543]. Accordingly, there is a need to validate our findings in a real-world application scenario.

Developing a model for an AI was not the goal of this work, even though conjectures for a possible category system were made. No concrete model is used in this paper, so no model of measurement and evaluation for an Artificial Intelligence is given. In the paper, the participants assume a normative evaluation, so that the categories correspond to normative assessments. Therefore, in the context of a classification model of an artificial intelligence, the question arises how and which classes are generated and how the data flow into these classes. This can lead to conflicts of interest between humans and artificial intelligence, which need to be examined in more detail in future research.

11.7 Conclusion

The sharing economy has experienced a veritable boom in recent years. Many sharing models have emerged as a consequence. In our study, we focused on P2P carsharing which depends on the trust that needs to be built between peers. Sharing platform providers therefore make use of reputation systems, which address trust by sharing information between peers. A new form of trust-building on P2P sharing platforms can be algorithm-based reputation systems. To bring people closer to the benefits of algorithm-based reputation systems, they need to gain an intuitive understanding of their functional principle.

Based on 16 interviews in our pre-study, we gave insights into how people think such scores work and how they should be designed for a better understanding. In a second step, we developed a prototype of an algorithm-based reputation system for P2P carsharing based within a co-creation workshop. The evaluation with 12 participants shows that people mostly understand an algorithm-based reputation systems (Driving Score) in the context of P2P carsharing as a kind of digital monitoring of driving behavior. The results of the pre-study as well as the evaluation indicate that algorithm-based reputation systems can indeed support trust-building. It was also confirmed by the evaluation participants that a drill-down of a score considerably increases the benefit of it and thus promotes trust. Further, the possibility of weighting preferences within the scoring parameters as well as feedback for the implications of different parameter values would support system intelligibility and therefore also trust in technology [602]. Nevertheless, the existing literature on biases caused by algorithms and a corresponding biased reality can also be confirmed. Algorithm-based reputation systems are not free of biases, especially in the case of unpredictable events such as strong braking maneuvers to prevent rear-end collisions. Here, our study confirmed that certain tolerance limits and also the inclusion of the driver's perspective, e.g., through a individual statement in certain situations, are necessary for a reasonable evaluation. Such scenarios have to be considered in the design of the interaction between the users and the algorithm-based systems and fair mechanisms have to be incorporated [251, 527]. In future work, we want to explore algorithm-

based reputation systems in a real-world application and other scenarios, e.g., selecting a Uber driver for a ride. Furthermore, there is a need to extend our understanding of algorithm-based reputation systems for reducing discrimination in real-world situations.

12 Comparative Summary of Findings

12.1 Data Strategies Shaping Consumer Practices: Negotiation of Disclosure or Rejection through Reciprocity

From section 10, we recognize selective data sharing by consumers. If consumers have the necessary options and information, they weigh up exactly how and for what purpose they use their data. Interestingly, the results reveal a reciprocal behavior in which both providers and consumers take on the other perspective in the mutual negotiation process with transparent information about data collection and purpose and evaluate it accordingly. Accordingly, consumers disclose information if they understand its purpose and can comprehend it from the provider's perspective.

The identified principle of data minimization and proportionality makes it clear that consumers tend to adopt a data-saving strategy. Consumers consider data they are already used to collecting in other contexts (e.g., mileage, tank fill level, etc.) harmless and disclose it. In addition, in the case of comprehensible purpose limitation, which supports the provider, the willingness to disclose is further increased by adopting perspective. However, in the case of a user-centered coordination process, the results also show that providers tend to make more data-efficient and proportional classifications of necessity. For example, the disclosure of the GPS location is seen as critical by consumers, but in the voting process, a limitation took place that satisfied both sides. An approximate location was sufficient for providers to check that the contract was treating the vehicle. Consumers were able to understand the necessity of the survey for the provider. They agreed to disclose the data, including an exact location, in case of a border crossing beyond the agreements.

In addition to understanding the provider, the negotiation is also determined by possible benefits for the consumer. The resulting added value of the disclosure is evaluated and weighed. For example, it was stated that a critical dashcam recording would only be disclosed if it resulted in legal advantages for the consumer.

The findings show that negotiation support is needed. This support should

contribute to a common, informed understanding of the negotiation processes, accompanied by adopting perspectives. To this end, this process must be transparent and open, and data disclosure's purposes and added value must be clearly communicated.

The findings from the study in section 9 confirm this tendency towards selective disclosure of information among critical consumers. Consumers are willing to disclose sensitive data such as dietary preferences or shopping habits to develop a personalized food recommender system if they know the transparency of data use and the benefits for their purchasing decisions. This is particularly evident in the design of personalized profiles, which enable consumers to integrate their respective ethical values and personal restrictions, such as allergies, into the system to receive suitable product recommendations.

12.2 Human-Algorithm Interaction as an Expression of Consumer Practices

Building on section 10, the study in section 11 reveals that consumer decision-making practices to be supported by algorithm-based reputation systems, however, require a deeper understanding of the algorithms, including in terms of the trust placed in them.

Since data and algorithms act as a unit[6] [332], an understanding of algorithms is therefore also required for alignment with consumer practices. However, research shows that consumers only have a basic knowledge of algorithms, especially as they are usually hidden from them and therefore not transparent [488]. This is also confirmed by the study based on section 10 in section 11. The participants mainly chose road traffic regulations as an essential reference for describing their assessments of data for calculating a rating.

This study confirms prior research states in the perception of algorithms as elements of social order processes on regulative, normative, cognitive, and technological levels, whereby consumers mostly judge regulative (here accord-

[6]This is also confirmed by the von Neumann architecture as a reference model of computers, which stores instructions and data together[31]

ing to the road traffic regulations) considering their norms and values. The algorithm-based reputation system to support their decision-making should be customizable according to their ideas and requirements. To this end, the participants mention the interaction with the algorithm via customizable input variables to take their values and norms into account according to their ideas. Through the co-design approach and the derived customizable definitions of own values and norms in the algorithm, we show various design implications in section 11, which increase the understanding and usefulness of DDSs to support consumption practices.

Section 9 underlines the need for more understanding of algorithms due to a lack of transparency. Critical consumers also call for customizability as a path to a more understandable, practice-supporting solution that prioritizes their ethical values. Here, we also derive design implications that increase the understanding and usefulness of DDSs. The developed personalized profile system allows consumers to link individual preferences, such as vegan or sustainable diets, with their local shopping practices. This expands consumers' scope for decision-making and support for niche practices.

The studies' findings show that, like the data strategies described, DDSs require clear communication of the basis for decision-making and individual adaptation of the algorithms to build trust and encourage willingness to interact with them. Transparency and the possibility of individual control play a central role in empowering consumers to use data in accordance with their ethical convictions and required added value.

The following section 12.3 now compares the general findings for recommender and reputation systems in the respective fields of application.

12.3 Co-Creating Value, Acceptance, and Trust in Data-Driven Services

12.3.1 Individual Preferences and Flexibility Embedded in Situational Contexts and Actual Practices

In the food domain, section 9 showed that eating habits and dietary requirements are individual and cannot be differentiated into rigid categories. Many consumers pursue complex, possibly contradictory goals beyond simple categorizations such as 'vegan'. Accordingly, DDSs must offer a high degree of flexibility in recording personal preferences and restrictions. This enables a heterogeneous view of eating habits and requirements. In addition to considering personalized preferences, embedding in situational contexts and lived practices is also relevant. Dietary recommendations should, therefore, allow for simple adjustments to personal circumstances such as available budget, supplies, or shopping routines [475].

In the mobility domain, significant heterogeneity in perspectives and needs was revealed (see section 10 and 11). Vehicle rental companies want as much information as possible to secure their property, whereas tenants perceive intensive monitoring as an intrusion into their privacy. Algorithm-based reputation systems that process this data must bring these different requirements together and enable a high degree of customization. An algorithm-based reputation system must be able to evaluate situational contexts by integrating individual norms to ensure an individual assessment of the user's interests.

12.3.2 Trust through Transparency about Data Usage and Algorithm Logics

Especially when sensitive personal data is processed, DDSs must ensure user transparency and control. Furthermore, transparency about the underlying algorithms and decision-making logic must be created, as this is the only way users can understand and categorize the recommendations of food recommender systems [354], and classifications of algorithm-based reputation systems [441].

During the development of recommender systems for critical consumers in section 9, it became clear that false recommendations and incorrectly labeled data reduce trust in the recommender system. Reliability was vital in the tension between ethics, values, and restrictions [354].

In the context of mobility, the scoring of algorithm-based reputation systems must, therefore, be transparent [607]. Consumers want to understand how a driver evaluation works [83] and how the algorithms arrive at their results [441].

12.4 Implications for Support Instruments for Businesses

A user-centered approach should enable the individual adaptability of preferences and settings and the collaborative and user-based adaptation of DDSs. It was suggested that users should be able to define their own rules and ratings or use profiles provided by trusted third parties. Connecting to communities where consumers share experiences and product reviews would also strengthen trust in the systems.

Furthermore, DDSs should continuously adapt to changing needs and new findings. This can lead to long-term acceptance and appropriation of the services.

A tool to support businesses in developing and designing DDSs must capture an understanding of the relevant user target groups: Needs, limitations, everyday practices, and contextual factors.

Based on this, it should be possible to address flexible options for adapting and personalizing the services, both from the consumer and business perspective. The tool should, therefore, provide templates and best practices for modular, expandable information architectures, user-centered interaction design, and interfaces for integrating external knowledge.

Another crucial aspect is ensuring transparency and trust. The tool should provide guidance and design patterns to create trust through insights into data processing and algorithm logic. At the same time, mechanisms for handling sensitive personal data should be integrated.

Overall, the tool must provide a holistic, user-centered development concept. It should support businesses in creating DDSs that fit seamlessly into consumers' lives and practices, consider their needs, and enable innovative solutions to create digital offerings that consumers accept and use.

Part IV

Aligning Businesses and Consumers: Instruments to Support Data-Driven Value Creation

13 Introduction

Digital transformation has made data a decisive factor in business success. Businesses face the challenge of using data in such a way that it optimizes operational processes and efficiency, drives data-driven innovation, and enables new value-creation potential. As a result, it is necessary to go beyond purely technical solutions and strategically align the business strategy with the data strategy through Business-Data Alignment. Businesses can develop DDBMs and DDSs through the targeted use of data resources, which serve as the basis for companies' competitiveness and create new added value (see part II).

While Business-Data Alignment focuses on the internal alignment of corporate strategies with data, the role of consumers in DDSs should not be neglected. Business-Consumer Alignment should ensure that the DDSs developed by companies meet the needs and preferences of consumers. Consumers are not merely passive recipients of these DDSs but are active in the role of data providers who pursue their strategies in dealing with data. To shed light on the consumer perspective, part III focuses on Consumer-Data Alignment and examines how consumers can shape their data strategies and align them with their consumption practices and goals.

As outlined in section 2.4.2, businesses lack data literacy and, therefore, the ability to align business and data strategies, leverage their data assets, and create value with data. To align with consumers in the context of Business-Consumer Alignment to develop user-centered DDSs, tools to support data-driven value creation are needed, especially in the creation of data-enhanced business models, since such businesses transfer from traditional business models to DDBMs and further lack of experience in this respect. Furthermore, as sections 11 and 9 showed, consumers need to understand data and algorithms used to be actively involved in user-centered development.

Based on the findings from part II and III, a framework is therefore developed in section 14, which supports businesses in understanding, developing, and designing DDBMs and DDSs in the context of Business-Data Alignment. This framework also serves as the basis for communication in Business-

Consumer Alignment. Therefore, the clear and understandable communication of data and its meanings is also crucial to support alignment processes between businesses and consumers. Data storytelling serves to communicate complex data through narrative and visual elements, thus representing a bridge between technical development and communication with non-data-savvy stakeholders. In Business-Consumer Alignment, data storytelling is vital in helping companies and consumers understand the added value of DDSs and recognize their significance for individual and corporate decisions. To this end, section 15 is developing a process to support internal and external communication with data, which contributes to democratizing the development of DDBMs and DDSs through a common language and understanding.

Given the challenges highlighted in part III in the development of DDSs in the consumer domains of food and mobility, there is a need for holistic concepts that support companies in consistently placing the perspective and needs of users at the center. Therefore, a holistic data science canvas and a structured data storytelling process can provide a valuable framework.

The Data Science Canvas helps companies develop data-driven solutions that consider user needs and preferences. It guides developers through the critical aspects of such a project and encourages them to incorporate the consumer's perspective at every stage actively.

As demonstrated in part III, a deep understanding of users is a prerequisite for developing DDSs tailored to their needs. The canvas, therefore, suggests using user research methods such as interviews, quantitative surveys, etc., to characterize the target groups comprehensively [400].

- Identify and characterize user target groups
- Recording users' motivations, goals, restrictions, and everyday practices
- Analyze the realworld context and relevant decision-making situations

As the section 5 showed, bringing together data from various sources in integrated systems is a key challenge. The canvas supports companies in system-

atically building a comprehensive, high-quality database that appropriately reflects the context and preferences of users [178].

- Building a database
- Identify internal and external data sources
- Obtain, prepare, and integrate data
- Analyze data gaps and quality

Furthermore, classic algorithms often fail due to the specifics of consumer data [62]. The canvas offers the opportunity to discuss alternative approaches incorporating individualization and transparency.

- Select suitable algorithms
- Adapt and train models to user needs
- Promote transparency and explainability

It is shown that user-centered development and collaboration forms hold great potential for adapting DDSs to changing needs in the long term [686, 126]. The canvas should support integrating participatory, user-centered design into the development process.

- Establish communication and feedback channels for users
- Enable collaborative customization and expansion of services
- Continuous further development based on user needs

14 Data Science Canvas: Evaluation of a Tool to Manage Data Science Projects

Abstract

Data emerged as a central success factor for companies to benefit from digitization. However, the skills in successfully creating value from data – especially at the management level – are not always profound. To address this problem, several canvas models have already been designed. Canvas models are usually created to write down an idea in a structured way to promote transparency and traceability. However, some existing data science canvas models mainly address developers and are thus unsuitable for decision-makers and communication within interdisciplinary teams. Based on a literature review, we identified influencing factors that are essential for the success of data science projects. With the information gained, the Data Science Canvas was developed in an expert workshop and finally evaluated by practitioners to find out whether such an instrument could support data-driven value creation.

14.1 Introduction

In times of ever-increasing data amounts and the ascribed value to data, companies are challenged by creating value from it. Due to volume, velocity, and validity of big data, the toolbox around machine learning and mathematical-statistical procedures becomes ever more important. However, the implementation of such procedures into valuable business models requires expert knowledge [67]. In view of today's fast-moving markets and the prevailing competition, it is urgently necessary for decision-makers to evaluate the opportunities that arise from their data, the existing expertise in their company, and bring both together to create value [442]. To drive digital transformation for their organizations, decision-makers should be aware of how to scale the value of their data assets and the capabilities and opportunities of analytics [472]. However, according to a study by the Data Literacy Project ($n = 7,377$), only 24% of these decision-makers consider themselves to be

data competent, although this is perceived as increasing in work performance [501]. Despite its high relevance, the introduction of a data strategy has also not yet reached any maturity in many companies [140]. This requires a basic knowledge regarding data analysis as well as an active data-driven corporate culture to enable data-driven value creation and business models (DDBM) [252].

Therefore there is a practical need to provide decision-makers with tools to support data-driven value creation processes [333]. For other fast-developing and complex business problems, a concept for a quick implementation in the sense of 'Learning by Doing' emerged. For example, Lean Manufacturing [687] and Lean Startup [519] are supported by the Lean Canvas as an orientation basis. Another example is the 'Business Model Canvas' developed by Alexander Osterwalder, which simplifies the development and communication of a business model by helping to visualize and understand its elements [464]. The intention of a canvas model is to summarize a complex problem as clearly and simply as possible and give guidance for solution-finding.

Also for DDBM some canvas approaches, e.g. the Data Canvas [404] or the canvas for data-driven ideation workshops [330] were explored. While those focus rather on the data and algorithmic value creation [335], canvases such as the Data Insight Generator seek to more strongly include the customers value proposition [335, 333]. Still, this canvas only focuses on "an illustration between the [...] two components, key resources and value proposition" [335] rather than providing a holistic approach that considers the cost and revenue structure as well as stakeholders of DDBMs [562, 711].

Having identified the need to support decision-makers regarding DDBMs, which qualifies as a "heretofore unsolved and important business problem" [266] and the lack of a holistic canvas approach, the purpose of this paper is to communicate the design (first design cycle) and evaluation (proof-of-concept) of the Data Science Canvas – a tool to holistically support decision making and stakeholder communication on DDBMs. Our canvas (artifact), thereby, aims at being sufficiently specific and holistic at the same time to be readily used by managers. This means it covers both the technical and managerial aspects of the data problem at hand, to address the specified problem.

Table 14: Overview of the First Cycle of the DSR Project [479]

Identify Problem and Motivate (Section 1)	Empowering the Management to understand and create value with data and enable collaboration of different stakeholders of data-driven business models is an unresolved problem.
Define Objectives of a Solution (Section 1)	A holistic canvas concept is required to support management in understanding and making decisions of data-driven business models. Such canvas concepts have also been successfully employed in other fast-developing and complex domains.
Design and Development (Section 2)	The holistic Data Science Canvas is designed on the basis of a literature review on other canvas approaches and an expert workshop.
Demonstration and Evaluation (Section 3)	We conducted a real-world trial with 10 participants including a post-interview to prove the usefulness of our concept and understand the design space for further design cycles.
Communication	The Canvas description and the evaluation results are communicated in this research paper.

Thereby, our research follows the Design Science Research [266] approach as outlined by Peffers et al. [479], to first design the canvas as an artifact that aims to solve the problem and derive insights and theoretical implications [43]. As shown in Table 14 the remainder of the paper is structured along the phases outlines by Peffers et al. [479]. Having motivated the problem and defined objectives of a holistic solution in this section, section 2 reports on the design and development of our canvas (artifact). This section includes a literature review of related artifacts as a presentation of the related work [232] and an expert workshop to develop our prototype. In section 3, the canvas is demonstrated and evaluated on real-life data projects as a naturalistic ex-post evaluation [659] to understand usefulness in the context of real projects and decision making. The results are discussed in section 4 and we draw conclusions in section 5.

14.2 Data Science Canvas Design & Development

14.2.1 Literature Review: Identifying and Understanding Relevant Elements

14.2.1.1 Procedure. To highlight important elements and delimitations of the concept to be evaluated, a literature review was conducted according to the four phases of Rowley and Slack [528]. We searched for scientific or practical literature that aims to solve the identified problem by means of designing an artifact. As the terms data science, big data, and machine learning are often used interchangeably [497], we used the following keywords and their combinations: canvas, data science, data analytics, machine learning, big data, business model. Based on the initial literature, a forward and backward reference search was conducted [367, 676]. The main inclusion criteria for the surveyed literature was the provision of an artifact in the form of a canvas and a description of the elements to allow for comparison.

Building on the SQ3R-Method [518] to review the found papers, we identified nine canvas models, which were then recorded in the form of a concept matrix [542] using a conceptual framework [676].[7] As the identified canvas models used different wordings, we analyzed the tiles according to their meaning and role and clustered and renamed them to form a unified scheme. This procedure was conducted collaboratively by two authors to ensure agreement on the resulting elements. Resulting from this, the initial 20 factors were reduced to 12. As a first result of the literature research, Table 15 shows the basis of comparison as a concept matrix and allows a more detailed analysis of the canvas model elements.

14.2.1.2 Findings: Identified Elements The concept matrix in Table 15 shows that there are in principle many elements for a Data Science Canvas concept. However, some of these elements are only considered to a minor extent in literature, as the only occur in single sources. Data source, methods, model quality, added value, and stakeholders are among the most important elements, with four out of five mentions each. Data collection, data quality,

[7] A list of all identified canvas models can be provided on request.

Table 15: Concept Matrix.

Canvas-Models \ Elements	1	2	3	4	5	6	7	8	9	10	11	12
Big Data Management Canvas [27]			X		X	X		X			X	X
Deep Learning Canvas [28]	X			X		X	X		X		X	X
AI Canvas [29]	X		X			X			X	X		
AI Project Canvas [30]			X		X		X		X	X	X	X
Machine Learning Canvas [31]	X	X	X			X			X		X	X
Data Insight Generator [13]	X	X		X	X	X		X			X	
Data Canvas [11]		X	X									
Data Innovation Board [12]	X	X	X							X	X	X
Data Project Canvas [32]		X					X	X		X	X	X
Legend:	1: Business Case			4: Data Quality			7: Skills			10: Costs, Revenues		
	2: Data Collection			5: Data Integration			8: Software			11: Value Added		
	3: Data Source			6: Methods			9: Model Quality			12: Stakeholders		

and software for data analysis with only one entry each are not often considered. In the following, we more deeply discuss the elements.

Business Case (1): Although the definition of a business case is one of the crucial positions in the data science process, it is integrated into the canvas model as an element from five sources. In some cases, however, an exact definition of a business case is lacking. Thus, this element is often oriented towards the selection of methods and quality rather than the problem definition. Furthermore, used headings and questions supporting the answer are partly misleading and therefore not unambiguous. For example, the AI Canvas uses the heading *"Prediction"* with the supporting question *"What do you need to know in order to make the decision?"* [15]. Moreover, this mainly addresses the ML area and ignores other areas such as clustering or classification.

Data Collection and Source (2, 3): Following the objective delimitation of the business case, data is collected within the framework of its operationalization – in analogy to a statistical investigation – to obtain the required data. If this involves data still to be collected, a survey method and the necessary scale level must be specified. Both the survey method and the scaling are determined by the problem and the methodology [416]. Five approaches have such an element, but they do not sufficiently highlight the data requirements and potentially necessary cleansing needs [173, 404, 322].

To ensure the correctness and accuracy of the model, the source of the data should be analyzed. This is particularly important in the case of external data.

Six sources examined represent such an element, whereby only one canvas concept explicitly addresses internal and external data resources [15, 322].

Data Quality (4): By means of explorative analyses and visualizations, the quality of the data – whether already available or not – should be validated. This may reveal the need for data cleansing, as the presence of erroneous or distorted data could lead to erroneous models and management making the wrong decisions. This element is only rudimentarily and casually represented by two canvas models [481, 335].

Data Integration (5): Once the data quality has been checked, the data (from the different sources) should be merged into one database. This will facilitate the access by the data analyst and improve the quality of information through accessibility, completeness, clarity, accuracy, and consistency [271]. This element occurs in three sources, but is explicit and comprehensible only in two cases [305, 335].

Methods (6): The heart of Data Science is the transformation of data into knowledge [442]. The conversion takes place through algorithms which fall back on mathematical-statistical procedures. The central task of an analyst is to select the right method for the respective problem based on the available data. In principle, there is a wide range of statistical instruments (multivariate methods) for data analysis, but the special nature of the data must be taken into account [416].

Procedures for segmentation, classification, estimation (regression) and association can be differentiated. Different methods are ranging from white box statistical approaches to black box machine learning applications to address these problems [66].

Five canvas concepts have an element for method selection. However, they usually limit themselves to a certain method category (e.g. deep learning or machine learning methods) [173, 481]. Furthermore, no concept offers support in the choice of methods beyond the canvas. This makes it impossible for decision-makers to use the respective canvas without methodological knowledge.

Skills (7): The analyses requires skills from the fields of statistics, computer

science and mathematics. Skill sets to be defined are also suitable for a job description and advertisement for the respective activities corresponding to the business case. Against the background of insufficiently though-out and formulated job advertisements to a data scientist, this seems to make sense [552]. In addition to methodological knowledge (methods, tools and libraries, programming, databases), a basic understanding of business and economics as well as domain knowledge from the respective company should also be addressed to generate deeper insights [714].

Furthermore, the various roles in connection with data preparation and analysis should also be differentiated: The Data Scientist is mainly concerned with complex methods of data analysis, while the Data Engineer is responsible for software solutions for Big Data and, e.g., creates data pipelines. This differentiation is still missing in many companies, although even success could be endangered due to a lack of understanding of the distribution of tasks and roles [26]. Only three of the identified sources make it possible to deal with the required skills of the (required) personnel.

Software (8): Based on the method, the existing skills/personnel, and resources, appropriate software must be selected. The provided program libraries also play a role in this, to fall back on predefined and useful functions. The 19th KDnuggets Software Survey from 2018 identifies Python, RapidMiner, and R as the most popular software of the 2,052 software users asked [486], but also Excel is present in the top 5. From a managerial perspective, software is an important issue, as it directly relates to the cost of procuring software packages or services for training and running the models.

There are only three sources that have an element for mapping the software. However, there is no clear distinction between method and software selection nor are libraries, as a fine-granular decision, addressed [305].

Model Quality (9): Once the model has been developed, its quality must be assessed. Depending on the chosen method, different indicators are used. For classification problems, a confusion matrix is often used, which compares actual with predicted classification based on the existing classes [220]. In the case of regression-based methods, criteria such as the mean square error or the (adjusted) coefficient of determination are used [416].

The literature research showed that four canvas models allow an evaluation of the model. One concept event differentiates evaluations before and during the use of the model. If the models are working in real-time, such a distinction proves to be useful. However, there is no further explanation or clarification of the indicators.

Costs and Revenues (10): In the business context, costs and revenues are key variables, especially for justifying (data analysis) projects to top-level management. In analogy to the Business Model Canvas, the costs and their structure on the one hand, and the possible revenue streams of the project on the other hand are listed here. Four of the sources have elements that include costs and revenues. One analyses only the costs with and without the model, while the other differs concretely costs and incomes [15, 703].

Value Added (11): This element also originates from the Business Model Canvas and is intended to show how the model can be used to support the internal and/or external customers [464]. Without a prior specification of the added value, any project is questionable. Seven sources have such an element.

Stakeholders (12): Finally, the results of the analysis and the resulting added value must be presented to the stakeholders involved in the project. If the model is implemented in the company, it is also important to support them in dealing with the model. For target-oriented control, the ability of Data Storytelling is also required here. Data Storytelling stands for the preparation and presentation of data in the form of a scenario within a story, tailored to the target group. This makes it easier for the target group to understand the meaning behind the data and motivates action [177, 443].

The analyzed canvas concepts underline the relevance of the consideration of stakeholders in a canvas model. Six concepts have such an element. Often, however, the requirements of the respective target group for the presentation and communication of the results are not mapped. In analogy to the various notation methods and their specific depth in process modeling, it makes sense to use a different representation, history, and complexity for the top-level management than for the specialist departments and their users [90].

14.2.1.3 Findings: Relation of Elements. To understand the relation of our elements and add a procedural view, we mapped them to a data science process (cf. Figure 22). The relation is thereby very close to the iterative processes as identified by Cielen et al. [131] but also CRISP-DM [685].

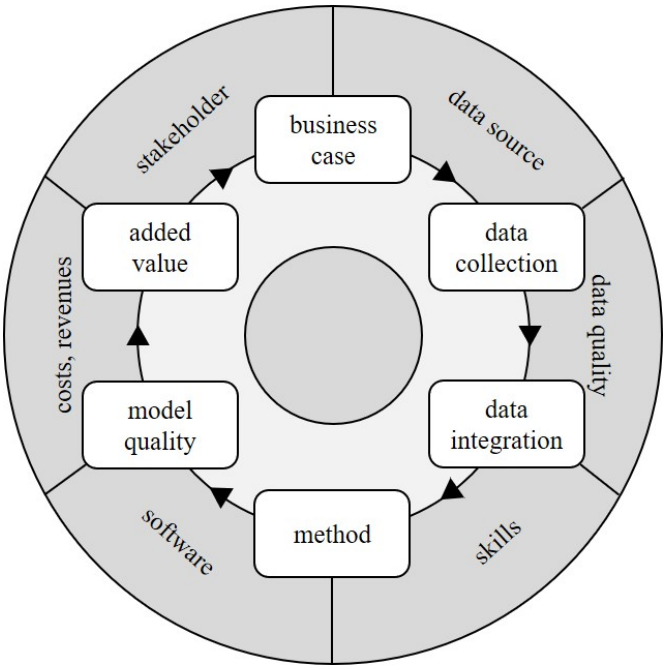


Figure 22: Data Science Process.

As a starting point and core of a new product or service, the business case represents a strategic decision to focus data analysis on a specific business objective. At the same time, this also implies an orientation to the requirements and needs of the customer (stakeholder) and commits to continuous measurement and optimization [624].

The data required to map the business case may not yet be available or its quality (data quality) may not be acceptable. Therefore, the data must (additionally) be collected (data collection), possibly cleaned up, and integrated into an existing system (data integration). To ensure sustainable quality, the data source should also be evaluated [271, 88].

If the required quality and quantity of the data are given, a model is to be developed according to the problem definition of the business case, which supplies the demanded result through a method of data analysis. Often a

software is used, which can turn out differently depending upon the question, availability, and existing competencies. Furthermore, the analytical skills of the analyst are required, which should be specified beforehand. This can also be helpful for the exact definition of a position [157].

Finally, the model quality measures the adaptability to the respective data for the evaluation of the model and describes its ability to apply these data to the business case. Depending on the method chosen, different quality criteria have to be applied [416].

A cost-benefit-analysis and (expected) added values should be highlighted for project prioritization, as they improve the possibilities of management and controlling. Due to result-oriented companies, many stakeholders are particularly interested in the resulting added value. The challenge is to present the right information to the right target group and thus avoid complexity. This is where Data Storytelling plays a significant role [443].

14.2.2 Expert Workshop

14.2.2.1 Procedure. Based on the results of the literature research, it becomes clear that most canvas concepts are more technically oriented, neglecting the cost and revenue structure or even the skills needed in the organization. In many cases, the derivation of the problem via a business case is missing. No single canvas provides additional support for the choice of methods and software, and even the completion of supporting questions is not always available.

Building on these results, we developed a model that can be used for any data analysis problem. Our canvas model should address non-data-affine managers and executives to make data science projects more plannable, controllable, and manageable. This allows companies to gain an overview of their resources and the utility of their data. The model should further help to evaluate the use of such a canvas in practice.

For the development of the first version of our Data Science Canvas, we used the design thinking process by Brown and Katz [94]. In the beginning, the problem space and the interest groups were defined. With the help of a work-

shop on data analysis with eight researchers and scientific experts from Germany from the fields of informatics, statistics, mathematics, economics, and business administration, it was examined how they proceed with the analysis of data and what problems they encounter. Researchers act at the transition point between science and practice by communicating research data and project results and, thus, can argue from both perspectives.

The insights gained were processed and served to define the viewpoints of the interest groups (managers, specialists, users). Further they were used in the idea generation phase to discuss the instantiation of a canvas through group dynamics and with discussion-enhancing "How can we...?" questions. The elements were therefore placed on a whiteboard via sticky notes and their meaning and arrangement were critically discussed by the group.

The interim result of the workshop was an artifact, which was finally tested and evaluated by each participant based on a use case in the form of a project to optimize a sustainable energy supply of a production plant. One of the main findings was that selecting the appropriate method still encounters problems. In comparison to the identified results of the literature review, explorative data analysis and model requirements were here mentioned as additional elements. The insights gained in this process were incorporated into the actual version of our Data Science Canvas (see Figure 23).

14.2.2.2 The Data Science Canvas. Consisting of three parts, our Data Science Canvas addresses the different aspects of data analysis: The dark grey elements are dedicated to the definition of the (business) problem (**Problem Statement**). The light grey fields deal with the collection and preparation of data for the model (**Data Collection and Preparation**). Finally, the white elements describe the implementation and evaluation of the previously derived method (**Execution and Evaluation**).

Users can access the Data Science Canvas basically from two different directions: On the one hand, it allows them to start by defining the business case via the data collection and preparation to derive and evaluate a model. On the other hand, it is possible to start from the data collection and preparation perspective (e.g. in the context of a research project) to derive an application

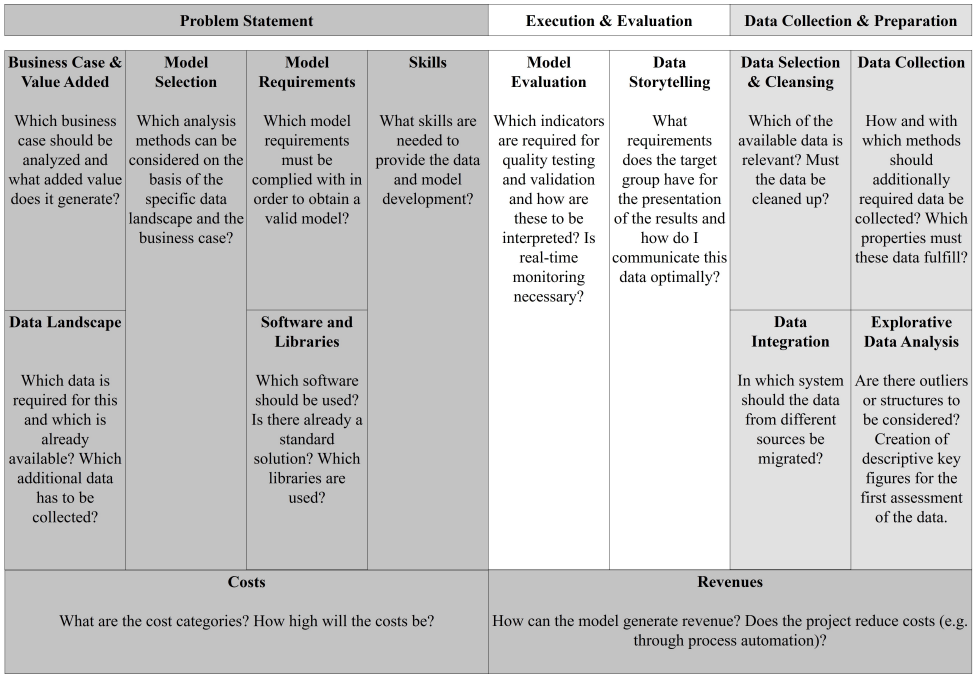


Figure 23: Data Science Canvas.

case which leads to a model.

Problem statement: Based on the business case and the expected added value, the already existing data landscape of the company is to be analyzed. The actual data situation shows whether additional (external) data is required under certain circumstances. The information about the existing data and the application case is used to select the specific model. If an analysis method was chosen, the requirements of the model (e.g. regarding scaling) should be noted. The software used is also influenced by the choice of model. Thus, several software tools are usually suitable, depending on the existing resources in the company as well as the skills to be defined subsequently. These are equally determined by the method used. For meaningful resource planning and project prioritization, the costs and revenues are also defined. In the case of values that are difficult to estimate, an enumeration of possible cost and revenue categories can also be helpful.

Data Collection and Preparation: If not all the data required for the model are available, additional data must be collected. This can be achieved using

various methods and sources (e.g. surveys, open data, etc.). The requirements of the model, e.g. for the scale level, must be considered. Following their collection, the quality of the data must be checked. Data inconsistencies and outliers can be quickly identified using descriptive statistics and visualizations. Furthermore, meaningful aggregations of the data to key figures (e.g. mean values, scatter, and correlations measures) help to gain a deeper understanding of the circumstances to be modelled. If it is necessary to clean up the data (e.g. due to outliers or missing values), the required measures are documented (e.g. interpolation of missing values) and the data relevant to the model is selected. The split into training and test data can also be carried out at this point. If there are (no longer) any quality deficits, it should be clarified into which system the data possibly originating from different sources are to be integrated, so that the model can access them flawlessly. In principle, the tasks of data collection and preparation should be performed by a data engineer.

Execution and Evaluation: Finally, during the execution and evaluation phase, the model is evaluated using specific quality indicators that depend on the method. For some models it is also helpful if continuous monitoring is carried out in real time (e.g. in the case of predictive maintenance solutions). The remaining element can be used to present the stakeholders' requirements for the presentation of the data, its further communication, and the use of the model. The element should support a target group oriented communication of the results through data storytelling to create understanding, motivation, and change.

14.3 Demonstration & Evaluation

14.3.1 Data Collection and Analysis

To conduct an ex-post naturalistic evaluation [659], we introduced the canvas to 10 managers and data science practitioners to test and evaluate it in their organizations (cf. Table 16). We then interviewed all of them along a semi-structured interview guideline, to understand and evaluate the canvas in terms of its suitability and usefulness for empowering managerial-decisions

Table 16: Overview of Participants

ID	Age, Gender	Branch	Job
F01	38, f	Media	Business Development
F02	42, m	Engineering	Consultant
F03	41, f	Automotive	Project Manager
F04	35, m	Chemical	Developer
F05	37, m	IT	Data Scientist
F06	48, m	Manufacturing	Procurement Manager
F07	40, m	Trade	CDO, Head of IT
F08	43, m	Project Sponsoring	Head of Software Development
F09	29, f	University	Researcher
F10	27, f	University	Researcher

on DDBMs and effective stakeholder communications. In the case of the canvas artifact, a numbering of the elements was explicitly left out in the sense of a procedure to provide scope for own layouts and processes.

We have analyzed the data based on the thematic analysis approach, as described by Brown et al. [87]. We used MAXQDA to code the transcribed interviews and in an open coding procedure, the codes were developed and discussed in the team of researchers to derive the relevant themes. We ended this process after all researchers agreed on the low probability of finding additional themes. The results of the practical evaluation interviews can be divided into three categories: Use cases, application problems, and design implications. They are presented in the following chapter.

14.3.2 Findings

14.3.2.1 Use Cases of the Data Science Canvas. In general, almost all participants accepted the canvas positively and addressed various use cases. The canvas was often perceived as a structuring element, but also as a checklist to describe progress or the current situation. This is also discussed in the context of the respective target group. For example, a manager can use it to inform about a new project and as a checklist for **progress review**. Within

the project team, the canvas can be used to structure the project during the planning phase and to point out problems that need to be solved.

"I find many of the points in the canvas very useful, and I will surely use them for my further work, especially with regard to the evaluation of the data situation and the data still needed, I find this extremely helpful and would pass it on to colleagues [...] and tell them, please: Work on these points." –[F08]

"Well, I think that's a good basis for argumentation to structure one's project. That you can think about how to handle it. [...] Or I could think of it as a checklist to check off during the project to see if it's there." –[F03]

It was also mentioned that the canvas can be used as an **instrument of trust** to make it clear to the manager that the previous preparation has been dealt with in a structured way.

"Such a canvas can also create trust with a supervisor by showing that you have dealt with it in a structured way, with your project. With my colleagues I would consider it more exploratory, for my boss it would be more like a check-up." –[F09]

"Another extreme value is that you can show a booth by showing that you can see: "I've just got that far and this and that needs to be done." –[F10]

There is also an application for cross-departmental projects to solve silo thinking through the structured and **interdisciplinary communication** of the topic or to identify the resulting problems.

"It also allows you to discuss with different actors across departments and to address overlapping problems and recognize them in the run-up to a project." –[F08]

A participant from the area of research also notes that she likes the fact that a problem can be approached from different directions. So on the one hand, if the data is available and something is to be investigated, and on the other hand, from a research perspective: I still want to collect the data and I can deal with the model requirements beforehand.

"I think it's good that it can also be filled out from a research perspective, that the problem can also be approached through data collection." –[F09]

Another use case is that the Canvas can help to define a specific **job description**.

"You can use this to have a better understanding of requirements for the definition of a position, what you are lacking in skills." –[F03]

Two participants also see data storytelling as an important factor to communicate results and progression to stakeholders.

"Data storytelling is then the customer's viewpoint, with which you can communicate the revenue again. It's important that it's communicated correctly." –[F05]

"I also like data storytelling, how it is prepared for the target group. It also helps to think in terms of goals, that you don't forget anything." –[F09]

14.3.2.2 Limitations of Applicability. During the interviews, however, problems regarding the use of such a canvas also became apparent. These mainly address interface problems with other departments, the scope of a canvas at the beginning of a project, and a general understanding of the elements.

Some participants pointed out that such a canvas is difficult to fill in at some points, especially at the **beginning of the planning phase**. The majority, however, emphasizes the added value addressed in the use cases. The Data Science Canvas provides a good overview of the implementation and weaknesses of a project, as well as of the resources still lacking.

"Well, I think that especially in the planning phase, where not much is known, that it was a bit difficult to fill in completely there." –[F06]

In this context, the problem of the **interdisciplinarity** of some projects is addressed. The canvas is also perceived positively here, to establish structure and overview. Especially in non-data-driven companies, there might be little understanding of cooperation in this area.

"A project manager working in the problem-space is not necessarily interested in data quality, at least not yet. But maybe in the future and the canvas could serve as an opener if everyone understands that data-driven business is important. But unfortunately, there are still too few of these people today."

–[F08]

"Perhaps you should be able to specify which department the data comes from. Especially if you don't have a common database or silo thinking it might be important to have at least one contact person. I often had problems with that in practice." –[F05]

A further problem of use was seen in the **choice of methods**. Many in the project may have no experience with it and need support in assessing the use of a specific model.

"The choice of a method is difficult for me. I don't know anything about these things. I can deal with the data, but I find it difficult to understand how I evaluate it. Perhaps recommendations could be made." –[F04]

It was generally suggested that there should be **more detailed guidance** on such a canvas model in addition to the supporting questions, also with a view to a common understanding.

"I always think it makes sense for the canvas models to come with a little manual or something, so that everyone understands the same thing." –[F01]

14.3.2.3 Further Improvements. The design of the Data Science Canvas was widely regarded as very extensive. On the one hand, this helped to ensure a structured and comprehensive approach to the problem, on the other hand, it required a more intensive study of the instrument. This also reflected the demand for **guidance for a common understanding**.

"[I think] the process is a bit difficult, it should at least be numbered in terms of phases." –[F06]

"So after I filled out the canvas, I noticed that I would arrange it differently, I would sort of design the structure from left to right, so first problem statement, then data collection and then execution. so it would have a clearer flow." –[F05]

It was also noted that a **differentiation into existing and still needed skills** would be useful. This could help to fill in the costs to be able to plan more precisely and also to create job profiles.

"One could perhaps still divide skills into skills that we already have and skills that we still need. This is certainly also relevant in practice. You can then apply that directly to the cost of what you would have to buy." –[F10]

It was also mentioned that a time component should be queried to better estimate the costs.

"You ask here already, must the data be cleaned up. Maybe. If I were to ask, how much longer or is it too much work?" –[F09]

14.4 Discussion & Implications

14.4.1 Practical Implications

The predominant contribution of this research for practice is the provision of a holistic Data Science Canvas that covers both technological and managerial aspects. Regarding the use of a Data Science Canvas model in practice, it becomes clear that it fits for several functions. On the one hand, it can be used as a structuring tool to discuss a project deeply by analyzing the current organizational situation and, thus, it is possible to identify problems that have to be solved beforehand. If the elements were filled out, the interviewees see also a use case regarding a communication instrument to provide trust in the project team's work as well as to grant a common understanding of the goals and challenges of the project.

On the other hand, it can also be used as a checklist for monitoring the project's progress by getting visualized what is done and what is left, to be able to intervene if it is necessary. Furthermore, there is a possibility to start breaking up silo-mentality by addressing and discussing topics that are interesting for both stakeholders Nevertheless, there are still problems left that need to be considered in further developments of the Data Science Canvas to make it more relevant to practical applicability.

14.4.2 Theoretical Implications

From a theoretical perspective, our research contributes to literature through the following insights: First, in line with the practical contribution, our research provides an overview of yet existing data science canvas approaches and their respective elements. Furthermore, we combine those elements into a holistic canvas and discuss their relations.

An essential insight that contributes to design theory of data science canvases, is that such an artifact can be seen as both a structuring tool and as a checklist, depending on the phase of the project. This information can serve regarding a necessary adaption of the canvas structure by maybe offering different versions — one in the manner of a canvas and one as a checklist. If a canvas model is provided as software, this could simplify the filling out, because only the canvas needs to be prepared and the checklist is generated automatically.

For a common understanding, guidance can be helpful, which should be additionally available. To make the selection of a method more comprehensible for those who do not have an affinity for data, supporting small cards with the essential requirements and possibilities of the respective method could be offered. This could also improve interdisciplinary cooperation by giving everyone an equal understanding of the method. This also essentially supports the purpose as a structuring element of a project and shows where a participant still has to work on.

To improve the design, the canvas should further be numbered according to the elements or described accordingly in a manual. In our case, the numbering therefore should be presented differently depending on the initial situation in a manual, so that several applications are possible.

It is also an important finding that supporting questions time requirements should be addressed and queried to ensure a better presentation of the cost and benefit view and to support the decision for or against such a project. Furthermore, skills should be differentiated into those that are already available and those that are still needed in terms of personnel. This will also allow better planning of personnel resources. This can also help with job definitions, e.g. by making it clear if a data scientist or data engineer is required.

14.5 Conclusion

According to the Data Literacy Project study, data literacy is often poorly represented among decision-makers [501]. Regarding the central research question, a concept like the Data Science Canvas can therefore contribute to the success of a data science project through the clearly arranged presentation of relevant influencing factors and help to democratize Data Science and thus support digitization by making it accessible for non-data-affine project managers. According to the practical evaluation it serves as a communication base between the different project stakeholders and can be seen as a structuring element or checklist before, during, and after a project.

But the evaluation interviews also showed that there are some problems regarding the understanding, scope, and problems resulting from interdisciplinarity in projects. Therefore a Canvas concept should consider this information and solve the derived problems.

15 Data Storytelling as a Critical Success Factor of Data Science[8]

Abstract

Due to the ongoing digitalization and the Big Data trend, an increasing amount of data is available. This results in many potentials—especially for companies. The ability to cope with and evaluate this data is reflected in the role of the data scientist, which is one of the most popular jobs at present. However, challenges arise from the integration of data into corporate strategy and culture. For example, complex data and analysis results must be communicated to stakeholders who are not data-affine. Data storytelling plays a decisive role here, because to use data to initiate change, understanding and motivation for the issue must first be created for every target group. However, data storytelling is still a niche topic. This article uses a systematic literature analysis to derive the success factors of data storytelling for an effective and efficient communication of data to support Data Scientists in research and practice in communicating data and results.

15.1 Introduction

> ”Data does not go viral. Stories do.” [293]

Storytelling has always been an essential part of our society. Not without reason did UNESCO include storytelling, which is just one part of the narrative arts, in the list of intangible cultural heritage [654]. This is also evident, among other things, in the current discourse on the climate and coronavirus crisis, which demonstrates the relevance of scientific results and their communication as a basis for social and political debates and decisions [32, 361, 615].

[8]This paper was originally published in German and has been translated into English for this thesis. The German source is: Neifer, T., Lawo, D., Bossauer, P., et al. ”Data Storytelling as a Critical Success Factor of Data Science.” HMD, 57, 1033-1046 (2020). https://doi.org/10.1365/s40702-020-00662-3.

However, good storytelling is not something people are born with. Instead, it is a skill that must be learned to be able to capture and hold the attention of one's audience [155]. Stories help people understand complex issues and evoke emotions. As a result, they remain in the memory and make it possible to use the insights gained and pass on impressions.

Telling the story of complex issues is also essential in business and research. On the one hand, ever larger volumes of data and more complex analysis methods are opening up new possibilities for gathering facts. Still, on the other hand, increasing complexity also increases the challenges of evaluating these volumes of data and interpreting, understanding, and communicating the results [176, 272, 443].

Against this background, the ability of data storytelling, i.e., communicating data effectively and efficiently, also gains central relevance in the corporate context, e.g., internal corporate research [394, 442]. Similarly, in science, as is evident in the COVID-19 pandemic and climate change communications. Here, an increased need for comprehensible and citizen-oriented communication of complex data evaluations and their predictions is demonstrated. In both contexts, the data scientist or researcher - in addition to their unique role as an analyst of complex contexts - also has the task of communicating to decision-makers [32, 128, 705].

Successful data storytelling requires two basic skills, which, according to Nussbaumer Knaflic (2015), are already taught in primary and secondary education [319]. Mathematics is needed to grasp, evaluate, and interpret numbers. Language is then used to transform numbers into words, words into sentences, and finally, these into stories.

Despite the importance of data storytelling for business decisions [394] and science communication [705], it is still a niche topic in the scientific literature. While storytelling is repeatedly listed as an essential tool and was even listed in CIO Magazine's 2017 Creativity Trends [664], the factors for data storytelling success have not yet been systematically studied.

Based on a systematic literature review, this paper identifies eight key success factors and derives a process approach for effective and efficient data

storytelling to implement a sustainable culture of data-driven communication because ”Data gives you the what, but humans know the why.” [69].

15.2 Data Storytelling

The term *Data Storytelling* is composed of the terms *Data* and *Storytelling*. While *Data* represents the numbers and signs that have to be analyzed, processed, and communicated for decision-making [208], *Storytelling* is defined as the ”art of [telling a story]” [388].

Klaus (2019) takes up these definitions and describes data storytelling as a communication method for ”relevante Entwicklungen verständlich zu kommunizieren mit dem Ziel, notwendige Entscheidungen zu treffen.” [relevant developments [...] with the goal of making necessary decisions.] [316]. Furthermore, he describes the role of the storyteller as an interpreter of data analysis into a language that everyone can understand [316]. Bladt and Filbin (2013) follow this exposition and understand data storytelling as ”humanizing” the data [69]. In addition to these attempted definitions, no clear definition of the term has emerged in the literature. However, based on the different perspectives and highlighted aspects [535, 662, 665, 177] of data storytelling, it can be understood as *process of preparing and presenting information from the results of a data analysis to motivate a decision or action in a language and visualization appropriate to the target audience*. From this, three important elements of data storytelling already emerge: the data or data analysis as well as its target group-oriented communication by means of a linguistic story and comprehensible visualizations. To tell a relevant story with corporate or research data, it is necessary to understand and observe all three key elements of data storytelling (see Fig. 24).

This combination allows the data to be given context and thus meaning in terms of future action. By visualizing the data, audiences can gain insights and be enlightened to new perspectives and opportunities. Connecting the story with visuals also has benefits: If a narrative is successfully supported with visual components, the audience can be entertained and motivated [319]. The goal of data storytelling to prepare relevant decisions and initiate change

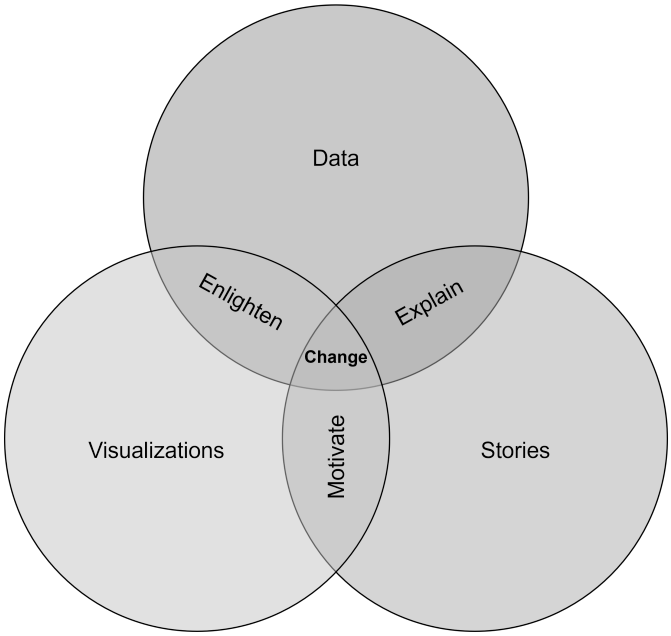


Figure 24: The Three Key Elements of Data Storytelling according to Dykes (2016) [177]

is best realized by linking all three components [177].

The data story does not differ from other stories in terms of structure. It usually consists of the initial situation (”in the past”), the climax and/or turning point (”then something happens”), and the end (”as a result”) [103]. According to Ryan (2016), a distinction can be made between eight plot frameworks of a data story [535] (see Tab. 17).

According to Segel and Heer (2010), seven genres of audiovisual representation of data continue to be distinguished at the intersection of ”story” and ”visualization” [564] (see Fig. 25).

The magazine style represents embedding a single illustration in a page of text, while a comic strip has many illustrations in a linear storyline. Comics offer the advantage, especially in printed media, of providing a data story with a lot of design space and expressiveness (sequences with images, words, and narrative elements) [39]. Partitioning posters, on the other hand, as they are often used in scientific contexts, usually feature a loose arrangement of

Plot	Description
Change over time	A visual story is created by a trend or a simple metric
Drill-down	Start with the big picture and work your way to the key message
Zoom-out	Inferring from the individual to larger groups
Contrast	Comparison of two sides: ”This” or ”That”
Distribution	Both the positive and the negative sides should be considered and the range of the data should be clarified
Transitions	Things that overlap or change from ”less than” to ”more than”
Factors	Two or more elements whose effect multiplies in combination
Special cases	An effective means is to point out something outside the norm

Table 17: Story Scaffolds of Data Stories according to Ryan (2016) [535]

images, this is especially exciting for easily condensable information to avoid countless presentation slides or to present several aspects divided and clearly arranged on one poster. An annotated diagram, as a simple form of image-language relationship, achieves added value by enriching familiar diagram types, which are supplemented with explanatory text. Flowcharts support a processual understanding concerning the connection between data and events. In the context of entrepreneurial, but also research presentations, instead of a comic strip, the screen presentation is mostly used to prepare a data story and to provide it with narrative elements. The presentation via films and videos is typically used in television commercials or explanatory videos [564].

Data storytelling can have an informative, convincing, entertaining, calming, explaining, or frightening character [308, 359] (see Tab 18).

Ojo and Heravi (2018) tried to explain which patterns and structures successful data stories follow based on successful data stories from practice (see Tab. 19). They also analyzed which technologies and tools are used in the process [456].

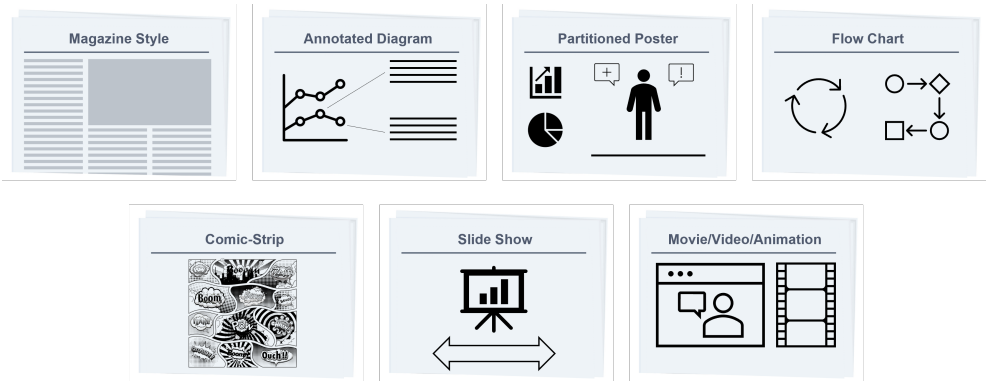


Figure 25: Genres of Narrative Visualization according to Segel and Heer (2010) [564]

Purpose	Example
Inform	Clarify about a situation based on current data
Convince	Demonstrating the effectiveness of measures
Entertain	Information of personal interest to stakeholders creates awareness
Calm	Present data that shows continuous development in turbulent times
Explain	Create a deeper understanding of a phenomenon based on detailed facts
Frighten	Show consequences to motivate action

Table 18: Different Purposes of Data Stories

15.3 Methodology

To get an overview of the success factors of data storytelling, a systematic literature analysis, according to the four phases of Rowley and Slack (2004) [528], was conducted (see Tab. 20).

Element	The most frequent Data Story aspects (in %)
Purpose	Inform (73)
	Convince (41)
	Explain (39)
Interactivity	Interactive (59)
	Possibility of searching, filtering, and selecting data (27)
	Static graphics (9)
Illustration	Annotated Graphic (77)
	Various / Combination (18)
	Video (18)
	Magazine style (14)
	Images (11)
	Slideshow (7)

Table 19: Important Patterns of Successful Data Stories according to Ojo and Heravi (2018) [456]

Phase	Procedure
1. Search and evaluation of the literature	Searching databases for research on data storytelling as well as literature with its practical application & quality assessment according to Cooper (1988) [143]
2. Filtering and retaining relevant information	Analyze and summarize the literature using the SQ3R (Survey, Question, Read, Recall, Review) process to identify essential information and literature [518]
3. development of a conceptual framework	Creation of a concept matrix to sort and rank the identified literature [676]
4. Formulating the Literature study	Derivation, critical appraisal, and documentation of a theory from the previously gained knowledge

Table 20: Four Phases of the Conducted Literature Study

For this purpose, in the first phase of the literature study, research results and practice-relevant articles were searched for using keywords such as ”data storytelling”, ”storytelling” or ”data science”. According to Cooper’s (1988) quality criteria [143] and a review of the abstracts regarding relevance, these were pre-filtered so that 24 relevant articles were available.

Building on this selection of literature, the second phase drew on the SQ3R (Survey, Question, Read, Recall, Review) process to identify essential lit-

erature [518]. Further selection focused on research findings, practices, or applications with a neutral perspective and the goal of generalization [143], extracting the ten most relevant articles from 24 sources.

After the appropriate sources of information were selected in the first two steps of the SQ3R process (Survey, Question), the three Rs were used (Read, Recall, Review) to identify the relevant success factors. For this purpose, a concept matrix was created [542, 676], which shows the results of the literature analysis in relation to the success factors (see Tab. 21). In this step, similar success factors were also summarized so that eight factors could be identified.

Finally, in phase four, the results were compiled into an overview of data storytelling success factors.

Source	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**
Nussbaumer Knaflic (2015) [319]	X	X	–	–	X	–	X	–
Klaus (2019) [316]	X	–	X	X	–	–	–	–
Bladt and Filbin (2013) [69]	X	X	–	X	–	–	–	–
Stone (2015) [614]	X	X	–	–	X	X	–	–
Morgan (2016) [429]	–	–	–	–	X	X	–	X
Wellington (2015) [679]	–	X	–	–	X	X	X	–
Samuel (2015) [544]	X	X	X	X	X	–	X	–
Heeg (2015) [260]	X	–	–	–	X	X	X	–
Pyczak (2017) [498]	–	X	–	–	X	X	X	X
Ryan (2016) [535]	X	–	–	X	X	–	–	–

A: Visualize data, *B*: Focus on the core idea, *C*: Motivate to act, *D*: Select the right data, *E*: Avoid complexity, *F*: Adapt Data Story to the audience, *G*: Telling the adequate story correctly, *H*: Develop story together

Table 21: Concept Matrix

15.4 The Success Factors of Data Storytelling

In the following, the success factors (see Tab. 21) are listed according to their frequency, described in more detail, and classified using an overarching example. The example describes the results of an app usage analysis based on log

files, which are to be processed by a researcher and presented for analyzing the app's target group (see Fig. 26).

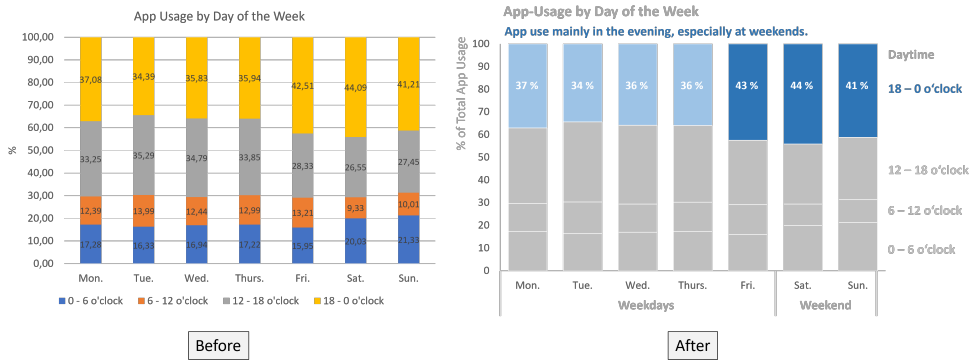


Figure 26: Analysis of App Usage to Derive the Relevant Target Group

15.4.1 Avoid Complexity

The most frequently mentioned factor for a successful data story is the avoidance of complexity. Nussbaumer Knaflic (2015) states that complexity can be reduced using visual aids. In general, care should be taken to ensure that a data story is easy to read and simply told. She substantiates this with the fact that listeners quickly switch off when explanations are presented in a way that is too complicated, and they feel uncomfortable or even become annoyed if the language is unnecessarily complex [319]. Also, complex visualizations ensure that readers develop their own story according to their understanding and thus may deviate from the desired purpose [614]. Illustrations and text, then, should include only the most important data and trends for your story. Rather than flooding the reader with all the data, it should be used selectively to support the message [429]. Implicitly, the argumentation for finding a suitable framework of the data story for the respective target group can already be recognized in some authors [260, 679]. In all cases, it is helpful to obtain interdisciplinary feedback beforehand [319]. Fig. 26 illustrates that better coloring brings the relevant data to the forefront, reducing complexity. The eye is drawn directly to the essential information and comprehension is strengthened by the subtitle.

15.4.2 Visualize Data

Visualizing data goes hand in hand with reducing complexity and is mentioned by seven sources. According to Bladt and Filbin (2013), data are more difficult to contextualize. For example, using their example, they cite that presenting demographic and geographic trends in the context of regression tables is challenging to understand, especially when there are multiple characteristics with possible interdependencies. However, when they resorted to a visual presentation of the data in the form of maps, the audience could quickly grasp the data, stimulating discussion and enabling even non-data-savvy viewers to uncover trends [69]. However, graphics should be as self-explanatory as possible. Further, visualization should be viewed only as a supporting element. Images do not replace the story but guide it to the climax [316]. This is also emphasized by Heeg (2015), who provides some examples from practice. He describes that the derivation of insights from large amounts of data and their communication were not successful, especially when the data visualization was not perceived as part of the data storytelling [260]. Data visualizations should accordingly be actively integrated into the data storytelling process. Visualizing numbers in the context of the after-image of Fig. 26 improves data storytelling by making it self-explanatory and supporting the data story by focusing on the main insight. This clearly identifies a trend in terms of increased usage on weekends. This is not immediately apparent in the context of the before figure.

15.4.3 Focus on the Core Idea

Six of the studied sources advocate focusing on the story's core idea and building the story around it. Data should always be presented to address the question to be answered using the data [498]. Regardless of the underlying data, conveying a clear key message within the final data story is relevant. This can be achieved by presenting the intent at the outset and then building on it linearly. It is also helpful to reflect on how the story could be summarized in one sentence [679, 544]. In the example, the analysis could be used to push marketing measures, which here - due to the mainly evening usage - could be

done via an advertisement on a streaming service, since the target audience mainly watches a series in the evening. This connection is directly evident from the ”After”-image (see Fig. 26) and can be discussed. Thus, it could also make sense to address a family audience to expand the target group in the 6-12 o’clock usage range.

15.4.4 Adapt Data Story to the Audience

Only half of the sources explicitly made this point. However, the adaptation of the story to the audience is emphasized in the context of these contributions. Due to one-size-fits-all presentations, it so often happens that the preparation makes sense to the analysts but not to the audience. It is essential to know your audience and adapt the story accordingly. Connecting people to what the data is saying meaningfully helps overcome the barrier between data analysis and decision-making. For this, it is recommended to ask someone from outside the project for understanding [260, 429, 498]. Again, the after plot of Fig. 26 advantages over raw numbers as well as the before diagram. Due to the clear and focused representation, the marketing department can immediately recognize the essential point and quickly deduce based on it.

15.4.5 Telling the Adequate Story Correctly

The right way of telling the data story is also mentioned by 50% of the sources. It is important to understand the supporting data to make the right decision. What matters here is the ability to transform the numbers into good stories. According to Pyczak (2017), narrative structures and dramaturgy can help with this [498]. According to Stone (2015), it is relevant to focus on a story’s beginning and end for a proper narrative [614]. The story’s start motivates further attention, while the end conveys the central messages to be remembered and carried forward [614]. Data should be illustrated with human examples whenever possible. This makes for a more tangible story and makes it easier to interpret correctly. It can also increase the usability of the messages [679, 544]. A good way to start the story of the example could be to derive the audience analysis via an amusing picture of app usage while

streaming, as many can certainly identify with this.

15.4.6 Select the right data

Four sources guide the proper selection of data. The primary issue here is to understand that there is not necessarily a truth in data, and therefore, context takes on an important role. The data, as well as its visual presentation, should accordingly be carefully considered to present a strong message [535]. Bladt and Filbin (2013) also argue that only those data that impact key business metrics should be included, thus omitting data that do not answer the problem [69]. However, it is always necessary to evaluate whether the correct data and visualization selection do not convey a false understanding [535]. In Fig. 26, this idea has been taken into account by centering on the relevant numbers regarding identifying a user group.

15.4.7 Develop Story Together

Two sources state that data storytelling is an interdisciplinary task. A data storytelling team can consist of data analysts, IT staff, marketing experts, and editors. Data analysts extract the patterns and information in the data, visualization experts create a simple and understandable visualization of the data, marketing experts provide an understanding of the needs and thoughts of the target audience, and domain understanding enables them to answer the right questions [429]. However, Pyczak (2017) points out that one person should take the lead in creating the story. The team should contribute ideas and make the story understandable and plausible [498]. For example, the analysis in the example should be discussed from different perspectives through interdisciplinary coordination before adapting it to the target group.

15.4.8 Motivate to Act

Finally, data stories should motivate the audience to take action. Therefore, a relevant goal of data storytelling is to arouse curiosity, as this is the ”engine of innovation” [316]. Overall, the example of Fig. 26 shows that Data

Storytelling, by combining data and an engaging visualization, can provide enlightenment about the app's peak usage time for more efficient marketing efforts. If the illustration is combined with a story, it can create motivation by, for example, creating general awareness about a new fact. For example, a story could be built about streaming service customers since they correspond to the essential target group of the app. The data serves as the basis for explaining the story and can reinforce statements or illuminate other aspects, such as the negative trend among users between 6 a.m. and 12 noon.

Concerning the genres, there are a variety of presentation options for our example. For example, the screen presentation with Fig. 26; depending on the speaker's voice, the annotated diagram could also be used. If the information is to be disseminated internally in the marketing department afterward, a comic strip would be conceivable to evoke attention and understanding among the readers without participating in the presentation. In a report, the magazine style could be used as an element for external communication.

15.4.9 Summarizing View

It is evident that researchers and practitioners consider easy-to-understand visual representations the best way to create a story from pure data. Interestingly, three technical or narrative criteria were identified as essential factors, whereas factors related to the story were only mentioned by every second author.

If the various success factors are brought into a temporal-logical sequence, the result is the data storytelling process we recommend. This can be seen as a guideline for the optimal implementation of data storytelling (see Fig. 27). From a process perspective in the context of data science processes, such as the CRISP-DM process, data storytelling can be understood as an overlying, accompanying process for efficient communication, which accompanies the phases of data exploration. Against the backdrop of the planned story and the story's purpose, data is selected, analyzed, and insights are gathered. The story is formulated over several iterations of the process model. It is shaped by creating illustrations matching the data in the exploration phase [359].

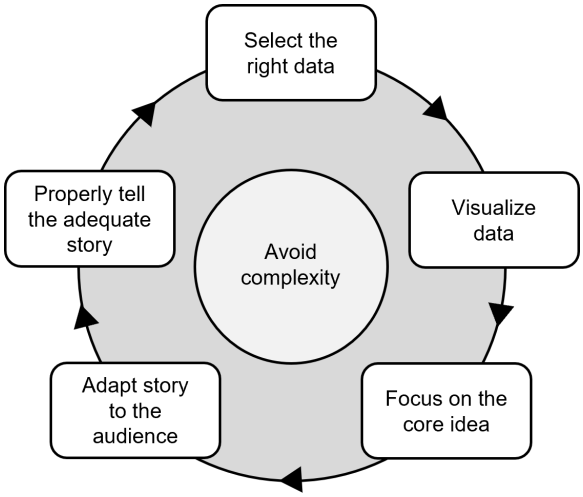


Figure 27: Data Storytelling Process

Ultimately, data storytelling combines a story, visualization, and data analysis processes in a way appropriate for the target audience, combined with the proper choice of presentation medium [456]. While data storytelling accompanies the entire process, the actual communication does not represent the end of the process but motivates new iterations of the exploration if necessary. For example, a formerly compelling data story that motivated action may be the starting point of evaluation and information in future stories.

15.5 Conclusion and Outlook

The starting point of the present work was the problem of communicating data using storytelling. Not only the exponential growth of data in today's age, but the combination of language and mathematics, especially, pose challenges for data storytelling. To address these difficulties and both transform data into relevant information and motivate audiences to take action, success factors for successful data storytelling were identified. This was done using a systematic literature review.

The evaluation of more than 20 sources results in eight success factors, six of which are considered particularly relevant due to the frequency of their occurrence. These relate both to the area of storytelling and data analysis and

visualization. The factors were put into a temporal-logical sequence, and a procedural recommendation was made based on the data storytelling process designed in this way. Data analysts can use this recommendation to tell successful data stories and convey the correct information in the right way to decision-makers.

However, data storytelling is still a niche topic [394]. However, the topic is expected to gain in importance due to increasing data volumes and the high demand for tailored information [443].

16 Comparative Summary of Findings

The comparative analysis of section 14 and 15 looks at different frameworks to support the negotiation process between businesses and consumers in the context of DDSs, which aim to improve communication between business requirements and consumer needs.

16.1 Data Science Canvas as a Structured Framework for Business-Data Alignment

Section 14 provides a structured framework for decision-makers and the operational level to support the co-evolution process of aligning strategic and operational levels (see section 2.2, figure 6). The Data Science Canvas enables the planning, communication, and control of the development of DDBMs and DDSs by harmonizing technical aspects with management perspectives. Furthermore, the precise preparation and creation of a common understanding in the context of the development of DDBMs and DDSs also align the technical with the social part of the organization, thereby promoting the co-evolution process of external (strategic) and internal (operational) levels called for in the literature [24]. As a tool for visualizing and organizing complex elements to be considered, such as data sources, methods, model quality, and stakeholders, it enables businesses to use data effectively to align their business strategy with the data strategy and achieve their strategic goals.

A vital contribution of the framework is the creation of transparency and collaboration between the various stakeholders and, through their definition, encourages the integration of consumer needs in the early stages of development by formulating business models and value propositions aimed at the quality, collection, and processing of data. By considering the relevant elements such as data collection, data integration, and model selection, the practices of companies identified in part II in the context of Business-Data Alignment, in particular, are supported by a holistic view and the uncovering of possible problem areas that hinder value-added data-driven development. It directly confronts the problems and offers solution options through further, consecu-

tive elements.

Structured planning of data-driven projects: The Data Science Canvas makes it possible to go beyond a purely technology-centered approach and align the data strategy with business objectives. Complex data analyses and models can be planned systematically and comprehensibly with the help of the tool, providing all relevant internal and external stakeholders with a common basis.

Promoting interdisciplinarity and transparency: A pivotal contribution to defining a successful data strategy is the promotion of transparency and interdisciplinary collaboration. As a communication tool, the Data Science Canvas supports the presentation of complex data projects to make technological decisions understandable to non-data-savvy stakeholders and to democratize the coordination processes based on clear and transparent elements.

Consideration of business models and stakeholders: The Data Science Canvas emphasizes the importance of the business case and promotes early coordination with the needs of internal and external stakeholders. Although the focus in section 14 is primarily on internal Business-Data Alignment, this represents a first step towards Business-Consumer Alignment by looking at data not in isolation but in the context of value creation for consumers.

Optimizing the management of data-driven business models: The canvas supports implementing DDBMs and services by providing a visual and structured basis for better managing and understanding data-driven projects. It serves as a checklist for project success and a tool for creating trust by clearly tracking progress and any arising strategic and operational challenges.

The Data Science Canvas provides a holistic framework for planning and developing DDBMs and services. It serves as a basis for integrating these into consumer-orientated coordination processes at a later stage. In Business-Consumer Alignment, the canvas acts as an internal business building block that ensures that data-driven innovations are in line with business objectives in preparation for communicating these externally and developing them together with consumers in the context of UCD.

16.2 Data Storytelling as a Bridging Method for Business-Consumer Alignment

Section 15 emphasizes the relevance of translating complex DDBMs and DDSs into understandable, comprehensible narratives to non-data-savvy stakeholders, especially consumers. Data storytelling is crucial in closing the communication gap between data scientists, businesses, and consumers. It provides data according to the developed process to make it understandable and transparent for consumers within the decision-making processes of Business-Consumer Alignment. Narrative techniques create understanding and transparency by integrating data into stories that encourage consumer participation and perspective-taking, thereby supporting their individual decisions to disclose data. While the data science canvas focuses on the use of data and its systematic preparation and documentation, data storytelling is a bridging function to establish a common language in the negotiation processes of Business-Consumer Alignment, strengthening consumers' roles through increased understanding. The identified success factors and the data storytelling process derived from them support businesses in the target group-orientated preparation of data problems and their communication with consumers.

Bridging the communication gap between businesses & consumers: Data storytelling closes the gap between transparency and understanding of businesses and consumers by transforming data into understandable, comprehensible, and appealing stories. Companies that collect and process complex data often face the challenge of communicating these insights so that consumers can recognize and understand their benefits and make informed decisions. As a communication tool, data storytelling, therefore, supports targeted negotiation in the context of Business-Consumer Alignment.

Creating motivation and perspective-taking through narrative elements: Showing the context of data through a story can motivate action and enable consumers to adopt a perspective (see part III). Stories help to present complex data, allowing the consumers to understand the importance of the data in the context of their personal decisions and increasing their trust in DDSs through transparency.

Success factors of data storytelling: Based on the systematic literature analysis in section 15, eight critical success factors were identified that promote compelling data storytelling. These mainly include reducing complexity, focusing on key messages, adapting the story to the target group, and visualizations to support the comprehensibility of the data. These factors ensure an understanding of data and the ability to translate data into action. In the context of Business-Consumer Alignment, it is relevant that consumers can recognize the individual benefits from their use of data and the corresponding added value of DDSs, which is why the traceability of the underlying data processes is essential. By democratizing the development of DDSs for consumers, data storytelling also leads to a more substantial alignment of corporate strategies with their needs.

Processual integration of data storytelling: Data storytelling should be understood as an end-to-end process that accompanies all data processing phases—from data exploration through data analysis to the final communication of the results. Data storytelling accompanies the internal exchange between specialist departments and other stakeholders and external communication with consumers. This holistic integration makes it possible to trace the entire value creation process of DDSs and thus support Business-Consumer Alignment.

Part IV offers a comprehensive approach to achieving Business-Consumer Alignment. The Data Science Canvas forms the basis for this by supporting Business-Data Alignment, the internal dynamics of decision-making, and the commitment of non-data-savvy specialists and managers. This allows the necessary development of data literacy to be additionally focused. In addition to the structuring function of the data science canvas, data storytelling complements this framework by enabling the effective communication of data with consumers.

Part V

Research Outcome

17 Discussion

Based on the comparative summaries of findings in sections 7, 12, and 16, this section discusses the results in the context of the research questions and the related work.

RQ1: What problems do businesses face within their practices in Business-Data Alignment and data-driven value creation, and what are recommendations for action to overcome these barriers?

RQ2: How do consumers perceive DDSs within Consumer-Data Alignment, and what supportive factors should be considered in a user-centered design of DDSs?

RQ3: How can businesses and consumers be supported in Business-Consumer Alignment by suitable tools and processes to create value for both?

To answer RQ1, this thesis examined the business perspective in part II. Technical and organizational challenges in DDS development were uncovered, and the resulting recommendations for action in the areas of data sourcing (see section 5) and DDS development (see section 6) were developed. For successful Business-Data Alignment, the case studies clarify that data model standardization and interoperability are crucial to integrating heterogeneous data sources. IPaaS can contribute to harmonizing data processes through flexible platform solutions and support companies' integration practices. However, according to the experts surveyed, several challenges must be considered when deciding on an IPaaS solution. In particular, the lack of data model standards, cost-benefit ratios, and data protection concerns come to the fore. The example of the development of a DDS in the form of a recommender system for food retail also makes it clear that data portability and interconnectivity are essential for developing DDSs. However, it becomes also apparent that both technological and organizational adaptations are required to make data-driven value creation effective. This manifests in particular in the need to involve users in the design processes to adequately consider their needs (see section 7). The findings are discussed in section 17.1 in the context of Business-Data Alignment through a practical lens and supplemented by the consumer studies of part III.

In part III, three case studies were conducted to explore how consumers perceive DDSs, integrate them into their consumption practices, and coordinate this with their data strategies as part of the discussion of RQ2, providing insights into the negotiation processes between added value and privacy. The results (see section 12) show that reciprocity helps to ensure that added value is perceived and trust in the DDSs can be created. Reciprocity takes place on both a sociological and an economic level. On a sociological level, reciprocity enables businesses and consumers to take on each other's perspectives, strengthening trust and reducing information asymmetries. One example is the exchange of vehicle data in peer-to-peer carsharing systems, which, through transparency about the mutual benefits of sharing driving behavior data, such as legal clarity in disputes, increased consumers' willingness to share data, enabling the development of algorithm-based reputation systems. On an economic level, reciprocity reduces the risk of opportunistic behavior, as clear rules and reciprocity structure the exchange of information.

However, this requires an understanding of the data and its intended use on the one hand and a deeper understanding of algorithms on the other, which consumers primarily abstract via tangible or familiar, apparent structures. The example of road traffic regulations clarifies that transparent cause-and-effect relationships ensure better comprehensibility and provide comprehensible consequences for certain behaviors (see section 11). Similar requirements are also evident in the development of recommender systems, where an "if-then" mechanism can make decisions and e-recommendations comprehensible (see section 9). This mainly addresses interaction with the algorithms in the form of customizable functionalities, which should be focused on as part of a participatory design approach. Adaptability to individual needs and personalized functions means that the consumer strategy - shaped by consumption practices - can be supported in the best possible way as part of Consumer-Data Alignment, and the data strategy can be informed accordingly using perspective adoption for the provision of the required data.

Perspective adoption describes the ability of consumers and companies to put themselves in the other side's position. This allows the competing strategies to be better aligned. Consumers are trying to understand how DDSs work, what

data is required, and how it is used, while companies should put themselves in the consumer's perspective to consider their needs and data strategies. The example of the negotiation processes of section 10 shows that consumers are more willing to disclose data if they understand the added value and the use of the data is made transparent. At the same time, consumers want personal needs and values to be taken into account when designing algorithms (see section 9 and 11). Perspective adoption thus ensures that consumers are not just passive data providers but that their active participation in the design of DDSs is made possible. In doing so, it represents a bridging function between companies and consumers to promote communication to enable shared added value. Section 17.2 discusses this concerning the development of tools to support the design of user-centered DDSs.

Building on this, two tools were developed in part IV to support companies in the user-centered design and development of DDSs based on the principles derived in part II and III, making theory actionable for practice. The Data Science Canvas can be primarily used for Business-Data Alignment as a structured method for planning and implementing data-driven projects. It helps businesses define and connect the expected added value, relevant data sources, required expertise, and suitable models to create a shared understanding among all stakeholders. In addition, the data storytelling process enables translating complex data into understandable, target group-specific stories that generate trust and support informed consumer decisions.

The results from section 10 and 11 in particular show that consumers have a strong need to be aware of the purpose of the data businesses collect and that this can promote a mutual adoption of perspectives within the framework of reciprocity. The Business-Consumer Alignment Canvas (BCA Canvas) is being developed in section 17.3 to address this need. The BCA Canvas is based on the findings of this thesis and directly addresses the area of tension between companies and consumers. The Data Science Canvas forms the basis of the BCA Canvas to provide the identified data needs of companies with a clearly defined purpose and to communicate them to consumers in a target group-specific manner using data storytelling. This enables consumers to make informed decisions about the disclosure of their data and actively

participate in a participatory design process.

17.1 Business-Data Alignment through a Practice Lens: From Integration Practices to Algorithm Development

Data integration is a critical issue in the Business-IT alignment process. Since this issue has become even more relevant and challenging in the context of DDSs and Business-Data Alignment [563], this thesis focuses in section 5 on data integration platforms that can support companies in their integration practices and thus ensure system flexibility.

The standardization of data models plays a vital role here in reducing the complexity of point-to-point integrations [415], facilitating data integrations from different sources, and providing Business-Data Alignment through a sufficient data basis for the implementation of data and business strategy. Nevertheless, this standardization is still missing. Furthermore, businesses need to pay attention to having access to many connectors to ensure the rapid integration of specific systems and their data transfer into their information system and to meet customer needs at the business model level [568]. Data protection, security, and transparency issues are also to be considered, resulting from the fact that the providers predominantly do not originate from Europe and do not or only partially meet local requirements. Therefore, this thesis motivates an open-source-based Integration platform that could create flexibility within the framework of aligning business and data strategy through data integration. An open-source-based IPaaS could counter the problems of standardizing data models through a common goal of contributors [36].

The benefits of data integration platforms also extend to the process integration of companies and their automation, as an integrated data basis from various systems is a fundamental prerequisite for the introduction of intelligent process automation [440]. This can be attributed to the upcoming services that rely on machine learning methods to empower the bots to make decisions, among other contributing factors. These require integrating (un)structured data to be trained effectively and benefit from synergy effects [171].

The development of DDSs as part of a DDBM thus reaches its limits in the case of a non-integrated IT system due to the possible existence of redundancies, a predominantly manual need for data input, poor or outdated data and unsustainable system interruptions [605]. In addition, for a successful Business-Data Alignment, it is necessary to define the connection between data and algorithms so that they can be regarded as a unit [332]. With a broad data basis, algorithms can be efficient, and obtaining information from raw data is only possible with algorithms.

However, the effectiveness and efficiency of algorithms depend not only on the pure availability of data but also on how this data is structured, integrated and represented. In the case of incomplete or unbalanced data integration, the algorithms may not sufficiently consider specific user groups or behavioral patterns, which limits performance and fairness [372].

Section 6 reveals that algorithms cannot reach their full potential in the case of incompletely integrated data and - using the example of food retail - tend not to take into account relevant but underrepresented minorities, such as critical consumers, and are therefore unable to provide adequate recommendations. Even if it is not always necessary for the efficiency of the algorithms to address all possible user groups with the same priority, the consideration of underrepresented groups can nevertheless improve the systems' adaptability, fairness, and relevance. Accounting for these groups can prevent systematic biases and improve the relevance of the recommendations for a broader spectrum of users, especially in areas like food retail, where consumption practices are diverse [644].

In addition, an integrated system like a household information system can support consumers' decision-making along the food practices [351, 605]. The consideration of these factors is particularly relevant from the point of view of interconnected, integrated and dispersed [212], entangled [447, 349, 577], or chronologically ordered [445] food practices to be able to provide suitable data-driven value-added services in this sense [351].

Here, too, the required diversity of connectors becomes evident, as many consumers interact with businesses across multiple digital touchpoints like mobile apps for grocery shopping, digital recipe platforms, or self-checkout sys-

tems in stationary retail. These touchpoints generate different data sets - from shopping lists to dietary habits and purchase histories. Without standardized integration, this data remains fragmented and cannot be used efficiently (by algorithms). At the same time, there is the problem that some consumption practices can only be covered incompletely by DDSs. An example is weekly grocery shopping, which does not necessarily have to take place only in the stores of a particular business. Still, it can also be continued in a competing discounter, for example. The transaction data distributed across different providers does not flow into the same system, which means that relevant information is missing and only a fragmentary consumption profile exists. As a result, recommender systems can only model actual shopping behavior to a limited extent, impairing personalized recommendations' quality and accuracy (see section 6).

These different touchpoints must be brought together for sufficient data integration [215], as integrated household information systems "should respect the individual artifacts choices, but still enable a common data and service platform for integration [...] since integration is perceived as a key feature that reduces effort and increases usability" [351]. Integrated household information systems can enable consumers to provide more consistent and appropriate recommendations by combining different data sources and interactions in a unified platform, allowing for a more complete picture of consumer habits.

In section 9, the unity of data and algorithms is also reflected in trusting data-driven value-added services. Reliable and open data and algorithms must properly consider restrictions or ethical choices, particularly in the sensitive consumption domain of food [644, 350]. But not only in food, other areas like health or finance also involve critical decisions for consumers, and a lack of transparency can result in a lack of trust, as consequences such as breaches of privacy, health risks, financial loss, or a breach of ethics cannot be assessed [458].

The studies in section 10 and 11 further emphasize the necessity of connecting and integrating various sensors and telematic solutions to enable DDSs. While the sharing economy is examined in this thesis, relying on decentralized and trust-dependent structures, DDSs like algorithm-based reputation systems can

solve current problems such as social bias in reputation mechanisms and offer added value as an objective measure for trustworthiness [602]. Nevertheless, the seamless integration of sensor and telematics data to provide reliable and transparent algorithmic solutions is also relevant for platform domains like smart home, digital health, and insurance [113, 677, 584]. Therefore, these principles can also be applied to other domains, especially where DDSs interact with end-users in sensitive or critical domains.

17.2 Consumer-Data Alignment: From Flexible Consumer Practices to Negotiating Between Privacy and Value

17.2.1 From Consumer to Co-creator: Flexible Human-Algorithm Interaction to Support Individual Consumption Practices and Foster Trust Through Understanding

To support consumer practices and, thus, consumer strategies through DDSs, this thesis presents design implications that provide added value for consumers. Therefore, section 6 shows that Business-Data Alignment relies on access to sufficient and integrated data to represent consumption patterns adequately. In practice, businesses (esp. low data business models [270]) often only have fragmented data, depending on their touchpoints with consumers [312]. Furthermore, the DDSs offered frequently do not sufficiently reflect diverse and evolving consumption practices [644, 352, 351], limiting the ability to provide relevant and personalized recommendations for specific user groups that do not match, e.g., an 'average consumer' or cross-retailer consumption patterns. As existing literature highlights, consumption strategies are not static but shaped by dynamic, situational, and context-dependent factors [672, 241]. Furthermore, they are deeply bound to habitual routines, emotional triggers, and external influences, such as pricing strategies or product availability [577, 212, 182]. Therefore, as claimed by UCD (see section 2.4.1), it is essential to consider individual needs and actively involve consumers in developing DDSs. These aspects motivate the analysis and design of user-centered DDSs through UCD to understand consumer strategies' diversity and perceptions [577, 212, 352].

Building on this, the sections 6 and 9 show that when developing DDSs such as personalized recommender systems, businesses do not adequately take into account the diverse consumption patterns of previously underrepresented consumer groups [644, 303]. While the metrics of recommender systems often focus on the average consumer, they fail to consider the broader and more complex dynamics of consumption habits [644, 5]. Consumers and their strategies need to be considered more holistically to take into account not only narrower factors such as health [436] or sustainability [303] but also consumption habits in all their complexity and diversity. This limitation represents a gap in academic discourse and practical implementations of personalized DDSs [643, 303] and requires a holistic understanding of consumer strategies, incorporating not only isolated factors but also the interplay of personal goals, values, and situational contexts [644, 349, 350, 303, 643].

Therefore, this thesis investigated the user-centered development of a food recommender system in section 9, enabling ethical [303] and practice-oriented [349] profiling. Furthermore, evaluating the system uncovers factors that should be considered beyond the average consumer in developing DDSs to support consumption practices more flexibly and enable consumers to make informed and value-driven choices. The creation of profiles allows for the integration of personal goals, altruistic values, allergic restrictions, and personal situations. This profile creation should be as flexible as possible since, in addition to the recourse to institutionalized and shared profiles and rules (such as the general direction ”vegan”), it must also be possible to modify rules and sub-rules to support own definitions of profiles (such as ”vegan”) in the context of a profound personalization. This further underlines the relevance of developing DDSs with end users [373].

The necessity of flexible profiling in the context of DDSs is also shown in section 11. However, the reason for this lies less in obtaining the ’right’ information and more in adapting algorithms towards normative values. For example, although the study participants interpret algorithmic ratings in the area of peer-to-peer carsharing as regulatory based on the rules of the road, they also want to flexibly incorporate subjective norms into the evaluation to help shape the algorithm itself and integrate corresponding values that are not

attributed to the algorithm. This aligns with prior research, which views algorithms as elements of social order processes [112], which differ according to regulative, normative, cognitive, and technological aspects. However, the cognitive and technical levels are also addressed, calling for weighting options and explanations of the factors influencing algorithm-based ratings and supplementing the weighting with a probability of occurrence and criticality within a social context. The evaluation interviews in section 11 also clarify that profiling through personas is desirable, understood as ”personification” of the algorithm for comparability and an indication of possible adaptations. The literature confirms this on the generation and incorporation of explanation sentences in the context of recommender systems [706, 119]. This also enables the filtering of a pre-selection of potential tenants in the area of P2P carsharing, thereby reducing information overload.

Similarly, trust in DDSs also shapes consumer strategy. The findings from sections 10 and 11 show that trust can be created particularly by understanding data and algorithms. Clear visualizations and comprehensible algorithms are used to understand the DDS. Nevertheless, this is often a fallacy in transparency, as a high level of complexity results in a low level of benefit for end-users, for example, through a precise breakdown of how DDSs work. Detailed algorithmic explanations can overwhelm users, reducing the utility of transparency. Therefore, there is a need to balance technical detail and practical interpretability [590].

In the context of HCI, the interaction between algorithms and users should be designed so that a deep technical understanding of the algorithm is no longer necessary to derive value. Therefore, provided explanations should enable interpretation of the results of DDSs rather than an explanation of the algorithm itself, focusing on making outcomes actionable and meaningful rather than simply explaining the algorithm’s underlying mechanisms [251, 228]. While understanding refers to a general comprehension of what the system is intended to achieve (what for?), explanation considers the reasoning behind specific outputs by uncovering regularities or causal relationships (why/how?). Interpretation focuses on how the outputs of algorithms can be applied in context-specific scenarios (what does this mean?) [435, 307].

These results partially negate the findings from section 10, which, in line with the literature on privacy, calls for an understanding of how algorithms work [671, 694]. These aspects are also confirmed by the more recent literature on explainable AI and the design of algorithmic interactive systems [546] as well as studies on interpretability by design [391, 473, 25, 710, 375]. To this end, the trade-off between application-based, human-based, and function-based interpretability should be considered in socio-technical systems such as algorithms. In the context of the application-based evaluation of interpretability, the focus is on the effect of interpretability in the course of a specific, precisely defined task or application. The human-based evaluation analyzes the quality and general comprehensibility of the terms of interpretations for non-expert users. Finally, function-based evaluation deals with a formal, mathematical definition of interpretability to evaluate the quality of interpretation [175]. For example, in section 11, a human-based assessment was examined through user feedback on their understanding of the algorithmic ratings and comparing them to familiar reputation mechanisms like user ratings.

To prevent algorithmic bias, it should also be ensured that a DDS can be compared with reality to ensure a fair and comprehensible result in specific situations, which Rouvroy (2013) [527] also confirms. If DDSs can create trust and be flexibly adapted to consumers' individual requirements, there is potential for increased added value in the context of consumption practices and consumer strategies.

In summary, this thesis shows that trust, interpretability, and adaptability are not isolated design factors but are interconnected and need to be approached holistically to maximize the added value of DDSs. Therefore, this thesis presents concrete examples from the food and mobility sectors that show how these principles can be operationalized to support meaningful and context-aware DDSs.

17.2.2 Consumer Data Strategies in the Field of Tension Between Privacy and Added Value: Reciprocity as the Key to User-centered Negotiation Processes

Section 10 adds a practical component to understanding privacy strategies and provides broader insights into consumers' perceptions of disclosing their data in specific contexts. The results support the previous literature in the aspect that consumers tend to prefer a data-sparing data strategy concerning the use of DDSs and that the data collected must be in proportion to the added value obtained but also - as an extension of the previous literature - to the perceived interests of the data processor or user. This phenomenon of consumers adopting providers' perspective stems from the principle of reciprocity on a social level, which - as a universal social principle - refers to a mutual exchange in social actions where both parties adjust their actions based on perceived fairness and benefits [413, 420]. However, the most important thing here is that this serves the purpose of seeing and understanding how the other party perceives a consumer through data disclosure. Accordingly, it should be possible for DDSs to give consumers access to the provider's perspective and, thus, into how their data shapes their perception of them. In this thesis, this was, for example, considered by the participants in section 11, where they wanted to know how they could improve their algorithm-based ratings when taking themselves into the perspective of a car rentee.

Based on the findings of this thesis, this can lead, in particular, to improved coordination in the event of unpredictable and uncertain events like intense braking maneuvers (see section 11). However, reciprocity at the economic level could also be observed, which in the context of information asymmetries should minimize the risk of opportunistic and unfair behavior [19]. This aligns with research on trade-offs and supports the idea that consumers find it more acceptable to disclose their data for expected benefits than if they cannot expect a reward [671, 694]. However, the results focused on the fact that both parties involved can access the data equally, and only if both agree to this can one-sided injuries be prevented. Data disclosure by consumers is seen as unobjectionable, especially if they expect the data to be available anyway and support the provider with digital information. The principle of

reciprocity, therefore, reflects balanced transparency to mitigate power imbalances (through information asymmetries) and foster long-term trust in DDSs [427].

Considering the various factors and a situation- and context-dependent discourse between data disclosure and expected added value, this thesis's findings demonstrate that an area of negotiability exists between consumers and providers of DDSs, which is not a static space but a dynamic process shaped by situational and contextual factors during different phases of service usage (see section 10).

The situatedness of information, which, according to the findings of this thesis, can be differentiated into a trust-supporting function before, during, and after the use of the DDS, thus complements the more traditional view of creating trust primarily before use [188, 278]. Furthermore, section 10 reveals that many consumer data disclosures depend on specific conditions, such as transparent data usage statements, assurances of privacy protection, and explanations of the added value resulting from the disclosure. These conditions form the basis of consumer trust and willingness to share information. Therefore, a flexible system design is advisable to ensure that consumers can actively and autonomously decide what data to disclose and under what conditions. Self-determination in data disclosure refers to the ability of consumers to retain control over their data disclosure decisions while being aware of the purposes for which their data will be used. This clarity fosters trust and increases the acceptance of DDSs [65, 59]. Here, the definition of individual conditions as negotiation support is desired to adequately ensure consumers' autonomy in the context of their data disclosure.

Section 11 also certifies the results and uses a design artifact to show how a DDS can be developed in a user-centric way by disclosing data. This requires a transparent and comprehensible added value for consumers and service providers; see also Athey et al. (2019) [35]. To this end, consumers must be involved in the technological processes to make them transparent, comprehensible, and user-centered and to increase interaction with the algorithms [251]. Section 9 supports this with the demand for transparent algorithms, which is not met even by supposedly consumer-friendly providers.

From the consumers' perspective, how a DDS decides often remains unclear. This lack of transparency can create mistrust and reduce user confidence in the system's outputs, significantly when decisions strongly impact consumer practices [458, 341].

Further studies on DDSs show that consumers must be aware of the data collected, especially since they do not experience a surprising and frighteningly accurate personalized service [353]. On the other hand, a paradox exists which goes hand in hand with a general quality of service and where personalization is unnecessary. In line with prior research [22], it is required that future designs provide a transparent overview of collected data and the rights guaranteed by the GDPR or make them available to consumers so that personalized (third-party) services can be enabled [601, 605]. In the context of peer-to-peer carsharing, investigating obstacles and barriers to sharing one's vehicle shows that consumers are even open to disclosing data about their driving behavior if this increases their chances of renting a car or if the rental price is lower. In this case, the abovementioned perspective also ensured that monitoring by vehicle providers is considered appropriate, as it is understandable that they should be aware of whether the car renter is using the car sensibly [469]. A study on blockchain-based sharing platforms shows that taking the principle of reciprocity into account helps to increase user acceptance and trust in the services concerning possible data misuse. A blockchain solution could contribute to improved security for privacy protection by mapping the negotiated data protection agreements via smart contracts and only disclosing data when events defined by the consumer occur [82]. Perspective-taking also plays a central role in the creation of digital sovereignty for consumers in the context of control over their data [356], which "provide the best balance for each [moral] calculus" [141]. This is also supported by identifying privacy problems in establishing a digital sovereignty concept in various research domains, which demand individual privacy and autonomy to prevent data misuse and surveillance that occurs, for example, through AI [355].

Overall, Consumer-Data Alignment describes how consumers align their data strategies with their consumption practices. This alignment can be divided into two main types, each reflecting different approaches to balancing data

disclosure and strategic goals (see Table 22 and 23). These types are derived from the empirical findings in part III, supported by theoretical foundations from prior research on the heterogeneity of consumer data strategies [471, 671, 694].

- **Data-Enhanced Alignment Strategy:** Consumers adapt their data strategy to their consumer strategy so that they only consciously use DDS by disclosing their data when they want to change their existing consumption habits and see a benefit or strategic advantage in usage. These benefits may include minimizing risk (e.g., using a food scanning app to receive personalized alerts based on allergy or intolerance data) or optimizing specific consumption habits (e.g., using a food recommendation app and disclosing personal health data to receive personalized meal recommendations based on specific dietary goals). For example, in section 11, the participants describe sharing their driving data under the specific condition of better rental prices.
- **Data-Dominated Alignment Strategy:** In contrast, consumers with a data-dominated alignment strategy use data disclosure as an essential part of their consumption practices with the primary goal of maximizing benefits derived from their data, such as increased convenience, personalization, or financial incentives. In this way, they often focus on these benefits rather than data privacy. For example, one participant in the 10 section was willing to disclose detailed data about their driving behavior to receive financial benefits and preferential booking options. The motivation appears primarily based on convenience and measurable benefits despite potential privacy concerns.

These two basic alignment types show the heterogeneity and complexity of Consumer-Data Alignment with varying priorities and strategies. Furthermore, the findings of part III motivate the distinction of the alignment between data strategy and consumer strategy on three levels [371, 587]:

1. **Individual Level:** The alignment processes are shaped by expected benefits, personal risk assessments, and subjective preferences [671]. In

the study in section 10, a participant describes sharing driving behavior when there is a discount for driving correctly.

2. **Situational Level:** The alignment is influenced by situated and context-dependent events. For example, a participant in section 10 describes that he would share GPS data on the condition that an accident has occurred.

3. **Strategical Level:** Alignment is determined by broader goals, e.g., using a food recommender system to achieve a long-term goal such as improving health (see section 9).

Strategy \ **Level**		**Individual Level**	**Situational Level**	**Strategical Level**
Data-Enhanced Alignment	Description	Targeted use of data for personal benefits and risk minimization	Targeted data disclosure in specific contexts to achieve situational advantages	Targeted data disclosure to achieve long-term personal goals
	Benefits	Personalized services that are tailored to specific goals	Problem solving in key situations	Long-term optimization of consumption practices
	Risks	Risks are reduced through selective and conscious disclosure	Data disclosure only under certain conditions	Long-term goals define the scope for disclosure
	Example	A user shares allergy data to receive personalized food recommendations	A user only shares GPS data in the event of an accident to enable quick assistance	A user uses a nutrition tracking tool for long-term health optimization

Table 22: Data-Enhanced Alignment Strategy at Individual, Situational, and Strategical Level

The findings highlight consumers' different strategies and alignment processes, revealing individual motivations, situational contexts, and broader strategic goals. Therefore, when considering Business-Consumer Alignment, these findings must be integrated into tools to support these complex, often competing alignments between companies and consumers.

Strategy / Level		Individual Level	Situational Level	Strategical Level
Data-Dominated Alignment	Description	Maximum disclosure of data to increase benefits	Full data disclosure to avoid situational challenges	Data disclosure for extensive, long-term optimization
	Benefits	Extensive personalization for maximum convenience	Fast problem solving thanks to detailed data	In-depth insights for optimizing long-term strategies
	Risks	Consciously accepted risks in exchange for maximum benefits	Justification for full disclosure by possible situational contexts	Full transparency to data owners due to long-term data
	Example	A user shares detailed driving data to receive financial benefits	A user accepts full GPS monitoring for optimal support in case of an accident	A user has his diet and transaction data comprehensively analyzed to receive optimal recommendations for weight reduction

Table 23: Data-Dominated Alignment Strategy at Individual, Situational, and Strategical Level

17.3 Business-Consumer Alignment: Co-Creating Data-Driven Value-Added Services

Successful Business-Consumer Alignment needs companies to be aware of their specific opportunities in developing and designing DDSs and overcome the existing problems and hurdles in Business-Data Alignment (see section 17.1). To this end, section IV motivates using the developed Data Science Canvas tool, which supports companies from a holistic modeling approach using a canvas. Section 14 shows that the Data Science Canvas can be used as a tool for structuring DDS development, as well as as a checklist and instrument of trust for internal and external stakeholders. The Data Science Canvas also enables non-data-literate specialists and managers to understand data science development projects through its clear and comprehensible presentation and to communicate within a multidisciplinary development team

using a generally understandable language based on the Data Science Canvas. Using elements for data availability, collection, preparation, and integration can counter the problems identified in section 17.1 or uncover and solve them in joint internal coordination if they lack awareness.

Furthermore, in section 15, a process was developed using derived success factors, which should improve the essential communication within development teams as well as with external stakeholders, such as consumers, through target group-specific and understandable language and thus promote overcoming (mutual) problems.

Derived from the basic principles of UCD, Business-Consumer Alignment requires an understanding of and for both sides. Businesses must understand consumers and their preferences and needs, and consumers must understand what businesses want to achieve with their value-added services and how these are designed. The principle of reciprocity, which has already been derived in the section 17.2, becomes apparent here. On a social level, the phenomenon of perspective-taking revealed in this thesis (see section 10) supports the understanding of both sides. For example, businesses must engage with consumers and incorporate their needs, wishes, requirements, and preferences into the development process of DDSs as part of UCD to create added value for both sides. At the same time, however, consumers must also adopt the perspective of businesses as part of their Consumer-Data Alignment strategy, try to understand the design of DDSs and data requirements, and evaluate them according to their individual data and consumer strategy.

To this end, it is helpful to define a purpose for the necessary data to be collected, which is understandable for the consumer and is differentiated before, during, and after use. This differentiation ensures that consumers understand what data is being collected and why at every stage of using a DDS, increasing transparency and reducing uncertainty about data usage. The importance of data transparency after use becomes apparent, for example, in section 10, when participants emphasized the importance of receiving reports on the condition of the vehicle after use, e.g., on unusual incidents or fuel status, as this simplifies the rental process and strengthens trust between the actors.

On an economic level, the Business-Consumer Alignment must also aim to

reduce the prevailing information asymmetries in the development of DDSs and thus minimize the risk of opportunistic and unfair behavior for consumers through transparency and UCD (see section III). The expected benefits or added value also play a leading role here, as consumers are more likely to act positively for the company as part of their strategy if they receive an acceptable value in return for disclosing their data (see section 10 and 17.2). Therefore, the co-creation of DDSs must enable the negotiation of situation- and context-related trade-offs to use the lessons from the principle of reciprocity accordingly. If both perspectives are combined via a coordination process between the actors, as highlighted in part III, this alignment is not just theoretical but practically achievable through transparent, participatory design approaches. In that case, this can encourage use and adoption for businesses on the one hand and strengthen or preserve consumers' autonomy on the other [356, 355]. However, if this coordination fails, there is a risk that consumer skepticism will increase, willingness to share data will decrease, and adoption of DDSs will be limited. A lack of coordination can also lead to inefficiencies in service design and missed opportunities to create reciprocally beneficial outcomes.

Since the alignment process between businesses and consumers is complex, mainly due to the problem of uncertainty and lack of clarity about a data strategy on the part of the consumer, as shown in section 17.2, this thesis recommends a tool that can be used in the context of co-creation workshops, which are used for alignment under the lens of UCD, to support precisely this alignment process for the consumer perspective. Canvas models are essentially suitable for summarizing a complex problem as clearly as possible and offering corresponding guidance for solution finding [464] (see also section 14). For this reason, the Business-Consumer Alignment Canvas (BCA Canvas) was derived based on the findings of the UCD studies from section III, which is intended to support this alignment process as part of the co-creation workshops in the development of DDSs.

The BCA canvas (see Fig. 28) encompasses the two perspectives of businesses and consumers. In preparation for the co-creation workshop, businesses can use the data science canvas to present the corresponding added

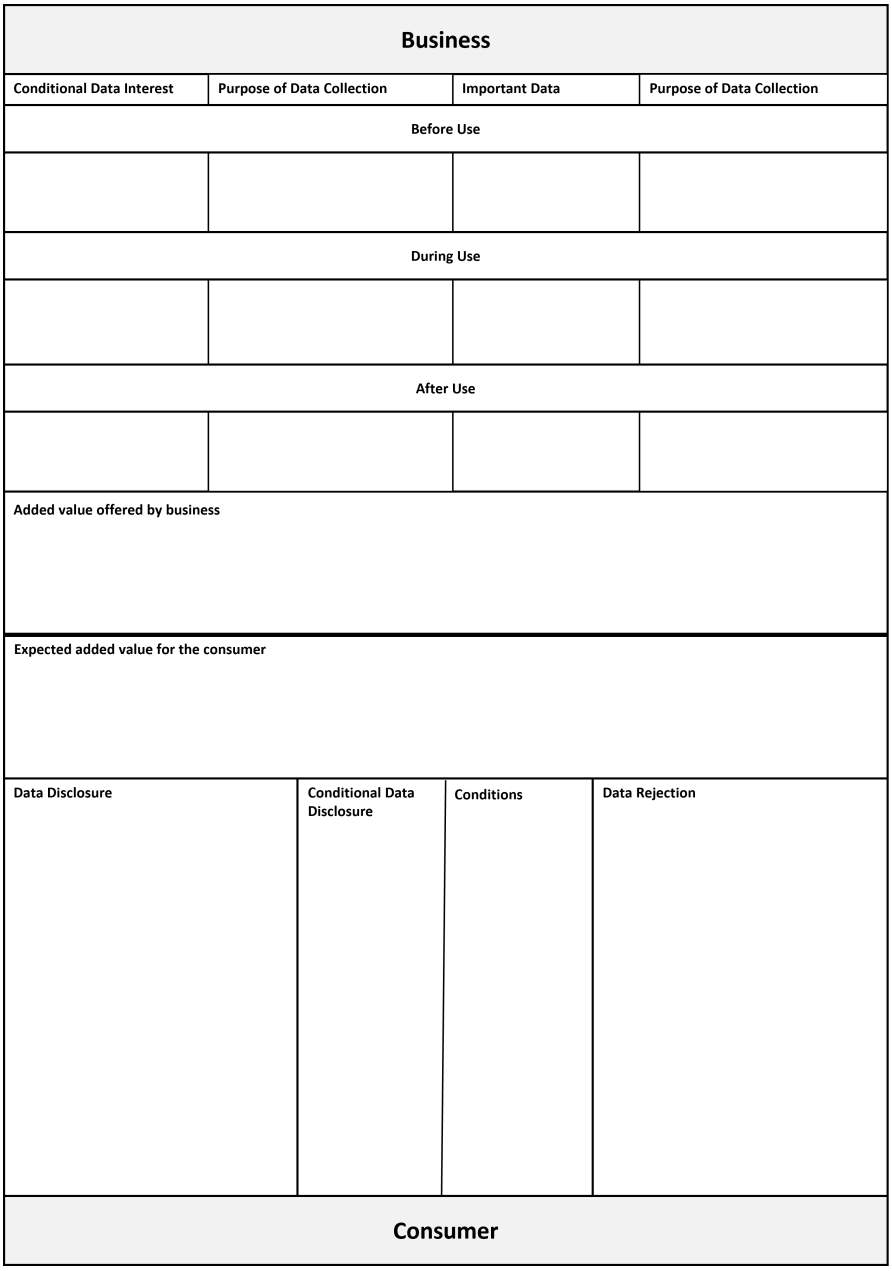


Figure 28: Business-Consumer Alignment Canvas

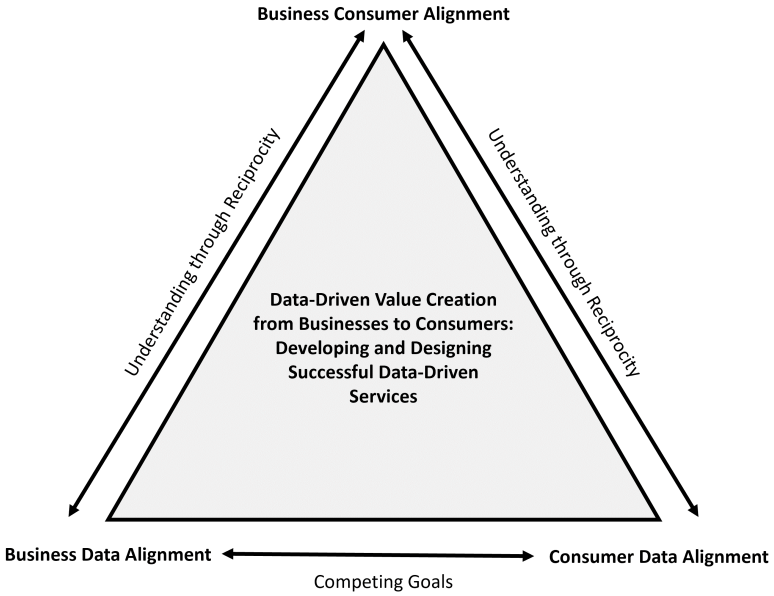


Figure 29: Triangulation of Business-Data, Consumer-Data, and Business-Consumer Alignment for Data-Driven Value Creation

value they would like to offer from developing a DDS based on their findings. At the same time, the various data can also be differentiated according to conditional data interest and essential data with their purpose of collection as an explanation and divided into before, during and after the use of the service to the consumer. Using this presentation, consumers can now interact with the company participants during the workshop to discuss which data they would disclose, conditionally disclose or reject. Particularly in the area of negotiability introduced in the sections 10 and 11, conditions for the release of conditional disclosure data can be defined, which are demanded by consumers with the expected added value in the context of their consumer strategies or practices and in coordination with the data strategy depicted there, which is abstract but supports the decision-making process.

Overall, the findings of this thesis show that promising and value-adding DDSs can only be developed and designed through the triangulation of Business-Data, Consumer-Data, and Business-Consumer Alignment (see Fig. 29). As Business-Data and Consumer-Data Alignment compete due to the different

motivations (profit generation vs. data protection), understanding both sides is required when aligning businesses and consumers. To this end, three tools were presented that can be used as part of a UCD approach, for example in co-creation workshops, to initiate precisely this triangulation: The Data Science Canvas is used to derive meaningful and value-adding DDSs on the part of businesses as well as their intensive examination of the problems and hurdles that arise from the development of DDSs (see part II and section 14). The data storytelling process also helps companies in the internal communication of these DDSs, the results of the analyses used, the demonstration of added value, and the external communication of these to consumers. The results of the Data Science Canvas can also be incorporated into the BCA Canvas, which, as a tool for alignment between businesses and consumers, promotes coordination between the two players and, through reciprocity, promotes the reduction of information asymmetries as well as the adoption of a perspective for understanding and strengthens the autonomy of the consumer.

18 Conclusion

This thesis is divided into five parts, each dedicated to investigating the tension between businesses and consumers in developing and designing DDSs and providing recommendations for action and tools to support them in this process.

Part I motivates this thesis, derives the research questions, and creates the theoretical foundations for DDBMs and DDSs for companies and consumers. Therefore, section 1 emphasizes the need to redefine traditional concepts such as Business-IT Alignment in the context of DDBMs and introduces Business-Data Alignment as the alignment of the business strategy with the data strategy of companies. Furthermore, the role of consumers in the development of DDSs is emphasized, and based on this, Consumer-Data Alignment is derived, which is of central relevance in the increasing integration of DDSs into consumers' living environments. Consumers must harmonize their consumer and data strategies to benefit from the data they provide. In this context, consumer strategy can be understood as a meta-level framework shaped by consumption practices and behaviors, focusing on specific goals (see section 2.3). This thesis further introduces Business-Consumer Alignment, conceptualized as a systematic negotiation process between businesses and consumers to harmonize competing goals, requirements, and needs (see also Fig 29) in developing DDSs.

In addition, the research domains to which this thesis contributes are presented. In particular, information systems and consumer informatics research are addressed in the context of data-driven value creation. The case studies conducted as part of this thesis in the selected application fields of consumer informatics, food, and mobility expand the understanding of recommender and reputation systems research by providing insights into the perceptions and practices of businesses and consumers.

Section 2 outlines this work's conceptual and empirical foundations, starting with transitioning from traditional business models to DDBMs, where data is at the center of value creation. This section introduces DDBMs and DDSs as well as data-driven value creation. The terms Business-Data Alignment,

Consumer-Data Alignment, and Business-Consumer Alignment developed as part of this work are presented. Business-Data Alignment is essential for companies to utilize data within their strategic and operational structures, thereby supporting innovative DDSs by aligning data with the business strategy. The focus is on the importance of flexible data infrastructures, interdisciplinary collaboration, and the development of DDSs. Consumer-Data Alignment examines the alignment processes of consumer practices with consumer data strategy. It highlights privacy, transparency, and reciprocity as critical factors influencing consumer engagement with DDS and emphasizes that consumers' willingness to use DDS depends on the perceived added value. Finally, Business-Consumer Alignment harmonizes the perspectives of businesses and consumers in developing and designing DDS to bridge the gap between business objectives and consumer expectations. This ensures that services are technically adequate and meet users' needs to create added value. Business-Consumer Alignment is based on the UCD concept, which advocates the active involvement of users in the design of services by promoting transparency and aligning the benefits of services with consumer values. It is, therefore, a co-creation model in which companies and consumers work together.

The following three parts are dedicated to the corresponding alignment terms, but this section also presents the state of the art of the respective case studies conducted.

Section 3 introduces the research framework and shows the structure of this thesis and the research questions to be answered. The research activities used in the case studies are also presented.

The case studies in **part II** deal with developing DDSs from a business perspective and highlight important aspects that should be considered for successful Business-Data Alignment. To this end, section 5 discusses challenges in integrating data from multiple sources and presents factors that businesses can consider when introducing IPaaS platforms to support data integration practices. Section 6 then highlights the development of a recommender system from a business perspective to illustrate issues like data portability and interconnectivity by uncovering problems in using real-world transaction data

to develop models for modeling repeat purchase behavior. Overall, this part shows that Business-Data Alignment must overcome technical and organizational obstacles, ensuring data can be effectively leveraged for added value for consumers and internal operational efficiency. To this end, factors and recommendations for action are derived to support companies in Business-Data Alignment.

Part III turns to the consumer perspective by examining how DDSs interact with consumer practices. Consumers of DDSs are not only passive recipients but active participants, and they provide data that influences service DDS design and personalization. Therefore, UCD serves as a framework for creating DDSs that actually meet consumer needs. The presented case studies on Consumer-Data Alignment in food and mobility application fields illustrate how trust, privacy concerns, and transparency shape consumer engagement and consumer practices. Balancing business objectives with consumer preferences and practices is highlighted by examining consumer data strategies, emphasizing participatory design methods to foster mutual benefits through reciprocity.

Traditional recommender systems are usually based on popularity or similarity metrics [14], so they are often not tailored to sustainable or conscious consumption practices. Therefore, section 9 builds on section 6 and looks at the development of recommender systems in the food sector that considers the niche group of critical consumers. This case study highlights the necessity for recommender systems to move beyond generalized metrics of popularity or similarity and incorporate user practices, values, and sustainability goals by design.

To further explore the alignment of consumer and data strategy, the focus of the 10 section shifts from personalization in food consumption to the negotiation processes underlying data disclosure for algorithm-based reputation systems in peer-to-peer carsharing. These DDSs aim to solve the problems of trust and reliability between peers in the sharing economy, such as social bias or fake reviews, through a data-driven approach. Building on this, section 11 expands the understanding of the topic by developing and evaluating a prototype algorithm-based reputation system using car telematics data. This

case study shows how telematics data can address trust deficits in peer-to-peer relationships but also highlights the importance of contextual customization to individual user preferences.

Overall, the case studies in part III show that consumers are not a homogeneous group but have differentiated demands on DDSs in the context of Consumer-Data Alignment. Businesses must consider individual consumer practices and data protection concerns to promote trust and the willingness to disclose data. It is also evident that a user-centered approach and transparent communication form the basis for successful DDSs. In general, DDSs should be tailored to the user needs, practices, and values of their specific application areas, thereby not focusing on a one-size-fits-all approach. These principles should also be considered as a basis for the design of DDS in other areas but must be tailored to the respective context and requirements.

Part IV proposes practical user-centered tools and methodologies to facilitate alignment between businesses and consumers in DDS design and development. The developed Data Science Canvas in section 14 and Data Storytelling Process in section 15 are introduced to bridge data literate and non-data literate stakeholders, promoting effective internal and external communication about DDSs. These instruments support Business-Consumer Alignment by helping businesses to incorporate consumer perspectives and data literacy into service design processes, ultimately aiming to create DDSs that add value for both consumers and businesses. Therefore, the framework of Business-Consumer Alignment is derived as a structured approach to managing data-driven projects focusing on clear communication and user-centered development.

Lastly, **part V** discusses the findings of this thesis by focusing on the necessary triangulation of Business-Data Alignment, Consumer-Data Alignment, and Business-Consumer Alignment. Furthermore, another tool to support Business-Consumer Alignment is derived to support the complex and often conflicting negotiation processes in the design or development of DDSs based on the overall findings and to support the reciprocity identified from the consumer studies in part III.

18.1 Contributions

This thesis contributes to Information Systems, Consumer Informatics, and, more specifically, Recommender Systems and Reputation Systems research.

Information Systems research is extended by exploring how data-driven value creation can be structured to align business and data strategies. Therefore, this thesis introduces Business-Data Alignment as a fundamental component that showcases how businesses can leverage data to optimize decision-making and create value with DDSs, reinforcing data strategy as an integral element in organizational decision support and value creation. Tools like the Data Science Canvas and Data Storytelling help to manage and implement data-driven business approaches effectively through systematic planning and communication, thereby better linking the strategic and operational levels. Therefore, the Data Science Canvas bridges strategic planning and the operational implementation of DDBMs and DDSs. For this purpose, it ensures a clear structure of business goals, data sources, and methods, enabling their conversion into concrete operational actions. Furthermore, by examining technical challenges such as data integration and organizational barriers in developing DDSs, solutions are identified to support businesses in the context of DDBMs and DDSs.

This thesis expands the understanding of consumer informatics research by providing insights into the perception and user-centered design of DDSs and how they influence consumption practices. To this end, this thesis introduces Consumer-Data Alignment, illustrating how consumers interact with DDSs in the respective examined application contexts.

In the application domain of food, this thesis focuses on personalized recommender systems that go beyond popularity or similarity metrics and consider critical consumers' practices and value choices. It is shown that trust and transparency are key factors in supporting sustainable and ethical consumption practices. Recommender systems contribute to reducing information overload in this area and can also actively shape the reflection and adaptation of consumer habits.

The two case studies conducted in the application areas of mobility are used to show how trust, transparency, and perceived added value influence consump-

tion practices and the willingness of consumers to share data. They provide insights into how DDSs can be adapted to individual values and preferences. They also show how technology and user-centered aspects combine to design DDSs that better support consumption practices. At the same time, the discourse on perceived added value and transparency, control, and privacy is addressed.

Overall, the case studies clarify that Consumer-Data Alignment has different requirements in different application contexts and dynamics, which must be addressed as part of UCD. In addition, the Business-Consumer Alignment Canvas was developed to structure the negotiation processes between companies and consumers and make them transparent. This supports the user-centered design of DDSs and promotes reciprocity.

A process was derived that supports the operationalization of data storytelling and thus acts as a bridge between communication within companies and between companies and consumers in Business-Consumer Alignment. This thesis, therefore, emphasizes that data storytelling can make development and design with data more transparent and accessible to consumers, helping to mitigate privacy concerns and foster trust. By translating complex data analytics into narratives that convey the service's value and purpose, businesses can engage consumers more meaningfully and reduce information asymmetries.

The triangulation of Business-Data, Consumer-Data, and Business-Consumer Alignment provides an integrative methodological perspective, going beyond traditional UCD approaches by embedding alignment mechanisms across three distinct but interconnected dimensions. In addition to UCD, which is often applied by businesses when focusing on user requirements in the design phase, this work complements the scope of UCD for businesses by systematically integrating strategic, operational, and negotiation processes into a coherent framework for DDS development. The dimensions ensure that technical feasibility (Business-Data Alignment), user needs and perceptions (Consumer-Data Alignment), and negotiation of (competing) goals are not addressed in isolation but viewed as reciprocal influencing processes throughout the development of DDSs (see Fig. 29). This triangulation is supported by the tools

developed in this thesis, which form a bridge between the stakeholders, as they not only cover technical and organizational requirements but also integrate the needs and expectations of consumers.

Furthermore, there are theoretical contributions to DDS design: Through the case studies and empirical results in this thesis, the theoretical understanding of how DDSs can be customized beyond average consumer profiles is extended, revealing the importance of personalized DDSs and demonstrating the practical implications of aligning consumer practices with flexible service models. The findings of this thesis underline the role of DDSs in driving value by recognizing and responding to diverse consumer practices by reducing information asymmetries and information overload.

Specifically, this thesis contributes to recommender systems research by investigating how these systems can be more suitable for critical consumer groups. This involves developing systems that go beyond promoting mainstream products and instead support niche or ethically motivated consumer practices, such as vegan diets, by considering user-centered values and preferences. Furthermore, it supports ongoing discourses on enhancing user trust and engagement in recommender systems by promoting transparency and ethical alignment in recommendations.

Regarding reputation systems research, this thesis extends the understanding of algorithm-based reputation systems, particularly within peer-to-peer carsharing platforms. Integrating objective data (such as car telematics) as an additional reputation mechanism reduces social biases and subjectivity issues, thereby improving trustworthiness. These findings are significant in designing more reliable and inclusive reputation systems, demonstrating how algorithm-based reputation systems can provide fairer assessments and mitigate discrimination in the sharing economy.

18.2 Limitations

This thesis is limited due to various factors, so the generalisability of the results must be discussed. The need for future research can be derived from this.

The case studies in this thesis focus on specific application areas, including integration practices, the food sector and mobility. In-depth insights into the requirements and practices of these sectors are provided, but the transferability of the results to other areas is only possible to a limited extent. For example, more complex regulatory frameworks or more sensitive data can lead to other data-driven practices. Therefore, the concepts of Business-Data, Consumer-Data, and Business-Consumer Alignment derived from the studies are sector-specialized. They must be validated for their generalizability and robustness through future research in other domains.

Although the intense focus on qualitative methods such as interviews, focus groups, and workshops enables detailed exploration of the respective perspectives under consideration, these methods are susceptible to subjective bias, and only limited insights into users' actual behavior in real application contexts are possible. Therefore, in addition to qualitative analyses in other domains, future research should also examine the robustness of the results using quantitative approaches.

The prototypes developed for recommender and algorithm-based reputation systems were developed and tested in controlled environments that do not cover the full complexity of real use cases. They should, therefore, be evaluated in real-world applications. In addition, there is the limitation that the effects on consumption practices by the analyzed DDSs should be investigated in the long term. Accordingly, the appropriation of the DDSs in the context of the practices is not examined. Therefore, future research should conduct long-term studies to assess whether and how the changes in consumption practices promoted by DDSs can be transferred.

Furthermore, the frameworks of Business-Data, Consumer-Data, and Business-Consumer Alignment derived from the context of this thesis are worked out based on the case studies. More empirical analyses must consider their triangulation in real-world DDS developments and demonstrate their operationalization. In addition, the tools developed to support the frameworks in larger and other real-world environments should also be evaluated.

References

[1] ABDELKAFI, N., RAASCH, C., ROTH, A., AND SRINIVASAN, R. Multi-sided platforms, 2019.

[2] ABDUL, A., VERMEULEN, J., WANG, D., LIM, B. Y., AND KANKANHALLI, M. Trends and trajectories for explainable, accountable and intelligible systems: An hci research agenda. In *Proceedings of the 2018 CHI conference on human factors in computing systems* (2018), pp. 1–18.

[3] ABERG, J. Dealing with malnutrition: A meal planning system for elderly. In *AAAI spring symposium: argumentation for consumers of healthcare* (2006), pp. 1–7.

[4] ABRAS, C., MALONEY-KRICHMAR, D., PREECE, J., ET AL. User-centered design. *Bainbridge, W. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications 37*, 4 (2004), 445–456.

[5] ABRELL, T. Business-centred, data-centred, or user-centred? a perspective on the role of designerly approaches to user centricity in big data innovation. In *Smart Services Summit: Digital as an Enabler for Smart Service Business Development*. Springer, 2021, pp. 131–141.

[6] ACKOFF, R. L. From data to wisdom. *Journal of applied systems analysis 16*, 1 (1989), 3–9.

[7] ACQUISTI, A. Privacy in electronic commerce and the economics of immediate gratification. In *Proceedings of the 5th ACM conference on Electronic commerce* (2004), pp. 21–29.

[8] ACQUISTI, A., BRANDIMARTE, L., AND LOEWENSTEIN, G. Privacy and human behavior in the age of information. *Science 347*, 6221 (2015), 509–514.

[9] ADOMAVICIUS, G., AND TUZHILIN, A. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible

extensions. *IEEE transactions on knowledge and data engineering 17*, 6 (2005), 734–749.

[10] ADRODEGARI, F., BACCHETTI, A., SACCANI, N., ARNAIZ, A., AND MEIREN, T. The transition towards service-oriented business models: A european survey on capital goods manufacturers. *International Journal of Engineering Business Management 10* (2018), 1847979018754469.

[11] AGARWAL, D., CHEN, B.-C., AND PANG, B. Personalized recommendation of user comments via factor models. In *Proceedings of the 2011 conference on empirical methods in natural language processing* (2011), pp. 571–582.

[12] AGARWAL, R., AND VENKATESH, V. Assessing a firm's web presence: a heuristic evaluation procedure for the measurement of usability. *Information systems research 13*, 2 (2002), 168–186.

[13] AGGARWAL, C. C. An introduction to recommender systems. In *Recommender systems*. Springer, 2016, pp. 1–28.

[14] AGGARWAL, C. C. *Recommender Systems*. Springer International Publishing, 2016.

[15] AGRAWAL, A., GANS, J., AND GOLDFARB, A. A simple tool to start making decisions with the help of ai. *Harvard Business Review* (2018).

[16] AHMED, M. A. O., DIDACI, L., LAVI, B., AND FUMERA, G. Using diversity for classifier ensemble pruning: an empirical investigation. *Theoretical and Applied Informatics 1*, 29 (2017).

[17] AHN, J., WILLIAMSON, J., GARTRELL, M., HAN, R., LV, Q., AND MISHRA, S. Supporting healthy grocery shopping via mobile augmented reality. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM) 12*, 1s (2015), 1–24.

[18] AIGNER, W., MIKSCH, S., SCHUMANN, H., AND TOMINSKI, C. *Visualization of time-oriented data*. Springer Science & Business Media, 2011.

[19] AKERLOF, G. A. The market for "lemons": Quality uncertainty and the market mechanism. In *Uncertainty in economics*. Elsevier, 1978, pp. 235–251.

[20] AL-DEBEI, M. M., AND AVISON, D. Developing a unified framework of the business model concept. *European journal of information systems 19*, 3 (2010), 359–376.

[21] ALANI, H., JONES, C., AND TUDHOPE, D. Associative and spatial relationships in thesaurus-based retrieval. In *International Conference on Theory and Practice of Digital Libraries* (2000), Springer, pp. 45–58.

[22] ALIZADEH, F., JAKOBI, T., BOLDT, J., AND STEVENS, G. Gdpr-reality check on the right to access data: claiming and investigating personally identifiable data from companies. In *Proceedings of Mensch und Computer 2019*. 2019, pp. 811–814.

[23] ALKHATIB, A. To live in their utopia: Why algorithmic systems create absurd outcomes. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (2021), pp. 1–9.

[24] AMARILLI, F., VAN DEN HOOFF, B., AND VAN VLIET, M. Business-it alignment as a coevolution process: An empirical study. *The Journal of Strategic Information Systems 32*, 2 (2023), 101776.

[25] ANAND, A., SEN, P., SAHA, S., VERMA, M., AND MITRA, M. Explainable information retrieval. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval* (2023), pp. 3448–3451.

[26] ANDERSON, J. Data engineers vs. data scientists, Apr. 2018.

[27] ANDREEVA, G., ANSELL, J., AND CROOK, J. Impact of anti-discrimination laws on credit scoring. *Journal of Financial Services Marketing 9*, 1 (2004), 22–33.

[28] ANSHARI, M., ALMUNAWAR, M. N., LIM, S. A., AND AL-MUDIMIGH, A. Customer relationship management and big data enabled: Personalization & customization of services. *Applied Computing and Informatics 15*, 2 (2019), 94–101.

[29] ANSOFF, H. I., DECLERCK, R. P., AND HAYES, R. L. From strategic planning to strategic management. *Strategische Unternehmungsplanung/Strategische Unternehmungsführung: Stand und Entwicklungstendenzen* (1997), 105–143.

[30] ARGYRIS, C. Actionable knowledge: Design causality in the service of consequential theory. *The Journal of applied behavioral science 32*, 4 (1996), 390–406.

[31] ARIKPO, I., OGBAN, F., AND ETENG, I. Von neumann architecture and modern computers. *Global Journal of Mathematical Sciences 6*, 2 (2007), 97–103.

[32] ARNOLD, A. *Climate change and storytelling: Narratives and cultural meaning in environmental communication*. Springer, 2018.

[33] ARRIETA, A. B., DÍAZ-RODRÍGUEZ, N., DEL SER, J., BENNETOT, A., TABIK, S., BARBADO, A., GARCÍA, S., GIL-LÓPEZ, S., MOLINA, D., BENJAMINS, R., ET AL. Explainable artificial intelligence (xai): Concepts, taxonomies, opportunities and challenges toward responsible ai. *Information Fusion 58* (2020), 82–115.

[34] ASCHEMANN-WITZEL, J., AND ZIELKE, S. Can't buy me green? a review of consumer perceptions of and behavior toward the price of organic food. *Journal of Consumer Affairs 51*, 1 (sep 2015), 211–251.

[35] ATHEY, S., CASTILLO, J. C., AND CHANDAR, B. Service quality in the gig economy: Empirical evidence about driving quality at uber. *Available at SSRN* (2019).

[36] AVELLANEDA, O., BACHMANN, A., BARBIR, A., BRENAN, J., DINGLE, P., DUFFY, K. H., MALER, E., REED, D., AND SPORNY,

M. Decentralized identity: Where did it come from and where is it going? *IEEE Communications Standards Magazine 3*, 4 (2019), 10–13.

[37] AWAD, E., AND GHAZIRI, H. *Knowledge Management*. Upper Saddle River, NJ, Pearson Education Inc., 2004.

[38] AYRES, L. Semi-structured interview. *The SAGE encyclopedia of qualitative research methods 8* (2008), 811–813.

[39] BACH, B., RICHE, N. H., CARPENDALE, S., AND PFISTER, H. The emerging genre of data comics. *IEEE computer graphics and applications 37*, 3 (2017), 6–13.

[40] BALLÚS-ARMET, I., SHAHEEN, S. A., CLONTS, K., AND WEINZIMMER, D. Peer-to-peer carsharing: Exploring public perception and market characteristics in the san francisco bay area, california. *Transportation Research Record 2416*, 1 (2014), 27–36.

[41] BARNES, J. G. *Build your customer strategy: a guide to creating profitable customer relationships*. John Wiley & Sons, 2007.

[42] BAROCAS, S., AND SELBST, A. D. Big data's disparate impact. *Calif. L. Rev. 104* (2016), 671.

[43] BASKERVILLE, R. L., KAUL, M., AND STOREY, V. C. Genres of inquiry in design-science research: Justification and evaluation of knowledge production. *Mis Quarterly 39*, 3 (2015), 541–564.

[44] BATMETAN, J. R., RAWIS, J. A., LENGKONG, J. S. J., AND ROTTY, V. N. J. Future trends for direction in enterprise architecture: systematic literature review. *International Journal of Information Technology and Education 2*, 3 (2023), 1–20.

[45] BAUMAN, Z. *In search of politics*. Stanford University Press, Stanford, Calif, 1999.

[46] BAUMEISTER, R. F. Writing a literature review. *The portable mentor: Expert guide to a successful career in psychology* (2013), 119–132.

[47] BEGGS, J. N. Private-sector nudging: the good, the bad, and the uncertain. In *Nudge Theory in Action: Behavioral Design in Policy and Markets*. Springer, 2016, pp. 125–158.

[48] BEIMBORN, D., FRANKE, J., WAGNER, H.-T., AND WEITZEL, T. The influence of alignment on the post-implementation success of a core banking information system: An embedded case study. In *2007 40th Annual Hawaii International Conference on System Sciences (HICSS'07)* (2007), IEEE, pp. 234b–234b.

[49] BEIMBORN, D., MILETZKI, T., AND WENZEL, S. Platform as a service (paas). *Business & Information Systems Engineering 3*, 6 (2011), 381–384.

[50] BÉLANGER, F., AND CROSSLER, R. E. Privacy in the digital age: a review of information privacy research in information systems. *MIS quarterly* (2011), 1017–1041.

[51] BELK, R. Why not share rather than own? *The Annals of the American Academy of Political and Social Science 611*, 1 (2007), 126–140.

[52] BELK, R. Sharing. *Journal of consumer research 36*, 5 (2010), 715–734.

[53] BELK, R. W. Possessions and the extended self. *Journal of consumer research 15*, 2 (1988), 139–168.

[54] BELL, E., BRYMAN, A., AND HARLEY, B. *Business research methods*. Oxford university press, 2018.

[55] BELLINGER, G., CASTRO, D., AND MILLS, A. Data, information, knowledge, and wisdom, 2004.

[56] BENJAAFAR, S., KONG, G., LI, X., AND COURCOUBETIS, C. Peer-to-peer product sharing: Implications for ownership, usage, and social welfare in the sharing economy. *Management Science 65*, 2 (2019), 477–493.

[57] BENTE, G., BAPTIST, O., AND LEUSCHNER, H. To buy or not to buy: Influence of seller photos and reputation on buyer trust and purchase behavior. *International Journal of Human-Computer Studies 70*, 1 (2012), 1–13.

[58] BERENDS, L., AND JOHNSTON, J. Using multiple coders to enhance qualitative analysis: The case of interviews with consumers of drug treatment. *Addiction Research & Theory 13*, 4 (2005), 373–381.

[59] BERNARDES, M. B., DE ANDRADE, F. P., AND NOVAIS, P. Smart cities, data and right to privacy: A look from the portuguese and brazilian experience. In *Proceedings of the 11th International Conference on Theory and Practice of Electronic Governance* (2018), pp. 328–337.

[60] BEVAN, N. What is the difference between the purpose of usability and user experience evaluation methods. In *Proceedings of the Workshop UXEM* (2009), vol. 9, Citeseer, pp. 1–4.

[61] BEVERUNGEN, D., MÜLLER, O., MATZNER, M., MENDLING, J., AND VOM BROCKE, J. Conceptualizing smart service systems. *Electronic Markets 29* (2019), 7–18.

[62] BHAGAT, R., MURALIDHARAN, S., LOBZHANIDZE, A., AND VISHWANATH, S. Buy it again: Modeling repeat purchase recommendations. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (2018), pp. 62–70.

[63] BIRAN, O., AND COTTON, C. Explanation and justification in machine learning: A survey. In *IJCAI-17 workshop on explainable AI (XAI)* (2017), vol. 8, pp. 8–13.

[64] BIRD, J., KALNIKAITÉ, V., AND ROGERS, Y. The augmented shopping trolley: An ambient display to provide shoppers with non-obvious product information. In *Proceedings of the 2nd International Workshop on Persuasion, Influence, Nudge & Coercion Through Mobile Devices* (2011), pp. 3–6.

[65] BIRNSTILL, P., AND BEYERER, J. Building blocks for identity management and protection for smart environments and interactive assistance systems. In *Proceedings of the 11th PErvasive Technologies Related to Assistive Environments Conference* (2018), pp. 292–296.

[66] BISHOP, C. M. *Pattern recognition and machine learning*. springer, 2006.

[67] BITKOM. Digital-Design-Manifest, 2018.

[68] BJÖRGVINSSON, E., EHN, P., AND HILLGREN, P.-A. Participatory design and" democratizing innovation". In *Proceedings of the 11th Biennial participatory design conference* (2010), pp. 41–50.

[69] BLADT, J., AND FILBIN, B. A data scientist's real job: Storytelling. *Harvard Business Review* (2013).

[70] BLÖCHER, K., REINHOLD, O., AND ALT, R. Managing customer data in data-driven service innovation: A framework of data principles. *18th International Conference on Wirtschaftsinformatik* (2023).

[71] BLOMBERG, J. On participation and service innovation. *Re) Searching the Digital Bauhaus. Springer* (2009).

[72] BLOMKVIST, S. Towards a model for bridging agile development and user-centered design. In *Human-centered software engineering—integrating usability in the software development lifecycle*. Springer, 2005, pp. 219–244.

[73] BOBADILLA, J., ORTEGA, F., HERNANDO, A., AND GUTIÉRREZ, A. Recommender systems survey. *Knowledge-based systems 46* (2013), 109–132.

[74] BODDY, D., BOONSTRA, A., AND KENNEDY, G. *Managing information systems: an organisational perspective*. Pearson Education, 2005.

[75] BODEN, A., STEVENS, G., BERKHOLZ, J., AND SCHREIBER, D. Einordnung und hintergrund. In *Verbraucherinformatik: Grundla-*

gen und Anwendungsfelder der digitalen Konsumgesellschaft. Springer Berlin Heidelberg Berlin, Heidelberg, 2024, pp. 1–27.

[76] BOGNER, A., LITTIG, B., AND MENZ, W. *Interviewing experts*. Springer, 2009.

[77] BOGNER, A., LITTIG, B., AND MENZ, W. Introduction: Expert interviews—an introduction to a new methodological debate. In *Interviewing experts*. Springer, 2009, pp. 1–13.

[78] BOGNER, A., AND MENZ, W. Das theoriegenerierende experteninterview: Erkenntnisinteresse, wissensformen, interaktion. *Das Experteninterview: Theorie, Methode, Anwendung* (2002), 33–70.

[79] BÖHM, L., PETERS, F., BOSSAUER, P., LAWO, D., PAKUSCH, C., AND STEVENS, G. Designing a mobility intelligence system for decision-making with shared mobility data. In *2022 International Conference on ICT for Sustainability (ICT4S)* (2022), IEEE, pp. 53–64.

[80] BOLLINGER, A. S., AND SMITH, R. D. Managing organizational knowledge as a strategic asset. *Journal of knowledge management 5*, 1 (2001), 8–18.

[81] BOLTON, G. E., KATOK, E., AND OCKENFELS, A. How effective are electronic reputation mechanisms? an experimental investigation. *Management science 50*, 11 (2004), 1587–1602.

[82] BOSSAUER, P., NEIFER, T., PAKUSCH, C., AND STASKIEWICZ, P. Using blockchain in peer-to-peer carsharing to build trust in the sharing economy.

[83] BOSSAUER, P., NEIFER, T., STEVENS, G., AND PAKUSCH, C. Trust versus privacy: Using connected car data in peer-to-peer carsharing. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (2020), pp. 1–13.

[84] BOSTMAN, R., AND ROGERS, R. What's mine is yours: How collaborative consumption is changing the way we live. *London: Collins* (2011).

[85] BRAUN, S., AND DESSLOCH, S. A classification of replicated data for the design of eventually consistent domain models. In *2020 IEEE International Conference on Software Architecture Companion (ICSA-C)* (2020), IEEE, pp. 33–40.

[86] BRAUN, V., AND CLARKE, V. Using thematic analysis in psychology. *Qualitative research in psychology 3*, 2 (2006), 77–101.

[87] BRAUN, V., CLARKE, V., AND WEATE, P. Using thematic analysis in sport and exercise research. *Routledge handbook of qualitative research in sport and exercise* (2016), 191–205.

[88] BRAZHNIK, O., AND JONES, J. F. Anatomy of data integration. *Journal of biomedical informatics 40*, 3 (2007), 252–269.

[89] BREESE, J. S., HECKERMAN, D., AND KADIE, C. Empirical analysis of predictive algorithms for collaborative filtering. *arXiv preprint arXiv:1301.7363* (2013).

[90] BREUKER, D., PFEIFFER, D., AND BECKER, J. Reducing the variation in intra-and interorganizanional business process modeling-an empirical evaluation. In *Wirtschaftsinformatik (1)* (2009), pp. 203–212.

[91] BRINKMANN, S. Narrative inquiry, field research, and interview methods, unstructured and semi-structured interviewing, p. leavy, 2014.

[92] BRITO, L. D. C., AND QUARESMA, M. M. R. Ucd: Influenced by an organizational culture and its maturity. In *Human-Computer Interaction. Theories, Methods, and Human Issues: 20th International Conference, HCI International 2018, Las Vegas, NV, USA, July 15–20, 2018, Proceedings, Part I 20* (2018), Springer, pp. 14–23.

[93] BROCKE, J. V., SIMONS, A., NIEHAVES, B., NIEHAVES, B., REIMER, K., PLATTFAUT, R., AND CLEVEN, A. Reconstructing the giant: On the importance of rigour in documenting the literature search process.

[94] BROWN, T., AND KATZ, B. Change by design. *Journal of product innovation management 28*, 3 (2011), 381–383.

[95] BRUSILOVSKY, P. Adaptive navigation support. In *The adaptive web*. Springer, 2007, pp. 263–290.

[96] BRYNJOLFSSON, E., AND MCAFEE, A. *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. WW Norton & Company, 2014.

[97] BRYNJOLFSSON, E., AND MCAFEE, A. The business of artificial intelligence. *Harvard Business Review* (2017), 1–20.

[98] BRYNJOLFSSON, E., AND MITCHELL, T. What can machine learning do? workforce implications. *Science 358*, 6370 (2017), 1530–1534.

[99] BULGER, M., TAYLOR, G., AND SCHROEDER, R. Data-driven business models: challenges and opportunities of big data. *Oxford Internet Institute. Research Councils UK: NEMODE, New Economic Models in the Digital Economy* (2014).

[100] BURRELL, J. How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big data & society 3*, 1 (2016), 2053951715622512.

[101] BUZZELL, R. D. *The PIMS principles: Linking strategy to performance*. The Free Press, 1987.

[102] CACHEDA, F., CARNEIRO, V., FERNÁNDEZ, D., AND FORMOSO, V. Comparison of collaborative filtering algorithms: Limitations of current techniques and proposals for scalable, high-performance recommender systems. *ACM Transactions on the Web (TWEB) 5*, 1 (2011), 1–33.

[103] CALLAHAN, S. The role of stories in data storytelling. `https://www.anecdote.com/2016/08/stories-data-storytelling/`, 2016. (Accessed on 19/05/2023).

[104] CALO, R., AND ROSENBLAT, A. The taking economy: Uber, information, and power. *Colum. L. Rev. 117* (2017), 1623.

[105] CARFAGNA, L. B., DUBOIS, E. A., FITZMAURICE, C., OUIMETTE, M. Y., SCHOR, J. B., WILLIS, M., AND LAIDLEY, T. An emerging eco-habitus: The reconfiguration of high cultural capital practices among ethical consumers. *Journal of Consumer Culture 14*, 2 (2014), 158–178.

[106] CARNEIN, M., HOMANN, L., TRAUTMANN, H., AND VOSSEN, G. A recommender system based on omni-channel customer data. In *2019 IEEE 21st Conference on Business Informatics (CBI)* (2019), vol. 1, IEEE, pp. 65–74.

[107] CAROL, S., EICH, D., KELLER, M., STEINER, F., AND STORZ, K. Who can ride along? discrimination in a german carpooling market. *Population, space and place 25*, 8 (2019), e2249.

[108] CARR, N. G. It doesn't matter. *Educause Review 38* (2003), 24–38.

[109] CARRINGTON, M. J., NEVILLE, B. A., AND WHITWELL, G. J. Lost in translation: Exploring the ethical consumer intention–behavior gap. *Journal of Business Research 67*, 1 (2014), 2759–2767.

[110] CARRINGTON, M. J., NEVILLE, B. A., AND WHITWELL, G. J. Lost in translation: Exploring the ethical consumer intention–behavior gap. *Journal of Business Research 67*, 1 (jan 2014), 2759–2767.

[111] CĂRUȚAȘU, G., BOTEZATU, C., BOTEZATU, M., ET AL. Expanding ecall from cars to other means of transport. *Journal of Information Systems & Operations Management 10*, 2 (2016), 354–363.

[112] CARVALHO, A. D. P., DA CUNHA, S. K., DE LIMA, L. F., AND CARSTENS, D. D. The role and contributions of sociological institutional theory to the socio-technical approach to innovation theory. *RAI Revista de Administração e Inovação 14*, 3 (2017), 250–259.

[113] CASTELLI, N., OGONOWSKI, C., JAKOBI, T., STEIN, M., STEVENS, G., AND WULF, V. What happened in my home? an end-user development approach for smart home data visualization. In *Proceedings*

of the 2017 CHI Conference on Human Factors in Computing Systems (2017), pp. 853–866.

[114] CASTILLO-LÓPEZ, S. E., FLORES GUIRAO, S. K., ROLDÁN-REYES, E., MALDONADO MACIAS, A. A., AND DUQUE, D. M. A service design process based on the business model canvas and the ck theory. *Managing Innovation in Highly Restrictive Environments: Lessons from Latin America and Emerging Markets* (2019), 279–301.

[115] CELESTE, E. Digital sovereignty in the eu: challenges and future perspectives. *Data Protection Beyond Borders: Transatlantic Perspectives on Extraterritoriality and Sovereignty* (2021), 211–228.

[116] CHAFFEY, D., AND WOOD, S. *Business information management: Improving performance using information systems*. Financial Times Prentice Hall Harlow, 2005.

[117] CHAKRABORTI, T., ISAHAGIAN, V., KHALAF, R., KHAZAENI, Y., MUTHUSAMY, V., RIZK, Y., AND UNUVAR, M. From robotic process automation to intelligent process automation. In *International Conference on Business Process Management* (2020), Springer, pp. 215–228.

[118] CHAROENSUK, S., WONGSURAWAT, W., AND KHANG, D. B. Business-it alignment: A practical research approach. *The Journal of High Technology Management Research 25*, 2 (2014), 132–147.

[119] CHEN, H., CHEN, X., SHI, S., AND ZHANG, Y. Generate natural language explanations for recommendation. *arXiv preprint arXiv:2101.03392* (2021).

[120] CHEN, H., CHIANG, R. H., AND STOREY, V. C. Business intelligence and analytics: From big data to big impact. *MIS quarterly* (2012), 1165–1188.

[121] CHEN, M., MAO, S., AND LIU, Y. Big data: A survey. *Mobile networks and applications 19* (2014), 171–209.

[122] CHEN, P.-Y., CHOU, Y.-C., AND KAUFFMAN, R. J. Community-based recommender systems: Analyzing business models from a systems operator's perspective. In *2009 42nd Hawaii International Conference on System Sciences* (2009), IEEE, pp. 1–10.

[123] CHEN, S. C., AND DHILLON, G. S. Interpreting dimensions of consumer trust in e-commerce. *Information technology and management 4* (2003), 303–318.

[124] CHEN, Y., KREULEN, J., CAMPBELL, M., AND ABRAMS, C. Analytics ecosystem transformation: A force for business model innovation. In *2011 Annual SRII global conference* (2011), IEEE, pp. 11–20.

[125] CHHETRI, C. Towards a smart home usable privacy framework. In *Companion Publication of the 2019 Conference on Computer Supported Cooperative Work and Social Computing* (2019), pp. 43–46.

[126] CHILANA, P. K., KO, A. J., AND WOBBROCK, J. From user-centered to adoption-centered design: a case study of an hci research innovation becoming a product. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (2015), pp. 1749–1758.

[127] CHO, Y. B., CHO, Y. H., AND KIM, S. H. Mining changes in customer buying behavior for collaborative recommendations. *Expert Systems with Applications 28*, 2 (2005), 359–369.

[128] CHRISTOZOV, D., RASHEVA-YORDANOVA, K., AND TOLEVA-STOIMENOVA, S. Risks management in data science training. In *Proceeding of Regional international conference on Applied Protection and ITS Trends, Zlatibor. S* (2018), pp. 7–10.

[129] CHU, W., AND PARK, S.-T. Personalized recommendation on dynamic content using predictive bilinear models. In *Proceedings of the 18th international conference on World wide web* (2009), pp. 691–700.

[130] CHUNG, S. H., RAINER JR, R. K., AND LEWIS, B. R. The impact of information technology infrastructure flexibility on strategic alignment

and application implementations. *Communications of the Association for Information Systems 11*, 1 (2003), 11.

[131] CIELEN, D., MEYSMAN, A., AND ALI, M. *Introducing data science: big data, machine learning, and more, using Python tools*. Manning Publications Co., 2016.

[132] CLAYPOOL, M., GOKHALE, A., MIRANDA, T., MURNIKOV, P., NETES, D., AND SARTIN, M. Combing content-based and collaborative filters in an online newspaper. In *Proc. of Workshop on Recommender Systems-Implementation and Evaluation* (1999).

[133] CLEAR, A., AND FRIDAY, A. Designing a food'qualculator'. In *DIS 2012 workshop on Food for Thought: Designing for Critical Reflection on Food Practices. Newcastle, UK* (2012).

[134] CLEMENT, R., SCHREIBER, D., BOSSAUER, P., AND PAKUSCH, C. Intermediation: Direkte und indirekte verknüpfung von angebot und nachfrage. In *Internet-Ökonomie*. Springer, 2019, pp. 153–204.

[135] CLEMENT, R., SCHREIBER, D., BOSSAUER, P., AND PAKUSCH, C. Internet-ökonomie, grundlagen und fallbeispiele der digitalen und vernetzten wirtschaft, 4. *Aufl., Berlin* (2019).

[136] COHEN, B., AND KIETZMANN, J. Ride on! mobility business models for the sharing economy. *Organization & environment 27*, 3 (2014), 279–296.

[137] COHEN, J. E. Examined lives: Informational privacy and the subject as object. In *Law and Society Approaches to Cyberspace*. Routledge, 2017, pp. 473–538.

[138] COLLIS, J., AND HUSSEY, R. *Business research: A practical guide for undergraduate and postgraduate students*. Macmillan International Higher Education, 2013.

[139] COMPANY, F. M. Ford Survey: Drive My Car? Many Europeans Open to Sharing Their Car, Home, Phone – Even Their Dog. Tech. rep., Ford Motor Company, 2016.

[140] COMUZZI, M., AND PATEL, A. How organisations leverage big data: A maturity model. *Industrial management & Data systems* (2016).

[141] CONRADIE, N. H., AND NAGEL, S. K. Digital sovereignty and smart wearables: Three moral calculi for the distribution of legitimate control over the digital. *Journal of responsible technology 12* (2022), 100053.

[142] CONSUL, P. C., AND JAIN, G. C. A generalization of the poisson distribution. *Technometrics 15*, 4 (1973), 791–799.

[143] COOPER, H. M. Organizing knowledge syntheses: A taxonomy of literature reviews. *Knowledge in society 1*, 1 (1988), 104.

[144] COPPOLA, R., AND MORISIO, M. Connected car: technologies, issues, future trends. *ACM Computing Surveys (CSUR) 49*, 3 (2016), 1–36.

[145] CORBITT, B. J., THANASANKIT, T., AND YI, H. Trust and e-commerce: a study of consumer perceptions. *Electronic commerce research and applications 2*, 3 (2003), 203–215.

[146] CRESWELL, J. W., AND CLARK, V. L. P. *Designing and conducting mixed methods research*. Sage publications, 2017.

[147] CURRY, E. The big data value chain: definitions, concepts, and theoretical approaches. *New horizons for a data-driven economy: A roadmap for usage and exploitation of big data in Europe* (2016), 29–37.

[148] DARAIO, C., AND GLÄNZEL, W. Grand challenges in data integration—state of the art and future perspectives: An introduction. *Scientometrics 108*, 1 (2016), 391–400.

[149] DATTA, A., SEN, S., AND ZICK, Y. Algorithmic transparency via quantitative input influence: Theory and experiments with learning systems. In *2016 IEEE symposium on security and privacy (SP)* (2016), IEEE, pp. 598–617.

[150] DAUTZENBERG, K., GRASL, K., HALM, T., AND SCHULZE, A. Buchungs- und vergleichs-portale – die schwierige suche nach dem günstigsten preis. 80.

[151] DAVENPORT, T. H. Analytics 3.0. *Harvard business review 91*, 12 (2013), 64–72.

[152] DAVENPORT, T. H., AND KIM, J. *Keeping up with the quants: Your guide to understanding and using analytics*. Harvard Business Review Press, 2013.

[153] DAVENPORT, T. H., AND PATIL, D. Data scientist. *Harvard business review 90*, 5 (2012), 70–76.

[154] DAVENPORT, T. H., AND WESTERMAN, G. Why so many high-profile digital transformations fail. *Harvard Business Review 9*, 4 (2018), 15.

[155] DAVISON, R. M. The art of storytelling, 2016.

[156] DE HAES, S., AND VAN GREMBERGEN, W. An exploratory study into it governance implementations and its impact on business/it alignment. *Information Systems Management 26*, 2 (2009), 123–137.

[157] DE MAURO, A., GRECO, M., GRIMALDI, M., AND NOBILI, G. Beyond data scientists: a review of big data skills and job families. *Proceedings of IFKAD* (2016), 1844–1857.

[158] DE OLIVEIRA, V. Hierarchical poisson models for spatial count data. *Journal of Multivariate Analysis 122* (2013), 393–408.

[159] DEHNING, B., RICHARDSON, V. J., AND ZMUD, R. W. The value relevance of announcements of transformational information technology investments. *MIS quarterly* (2003), 637–656.

[160] DELLAROCAS, C., AND NARAYAN, R. What motivates consumers to review a product online? a study of the product-specific antecedents of online movie reviews. In *WISE* (2006), pp. 1–6.

[161] DEMIRKAN, H., BESS, C., SPOHRER, J., RAYES, A., ALLEN, D., AND MOGHADDAM, Y. Innovations with smart service systems: analytics, big data, cognitive assistance, and the internet of everything. *Communications of the association for Information Systems 37*, 1 (2015), 35.

[162] DERIKX, S., DE REUVER, M., AND KROESEN, M. Can privacy concerns for insurance of connected cars be compensated? *Electronic Markets 26*, 1 (2016), 73–81.

[163] DESPOTOVIC, Z., AND ABERER, K. Possibilities for managing trust in p2p networks.

[164] DESYLLAS, P., AND SAKO, M. Profiting from business model innovation: Evidence from pay-as-you-drive auto insurance. *Research Policy 42*, 1 (2013), 101–116.

[165] DEY, S., MITRA, P., AND GUPTA, K. Recommending repeat purchases using product segment statistics. In *Proceedings of the 10th ACM Conference on Recommender Systems* (2016), pp. 357–360.

[166] DI VALENTIN, C., BURKHART, T., VANDERHAEGHEN, D., WERTH, D., AND LOOS, P. Towards a framework for transforming business models into business processes.

[167] DINEV, T., AND HART, P. An extended privacy calculus model for e-commerce transactions. *Information systems research 17*, 1 (2006), 61–80.

[168] DING, Y., LI, X., AND ORLOWSKA, M. E. Recency-based collaborative filtering. In *Proceedings of the 17th Australasian Database Conference-Volume 49* (2006), pp. 99–107.

[169] DO, M.-P. T., NGUYEN, D., AND NGUYEN, L. Model-based approach for collaborative filtering. In *6th International Conference on Information Technology for Education* (2010).

[170] DOKRAS, N. S. *Prime Pantry Optimization: a cost analysis and deep-dive in process improvement*. PhD thesis, Massachusetts Institute of Technology, 2017.

[171] DONG, X. L., AND REKATSINAS, T. Data integration and machine learning: A natural synergy. In *Proceedings of the 2018 international conference on management of data* (2018), pp. 1645–1650.

[172] DONOHO, D. 50 years of data science. *Journal of Computational and Graphical Statistics 26*, 4 (2017), 745–766.

[173] DORARD, L. Machine Learning Canvas.

[174] DÖRINGER, S. 'the problem-centred expert interview'. combining qualitative interviewing approaches for investigating implicit expert knowledge. *International journal of social research methodology 24*, 3 (2021), 265–278.

[175] DOSHI-VELEZ, F., AND KIM, B. Towards a rigorous science of interpretable machine learning. *arXiv preprint arXiv:1702.08608* (2017).

[176] DUAN, Y., EDWARDS, J. S., AND DWIVEDI, Y. K. Artificial intelligence for decision making in the era of big data–evolution, challenges and research agenda. *International journal of information management 48* (2019), 63–71.

[177] DYKES, B. Data storytelling: The essential data science skill everyone needs. *Forbes Magazine* (2016).

[178] EBERT, N., WEBER, K., AND KORUNA, S. Integration platform as a service. *Business & Information Systems Engineering 59*, 5 (2017), 375–379.

[179] EDELMAN, B., LUCA, M., AND SVIRSKY, D. Racial discrimination in the sharing economy: Evidence from a field experiment. *American economic journal: applied economics 9*, 2 (2017), 1–22.

[180] EDWARDS, L., AND VEALE, M. Slave to the algorithm: Why a right to an explanation is probably not the remedy you are looking for. *Duke L. & Tech. Rev. 16* (2017), 18.

[181] EHN, P. *Work-oriented design of computer artifacts*. PhD thesis, Arbetslivscentrum, 1988.

[182] EHRENBERG, A. S. Repeat buying. *Journal of Empirical Generalisations in Marketing Science 5*, 2 (2000).

[183] ELAHI, M., GE, M., RICCI, F., FERNÁNDEZ-TOBÍAS, I., BERKOVSKY, S., AND MASSIMO, D. Interaction design in a mobile food recommender system. In *IntRS 2015 Interfaces and Human Decision Making for Recommender Systems: Proceedings of the Joint Workshop on Interfaces and Human Decision Making for Recommender Systems co-located with ACM Conference on Recommender Systems (RecSys 2015)* (2015), vol. 1438, CEUR-WS, pp. 49–52.

[184] ELAHI, M., GE, M., RICCI, F., MASSIMO, D., AND BERKOVSKY, S. Interactive food recommendation for groups. In *Recsys posters* (2014), Citeseer.

[185] ELSWEILER, D., HARVEY, M., LUDWIG, B., AND SAID, A. Bringing the" healthy" into food recommenders. In *DMRS* (2015), pp. 33–36.

[186] ENTWISTLE, J. M., RASMUSSEN, M. K., VERDEZOTO, N., BREWER, R. S., AND ANDERSEN, M. S. Beyond the individual. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. ACM. https://doi. org/10* (2015), vol. 1145.

[187] EPPLER, M. J., AND MENGIS, J. The concept of information overload-a review of literature from organization science, accounting, marketing, mis, and related disciplines (2004) the information society: An international journal, 20 (5), 2004, pp. 1–20. *Kommunikationsmanagement im Wandel: Beiträge aus 10 Jahren= mcminstitute* (2008), 271–305.

[188] ERT, E., FLEISCHER, A., AND MAGEN, N. Trust and reputation in the sharing economy: The role of personal photos in airbnb. *Tourism management 55* (2016), 62–73.

[189] ESAU-HELD, M., KRAUSS, V., AND ESSING, B. Digitale gestaltung. In *Verbraucherinformatik: Grundlagen und Anwendungsfelder der digitalen Konsumgesellschaft*. Springer Berlin Heidelberg Berlin, Heidelberg, 2024, pp. 261–300.

[190] EVANS, D. S., AND SCHMALENSEE, R. Failure to launch: Critical mass in platform businesses. *Review of Network Economics 9*, 4 (2010).

[191] FADER, P. S., AND HARDIE, B. G. Probability models for customer-base analysis. *Journal of interactive marketing 23*, 1 (2009), 61–69.

[192] FADER, P. S., HARDIE, B. G., AND LEE, K. L. "counting your customers" the easy way: An alternative to the pareto/nbd model. *Marketing science 24*, 2 (2005), 275–284.

[193] FARR-WHARTON, G., CHOI, J. H.-J., AND FOTH, M. Food talks back: exploring the role of mobile applications in reducing domestic food wastage. In *Proceedings of the 26th Australian computer-human interaction conference on designing futures: The future of design* (2014), pp. 352–361.

[194] FEENEY, M. companies uber, r.(2015). *Is Ridesharing Safe* (2015).

[195] FERNANDES, G. B., AND ARTES, R. Spatial dependence in credit risk and its improvement in credit scoring. *European Journal of Operational Research 249*, 2 (2016), 517–524.

[196] FILIPCZUK, D., BAARSLAG, T., GERDING, E. H., AND SCHRAEFEL, M. Automated privacy negotiations with preference uncertainty. *Autonomous Agents and Multi-Agent Systems 36*, 2 (2022), 49.

[197] FILZMOSER, M., AND VETSCHERA, R. A classification of bargaining steps and their impact on negotiation outcomes. *Group Decision and Negotiation 17* (2008), 421–443.

[198] FINK, A. *Conducting research literature reviews: From the internet to paper*. Sage publications, 2019.

[199] FORBES, C., EVANS, M., HASTINGS, N., AND PEACOCK, B. *Statistical distributions*. John Wiley & Sons, 2011.

[200] FOURNIER, S. Consumers and their brands: Developing relationship theory in consumer research. *Journal of consumer research 24*, 4 (1998), 343–373.

[201] FRANSEN, M. L., VERLEGH, P. W., KIRMANI, A., AND SMIT, E. G. A typology of consumer strategies for resisting advertising, and a review of mechanisms for countering them. *International journal of advertising 34*, 1 (2015), 6–16.

[202] FREHNER, A., DE BOER, I. J., MULLER, A., VAN ZANTEN, H. H., AND SCHADER, C. Consumer strategies towards a more sustainable food system: insights from switzerland. *The American Journal of Clinical Nutrition 115*, 4 (2022), 1039–1047.

[203] FREYNE, J., AND BERKOVSKY, S. Intelligent food planning. In *Proceedings of the 15th international conference on Intelligent user interfaces* (2010), pp. 321–324.

[204] FRIEDMAN, B., AND NISSENBAUM, H. Bias in computer systems. *ACM Transactions on information systems (TOIS) 14*, 3 (1996), 330–347.

[205] FU, X., BUDZIK, J., AND HAMMOND, K. J. Mining navigation history for recommendation. In *Proceedings of the 5th international conference on Intelligent user interfaces* (2000), pp. 106–112.

[206] GABRIELLI, S., FORBES, P., JYLHÄ, A., WELLS, S., SIRÉN, M., HEMMINKI, S., NURMI, P., MAIMONE, R., MASTHOFF, J., AND JACUCCI, G. Design challenges in motivating change for sustainable urban mobility. *Computers in Human Behavior 41* (2014), 416–423.

[207] GADATSCH, A. Digitization: Changing framework conditions in work and society. In *IT Controlling: From IT cost and activity allocation to smart controlling*. Springer, 2023, pp. 1–10.

[208] GADATSCH, A., AND LANDROCK, H. Big data für entscheider. *Entwicklung und Umsetzung datengetriebener Geschäftsmodelle. Wiesbaden (essentials). Online verfügbar unter http://dx. doi. org/10.1007/978-3-658-17340-1* (2017).

[209] GAJARDO, P. A., AND LA PAZ, A. I. Business-it alignment in the digital age, an empirical analysis. *International Journal of Innovation in Management 7*, 2 (2019), 89–100.

[210] GANDHI, M., AND GANDHI, S. A case study on use of telematics for car sharing business in india. *Journal of Applied Management-Jidnyasa 10*, 2 (2018), 15.

[211] GANDOMI, A., AND HAIDER, M. Beyond the hype: Big data concepts, methods, and analytics. *International journal of information management 35*, 2 (2015), 137–144.

[212] GANGLBAUER, E., FITZPATRICK, G., AND COMBER, R. Negotiating food waste: Using a practice lens to inform design. *ACM Transactions on Computer-Human Interaction (TOCHI) 20*, 2 (2013), 1–25.

[213] GASSMANN, O., FRANKENBERGER, K., AND CSIK, M. *The business model navigator: 55 models that will revolutionise your business*. Pearson UK, 2014.

[214] GATERSLEBEN, B. Affective and symbolic aspects of car use. In *Threats from car traffic to the quality of urban life: Problems, Causes and Solutions*. Emerald Group Publishing Limited, 2007, pp. 219–233.

[215] GAY, V., AND LEIJDEKKERS, P. Bringing health and fitness data together for connected health care: mobile apps as enablers of interoperability. *Journal of medical Internet research 17*, 11 (2015), e260.

[216] GE, Y., KNITTEL, C. R., MACKENZIE, D., AND ZOEPF, S. Racial and gender discrimination in transportation network companies. Tech. rep., National Bureau of Economic Research, 2016.

[217] GEELS, F. W. From sectoral systems of innovation to socio-technical systems: Insights about dynamics and change from sociology and institutional theory. *Research policy 33*, 6-7 (2004), 897–920.

[218] GEELS, F. W. *Technological transitions and system innovations: a co-evolutionary and socio-technical analysis*. Edward Elgar Publishing, 2005.

[219] GEFEN, D. Reflections on the dimensions of trust and trustworthiness among online consumers. *ACM SIGMIS Database: the DATABASE for Advances in Information Systems 33*, 3 (2002), 38–53.

[220] GÉRON, A. *Praxiseinstieg Machine Learning mit Scikit-Learn und TensorFlow: Konzepte, Tools und Techniken für intelligente Systeme*. O'Reilly, 2018.

[221] GEYER-SCHULZ, A., HAHSLER, M., AND JAHN, M. A customer purchase incidence model applied to recommender services. In *International Workshop on Mining Web Log Data Across All Customers Touch Points* (2001), Springer, pp. 25–47.

[222] GIESEL, F., AND NOBIS, C. The impact of carsharing on car ownership in german cities. *Transportation Research Procedia 19* (2016), 215–224.

[223] GIROD, B., VAN VUUREN, D. P., GRAHN, M., KITOUS, A., KIM, S. H., AND KYLE, P. Climate impact of transportation a model comparison. *Climatic change 118* (2013), 595–608.

[224] GJERRIS, M., GAMBORG, C., AND SAXE, H. What to buy? on the complexity of being a critical consumer. *Journal of Agricultural and Environmental Ethics 29*, 1 (dec 2016), 81–102.

[225] GLÄSER, J., AND LAUDEL, G. *Experteninterviews und qualitative Inhaltsanalyse*. Springer-Verlag, 2010.

[226] GOMEZ-URIBE, C. A., AND HUNT, N. The netflix recommender system: Algorithms, business value, and innovation. *ACM Transactions on Management Information Systems (TMIS) 6*, 4 (2015), 1–19.

[227] GONG, S. A collaborative filtering recommendation algorithm based on user clustering and item clustering. *JSW 5*, 7 (2010), 745–752.

[228] GOODMAN, B., AND FLAXMAN, S. European union regulations on algorithmic decision-making and a "right to explanation". *AI magazine 38*, 3 (2017), 50–57.

[229] GOPALAN, P., HOFMAN, J. M., AND BLEI, D. M. Scalable recommendation with hierarchical poisson factorization. In *UAI* (2015), pp. 326–335.

[230] GRANDHI, S., AND CHUGH, R. Implementation strategies for erp adoption by smes. In *International Conference on Future Generation Communication and Networking* (2012), Springer, pp. 210–216.

[231] GREENBAUM, J. M., AND KYNG, M. Design at work: Cooperative design of computer systems, 1991.

[232] GREGOR, S., AND HEVNER, A. R. Positioning and presenting design science research for maximum impact. *MIS quarterly* (2013), 337–355.

[233] GROFF, T., AND JONES, T. *Introduction to knowledge management*. Routledge, 2012.

[234] GULYÁS, E. Interpretations of ethical consumption. *Review of Sociology 14*, 1 (jun 2008), 25–44.

[235] GÜNTHER, W. A., MEHRIZI, M. H. R., HUYSMAN, M., AND FELDBERG, F. Debating big data: A literature review on realizing value from big data. *The Journal of Strategic Information Systems 26*, 3 (2017), 191–209.

[236] GUPTA, U., AND CANNON, S. Data strategy. In *A Practitioner's Guide to Data Governance: A Case-based Approach*. Emerald Publishing Limited, 2020, pp. 71–99.

[237] H. DAVENPORT, T. How strategists use "big data" to support internal business decisions, discovery and production. *Strategy & leadership 42*, 4 (2014), 45–50.

[238] HÄBERLE, T., CHARISSIS, L., FEHLING, C., NAHM, J., AND LEYMANN, F. The connected car in the cloud: a platform for prototyping telematics services. *IEEE Software 32*, 6 (2015), 11–17.

[239] HÅKANSSON, M., AND SENGERS, P. Beyond being green: simple living families and ict. In *Proceedings of the SIGCHI conference on human factors in computing systems* (2013), pp. 2725–2734.

[240] HÅKANSSON, M., AND SENGERS, P. Beyond being green: Simple living families and ICT. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (apr 2013), ACM, p. 2725.

[241] HALKIER, B., KATZ-GERRO, T., AND MARTENS, L. Applying practice theory to the study of consumption: Theoretical and methodological considerations, 2011.

[242] HALLIKAINEN, H., ALAMÄKI, A., AND LAUKKANEN, T. Individual preferences of digital touchpoints: A latent class analysis. *Journal of Retailing and Consumer Services 50* (2019), 386–393.

[243] HAMARI, J., SJÖKLINT, M., AND UKKONEN, A. The sharing economy: Why people participate in collaborative consumption. *Journal of the association for information science and technology 67*, 9 (2016), 2047–2059.

[244] HAMILTON, N. D. Food democracy ii: Revolution or restoration. *J. Food L. & Pol'y 1* (2005), 13.

[245] HAMILTON, R. Consumer-based strategy: Using multiple methods to generate consumer insights that inform strategy, 2016.

[246] HAMILTON, R., AND PRICE, L. L. Consumer journeys: Developing consumer-based strategy, 2019.

[247] HAMPSHIRE, R. C., AND GAITES, C. Peer-to-peer carsharing: Market analysis and potential growth. *Transportation research record 2217*, 1 (2011), 119–126.

[248] HAN, Y. Cloud computing: case studies and total cost of ownership. *Information technology and libraries 30*, 4 (2011), 198–206.

[249] HANDEL, P., SKOG, I., WAHLSTROM, J., BONAWIEDE, F., WELCH, R., OHLSSON, J., AND OHLSSON, M. Insurance telematics: Opportunities and challenges with the smartphone solution. *IEEE Intelligent Transportation Systems Magazine 6*, 4 (2014), 57–70.

[250] HANRAHAN, B. V., MA, N. F., AND YUAN, C. W. The roots of bias on uber. *arXiv preprint arXiv:1803.08579* (2018).

[251] HARPER, R. H. The role of hci in the age of ai. *International Journal of Human–Computer Interaction 35*, 15 (2019), 1331–1344.

[252] HARTMANN, P. M., ZAKI, M., FELDMANN, N., AND NEELY, A. Big data for big business? a taxonomy of data-driven business models used by start-up firms. *Cambridge Service Alliance* (2014).

[253] HARTMANN, P. M., ZAKI, M., FELDMANN, N., AND NEELY, A. 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 (2016), 1382–1406.

[254] HARTSON, R., AND PYLA, P. S. *The UX Book: Process and guidelines for ensuring a quality user experience*. Elsevier, 2012.

[255] HARVEY, M., LUDWIG, B., AND ELSWEILER, D. You are what you eat: Learning user tastes for rating prediction. In *International symposium on string processing and information retrieval* (2013), Springer, pp. 153–164.

[256] HASSELQVIST, H., HESSELGREN, M., AND BOGDAN, C. Challenging the Car Norm: Opportunities for ICT to Support Sustainable Transportation Practices. In *Proceedings of the 2016 CHI Conference on*

Human Factors in Computing Systems (may 2016), ACM, pp. 1300–1311.

[257] HAVARD, C. J. On the take: The black box of credit scoring and mortgage discrimination. *BU Pub. Int. LJ 20* (2010), 241.

[258] HAWLITSCHEK, F., TEUBNER, T., AND GIMPEL, H. Understanding the sharing economy–drivers and impediments for participation in peer-to-peer rental. In *2016 49th Hawaii International Conference on System Sciences (HICSS)* (2016), IEEE, pp. 4782–4791.

[259] HAWLITSCHEK, F., TEUBNER, T., AND WEINHARDT, C. Trust in the sharing economy. *Die Unternehmung 70*, 1 (2016), 26–44.

[260] HEEG, R. Data storytelling: All's well that ends well. *Research World 2015*, 53 (2015), 10–15.

[261] HELBERGER, N., LYNSKEY, O., MICKLITZ, H.-W., ROTT, P., SAX, M., AND STRYCHARZ, J. Eu consumer protection 2.0. *Structural asymmetries in digital consumer markets. Joint Report from EUCP2. 0 Project BEUC* (2021).

[262] HENDRIKX, F., BUBENDORFER, K., AND CHARD, R. Reputation systems: A survey and taxonomy. *Journal of Parallel and Distributed Computing 75* (2015), 184–197.

[263] HENGSTLER, M., ENKEL, E., AND DUELLI, S. Applied artificial intelligence and trust—the case of autonomous vehicles and medical assistance devices. *Technological Forecasting and Social Change 105* (2016), 105–120.

[264] HENNIG-THURAU, T., WALSH, G., AND WALSH, G. Electronic word-of-mouth: Motives for and consequences of reading customer articulations on the internet. *International journal of electronic commerce 8*, 2 (2003), 51–74.

[265] HERBIG, N., KAHL, G., AND KRÜGER, A. Design guidelines for assistance systems supporting sustainable purchase decisions. In *Pro-*

ceedings of the 2018 Designing Interactive Systems Conference (2018), pp. 1333–1344.

[266] HEVNER, A. R., MARCH, S. T., PARK, J., AND RAM, S. Design science in information systems research. *MIS quarterly* (2004), 75–105.

[267] HEWETT, T. T., BAECKER, R., CARD, S., CAREY, T., GASEN, J., MANTEI, M., PERLMAN, G., STRONG, G., AND VERPLANK, W. *ACM SIGCHI curricula for human-computer interaction*. ACM, 1992.

[268] HEYDE, C. C. On a property of the lognormal distribution. *Journal of the Royal Statistical Society: Series B (Methodological) 25*, 2 (1963), 392–393.

[269] HIDASI, B., KARATZOGLOU, A., BALTRUNAS, L., AND TIKK, D. Session-based recommendations with recurrent neural networks. *arXiv preprint arXiv:1511.06939* (2015).

[270] HILBIG, R., ETSIWAH, B., AND HECHT, S. Berlin start-ups–the rise of data-driven business models. In *ISPIM Innovation Symposium* (2018), The International Society for Professional Innovation Management (ISPIM), pp. 1–19.

[271] HILDEBRAND, K., GEBAUER, M., HINRICHSEN, H., AND MIELKE, M., Eds. *Daten- und Informationsqualität: auf dem Weg zur Information Excellence*, 4., überarbeitete und erweiterte auflage ed. Springer Vieweg, Wiesbaden, 2018. OCLC: 1044698251.

[272] HIND, M. Explaining explainable ai. *XRDS: Crossroads, The ACM Magazine for Students 25*, 3 (2019), 16–19.

[273] HOFMAN-KOHLMEYER, M. Customer loyalty program as a tool of customer retention: literature review. In *CBU International Conference Proceedings* (2016), vol. 4, pp. 199–203.

[274] HOLMES, D., ZOLKIEWSKI, J., AND BURTON, J. The outcomes of b2b data-driven customer focused value creation. *Journal of Business & Industrial Marketing 38*, 6 (2023), 1295–1315.

[275] HOLTZBLATT, K., AND BEYER, H. *Contextual design: defining customer-centered systems*. Elsevier, 1997.

[276] HOLZINGER, A. From machine learning to explainable ai. In *2018 World Symposium on Digital Intelligence for Systems and Machines (DISA)* (2018), IEEE, pp. 55–66.

[277] HONG, J.-H., MARGINES, B., AND DEY, A. K. A smartphone-based sensing platform to model aggressive driving behaviors. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2014), pp. 4047–4056.

[278] HSIAO, J. C.-Y., MOSER, C., SCHOENEBECK, S., AND DILLAHUNT, T. R. The role of demographics, trust, computer self-efficacy, and ease of use in the sharing economy. In *Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies* (2018), pp. 1–11.

[279] HU, Q., XIE, J., AND ZHANG, G. The role of data-driven services strategy in platform competition: A system performance perspective. *Plos one 18*, 1 (2023), e0272547.

[280] HUNKE, F., ENGEL, C., SCHÜRITZ, R., AND EBEL, P. Understanding the anatomy of analytics-based services-a taxonomy to conceptualize the use of data and analytics in services. In *ECIS* (2019).

[281] HUNKE, F., SEEBACHER, S., SCHÜRITZ, R., AND ILLI, A. Towards a process model for data-driven business model innovation. In *2017 IEEE 19th Conference on Business Informatics (CBI)* (2017), vol. 1, IEEE, pp. 150–157.

[282] HUSSIN, H., KING, M., AND CRAGG, P. It alignment in small firms. *European journal of information systems 11*, 2 (2002), 108–127.

[283] HUTAPEA, L., AND MALANOWSKI, N. Neue geschäftsmodelle in der ernährungsindustrie und im lebensmitteleinzelhandel. Tech. rep., Working Paper Forschungsförderung, 2019.

[284] IMRAN, M., YIN, H., CHEN, T., NGUYEN, Q. V. H., ZHOU, A., AND ZHENG, K. Refrs: Resource-efficient federated recommender system for dynamic and diversified user preferences. *ACM Transactions on Information Systems 41*, 3 (2023), 1–30.

[285] ISINKAYE, F. O., FOLAJIMI, Y., AND OJOKOH, B. A. Recommendation systems: Principles, methods and evaluation. *Egyptian informatics journal 16*, 3 (2015), 261–273.

[286] JACKSON, P., ET AL. *The lord of the rings: The fellowship of the ring*. USA New Line, 2001.

[287] JACKSON, P., ET AL. *The lord of the rings: The two towers*. USA New Line, 2002.

[288] JACKSON JR, J. L. *Thin description*. Harvard University Press, 2013.

[289] JAKOBI, T., PATIL, S., RANDALL, D., STEVENS, G., AND WULF, V. It is about what they could do with the data: a user perspective on privacy in smart metering. *ACM Transactions on Computer-Human Interaction (TOCHI) 26*, 1 (2019), 1–44.

[290] JAKOBI, T., STEVENS, G., AND SEUFERT, A.-M. Privacy-by-design für das connected car: Architekturen aus verbrauchersicht: Eine nutzerorientierte diskussion. *Datenschutz und Datensicherheit-DuD 42*, 11 (2018), 704–707.

[291] JAKOBI, T., STEVENS, G., SEUFERT, A.-M., BECKER, M., AND VON GRAFENSTEIN, M. Web tracking under the new data protection law: Design potentials at the intersection of jurisprudence and hci. *i-com 19*, 1 (2020), 31–45.

[292] JANNACH, D., ZANKER, M., FELFERNIG, A., AND FRIEDRICH, G. *Recommender systems: an introduction*. Cambridge University Press, 2010.

[293] JOHNSON, L., AND FUOTI, J. Story juice. *How ideas spread and brands grow. Reach Group, Portland* (2013).

[294] JOHNSON, M. W. *Seizing the white space: Business model innovation for growth and renewal*. Harvard Business Press, 2010.

[295] JOHNSON, S. C. Hierarchical clustering schemes. *Psychometrika 32*, 3 (1967), 241–254.

[296] KAAS, K. P., AND DIETERICH, M. Die entstehung von kaufgewohnheiten bei konsumgütern. *Marketing: Zeitschrift für Forschung und Praxis* (1979), 13–22.

[297] KAISER, C., STOCKER, A., VISCUSI, G., FELLMANN, M., AND RICHTER, A. Conceptualising value creation in data-driven services: The case of vehicle data. *International journal of information management 59* (2021), 102335.

[298] KALLEHAVE, O., SKOV, M. B., AND TIAINEN, N. Persuasion in situ: shopping for healthy food in supermarkets. In *Proceedings of PINC 2011 workshop at CHI* (2011).

[299] KALNIKAITE, V., KREITMAYER, S., ROGERS, Y., BIRD, J., VILLAR, N., BACHOUR, K., PAYNE, S., TODD, P. M., SCHÖNING, J., AND KRÜGER, A. How to nudge in situ: designing lambent devices to deliver salient information in supermarkets. In *Proceedings of the 13th international conference on Ubiquitous computing* (2011), pp. 11–20.

[300] KALTE, D. Political veganism: An empirical analysis of vegans' motives, aims, and political engagement. *Political Studies* (2020), 003232172093017.

[301] KAMAL, P., AND CHEN, J. Q. Trust in sharing economy. pacis 2016 proceedings, 2016.

[302] KARAT, J., AND KARAT, C.-M. The evolution of user-centered focus in the human-computer interaction field. *IBM Systems Journal 42*, 4 (2003), 532–541.

[303] KARPATI, D., NAJJAR, A., AND AMBROSSIO, D. A. Ethics of food recommender applications. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* (2020), pp. 313–319.

[304] KATZEFF, C., MILESTAD, R., ZAPICO, J. L., AND BOHNÉ, U. Encouraging organic food consumption through visualization of personal shopping data. *Sustainability 12*, 9 (apr 2020), 3599.

[305] KAUFMANN, M. Big data management canvas: a reference model for value creation from data. *Big Data and Cognitive Computing 3*, 1 (2019), 19.

[306] KAWGAN-KAGAN, I. Early adopters of carsharing with and without bevs with respect to gender preferences. *European Transport Research Review 7* (2015), 1–11.

[307] KEIL, F. C. Explanation and understanding. *Annu. Rev. Psychol. 57*, 1 (2006), 227–254.

[308] KELLIHER, A., AND SLANEY, M. Tell me a story. *IEEE multimedia 19*, 1 (2012), 4–4.

[309] KEMP, R., AND VAN LENTE, H. The dual challenge of sustainability transitions. *Environmental Innovation and Societal Transitions 1*, 1 (2011), 121–124.

[310] KENENI, B. M., KAUR, D., AL BATAINEH, A., DEVABHAKTUNI, V. K., JAVAID, A. Y., ZAIENTZ, J. D., AND MARINIER, R. P. Evolving rule-based explainable artificial intelligence for unmanned aerial vehicles. *IEEE Access 7* (2019), 17001–17016.

[311] KHUSRO, S., ALI, Z., AND ULLAH, I. Recommender systems: issues, challenges, and research opportunities. In *Information science and applications (ICISA) 2016* (2016), Springer, pp. 1179–1189.

[312] KIHN, M., AND O'HARA, C. B. *Customer data platforms: Use people data to transform the future of marketing engagement*. John Wiley & Sons, 2020.

[313] KING, N. J., AND RAJA, V. Protecting the privacy and security of sensitive customer data in the cloud. *Computer Law & Security Review 28*, 3 (2012), 308–319.

[314] KITCHENHAM, B. A., DYBA, T., AND JORGENSEN, M. Evidence-based software engineering. In *Proceedings. 26th International Conference on Software Engineering* (2004), IEEE, pp. 273–281.

[315] KITZINGER, J. Qualitative research: introducing focus groups. *Bmj 311*, 7000 (1995), 299–302.

[316] KLAUS, L., ET AL. *Data-driven marketing und der Erfolgsfaktor Mensch*. Springer, 2019.

[317] KLEIN, N. J., AND SMART, M. J. Millennials and car ownership: Less money, fewer cars. *Transport Policy 53* (2017), 20–29.

[318] KLOPFER, P. H., AND RUBENSTEIN, D. I. The concept privacy and its biological basis. *Journal of social Issues 33*, 3 (1977), 52–65.

[319] KNAFLIC, C. N. *Storytelling with data: A data visualization guide for business professionals*. John Wiley & Sons, 2015.

[320] KOHLBACHER, F. The use of qualitative content analysis in case study research. In *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research* (2006), vol. 7, Institut für Qualitative Forschung, pp. 1–30.

[321] KOKOLAKIS, S. Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon. *Computers & security 64* (2017), 122–134.

[322] KOLKMAN, D., AND SNEEP, R. Challenges to data science projects with smes: An analysis and decision support tool. *Available at SSRN 3343092* (2019).

[323] KOLLOCK, P., ET AL. The production of trust in online markets. *Advances in group processes 16*, 1 (1999), 99–123.

[324] KOMMERA, A. Future of enterprise integrations and ipaas (integration platform as a service) adoption. *NeuroQuantology 13* (2015), 176–186.

[325] KONSTAN, J. A., AND RIEDL, J. Recommender systems: from algorithms to user experience. *User modeling and user-adapted interaction 22* (2012), 101–123.

[326] KOREN, Y. Collaborative filtering with temporal dynamics. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining* (2009), pp. 447–456.

[327] KOSKELA, L., AND KAGIOGLOU, M. On the metaphysics of production.

[328] KOUTROULI, E., AND TSALGATIDOU, A. Taxonomy of attacks and defense mechanisms in p2p reputation systems—lessons for reputation system designers. *Computer Science Review 6*, 2-3 (2012), 47–70.

[329] KRCMAR, H. *Einführung in das Informationsmanagement*. Springer, 2011.

[330] KRONSBEIN, T., AND MUELLER, R. Data thinking: A canvas for data-driven ideation workshops. In *Proceedings of the 52nd Hawaii International Conference on System Sciences* (2019).

[331] KRUEGER, R. A., AND CASEY, M. A. Focus group interviewing. *Handbook of practical program evaluation* (2015), 506–534.

[332] KUBOTA, K. J., CHEN, J. A., AND LITTLE, M. A. Machine learning for large-scale wearable sensor data in parkinson's disease: Concepts, promises, pitfalls, and futures. *Movement disorders 31*, 9 (2016), 1314–1326.

[333] KÜHNE, B., AND BÖHMANN, T. Requirements for representing data-driven business models-towards extending the business model canvas.

[334] KÜHNE, B., AND BÖHMANN, T. Data-driven business models-building the bridge between data and value. In *ECIS* (2019).

[335] KÜHNE, B., AND BÖHMANN, T. Formative evaluation of data-driven business models–the data insight generator. In *Proceedings of the 53rd Hawaii International Conference on System Sciences* (2020).

[336] KUMAR BOKDE, D., GIRASE, S., AND MUKHOPADHYAY, D. Role of matrix factorization model in collaborative filtering algorithm: A survey. *CoRR, abs/1503.07475* (2015).

[337] KUSHWAHA, N., ROGUSKI, P., AND WATSON, B. W. Up in the air: Ensuring government data sovereignty in the cloud. In *2020 12th International Conference on Cyber Conflict (CyCon)* (2020), vol. 1300, IEEE, pp. 43–61.

[338] KVALE, S., AND BRINKMANN, S. *Interviews: Learning the craft of qualitative research interviewing*. sage, 2009.

[339] KWIATKOWSKA, J., SZÓSTEK, A., AND LAMAS, D. Design and business gaps: From literature to practice. In *Proceedings of the 2014 Mulitmedia, Interaction, Design and Innovation International Conference on Multimedia, Interaction, Design and Innovation* (2014), pp. 1–7.

[340] LAM, S. K., AND RIEDL, J. Shilling recommender systems for fun and profit. In *Proceedings of the 13th international conference on World Wide Web* (2004), pp. 393–402.

[341] LAM, T. K., HEALES, J., HARTLEY, N., HODKINSON, C., ET AL. Consumer trust in food safety requires information transparency. *Australasian Journal of Information Systems 24* (2020).

[342] LATHIA, N., HAILES, S., AND CAPRA, L. Temporal collaborative filtering with adaptive neighbourhoods. In *Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval* (2009), pp. 796–797.

[343] LAUDON, K. C., AND LAUDON, J. P. *Management information systems: Managing the digital firm*. Pearson Educación, 2004.

[344] LAUFER, R. S., AND WOLFE, M. Privacy as a concept and a social issue: A multidimensional developmental theory. *Journal of social Issues 33*, 3 (1977), 22–42.

[345] LAUTERBACH, D., TRUONG, H., SHAH, T., AND ADAMIC, L. Surfing a web of trust: Reputation and reciprocity on couchsurfing. com. In *2009 International Conference on Computational Science and Engineering* (2009), vol. 4, IEEE, pp. 346–353.

[346] LAVALLE, S., LESSER, E., SHOCKLEY, R., HOPKINS, M. S., AND KRUSCHWITZ, N. Big data, analytics and the path from insights to value. *MIT sloan management review* (2010).

[347] LAWO, D., BÖHM, L., AND ESAU, M. Supporting plant-based diets with ingredient2vec. *7th International Conference on ICT for Sustainability* (2020).

[348] LAWO, D., BÖHM, L., FLÜGGE, A.-K., PAKUSCH, C., AND STEVENS, G. Going car-free: investigating mobility practice transformations and the role of ict. In *Proceedings of the 5th International Conference on Computer-Human Interaction Research and Applications (CHIRA 2021), October 28-29, 2021* (2021), SciTePress, Science and Technology Publications, pp. 36–47.

[349] LAWO, D., ENGELBUTZEDER, P., ESAU, M., AND STEVENS, G. Networks of practices: exploring design opportunities for interconnected practices.

[350] LAWO, D., ESAU, M., ENGELBUTZEDER, P., AND STEVENS, G. Going vegan: The role(s) of ict in vegan practice transformation. *Sustainability 12*, 12 (2020), 5184.

[351] LAWO, D., ESAU, M., NEIFER, T., AND STEVENS, G. Probing integrated household information systems for integrated food practices. In *Proceedings of the 5th International Conference on Computer-Human Interaction Research and Applications (CHIRA 2021), October 28-29, 2021* (2021), SciTePress, Science and Technology Publications, pp. 15–23.

[352] LAWO, D., LITZ, K., GROMOV, C., SCHWÄRZER, H., AND STEVENS, G. Going vegan: The use of digital media in vegan diet

transition. *Proceedings of Mensch und Computer 2019* (2019), 503–507.

[353] LAWO, D., NEIFER, T., ESAU, M., ENGELBUTZEDER, P., AND STEVENS, G. Scan&go: Understanding adoption and design of smartphone-based self-checkout. In *Wijnhoven, van Sinderen (Eds.): Proceedings of the 18th International Conference on e-Business, ICE-B 2021, July 7-9, 2021* (2021), SciTePress, Science and Technology Publications, pp. 183–194.

[354] LAWO, D., NEIFER, T., ESAU, M., AND STEVENS, G. Buying the 'right'thing: Designing food recommender systems with critical consumers. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (2021), pp. 1–13.

[355] LAWO, D., NEIFER, T., ESAU, M., AND STEVENS, G. Human-centred digital sovereignty: Explorative conceptual model and ways forward. In *International Conference on Computer-Human Interaction Research and Applications* (2023), Springer, pp. 84–103.

[356] LAWO, D., NEIFER, T., ESAU-HELD, M., AND STEVENS, G. Digital sovereignty: what it is and why it matters for hci. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* (2023), pp. 1–7.

[357] LAWSON, P., MCPHAIL, B., AND LAWTON, E. The connected car: who is in the driver's seat. *British Columbia Freedom of Information and Privacy Association 20* (2015).

[358] LE VINE, S., AND POLAK, J. The impact of free-floating carsharing on car ownership: Early-stage findings from london. *Transport Policy 75* (2019), 119–127.

[359] LEE, B., RICHE, N. H., ISENBERG, P., AND CARPENDALE, S. More than telling a story: Transforming data into visually shared stories. *IEEE computer graphics and applications 35*, 5 (2015), 84–90.

[360] LEE, G., AND TRUONG, N. A reputation and knowledge based trust service platform for trustworthy social internet of things.

[361] LEE, H., AND JAHNG, M. R. The role of storytelling in crisis communication: A test of crisis severity, crisis responsibility, and organizational trust. *Journalism & Mass Communication Quarterly 97*, 4 (2020), 981–1002.

[362] LEE, J., SUN, M., AND LEBANON, G. A comparative study of collaborative filtering algorithms. *arXiv preprint arXiv:1205.3193* (2012).

[363] LEE, T. Q., PARK, Y., AND PARK, Y.-T. A time-based approach to effective recommender systems using implicit feedback. *Expert systems with applications 34*, 4 (2008), 3055–3062.

[364] LEE, Y., MADNICK, S. E., WANG, R. Y., WANG, F., AND ZHANG, H. A cubic framework for the chief data officer: Succeeding in a world of big data.

[365] LEMON, K. N., AND VERHOEF, P. C. Understanding customer experience throughout the customer journey. *Journal of marketing 80*, 6 (2016), 69–96.

[366] LEPRI, B., OLIVER, N., LETOUZÉ, E., PENTLAND, A., AND VINCK, P. Fair, transparent, and accountable algorithmic decision-making processes: The premise, the proposed solutions, and the open challenges. *Philosophy & Technology 31*, 4 (2018), 611–627.

[367] LEVY, Y., AND ELLIS, T. J. A systems approach to conduct an effective literature review in support of information systems research. *Informing Science 9* (2006).

[368] LEWIS, A., AND SIMMONS, M. P2p carsharing service design: informing user experience development, 2012.

[369] LI, H., WU, J., GAO, Y., AND SHI, Y. Examining individuals' adoption of healthcare wearable devices: An empirical study from privacy calculus perspective. *International journal of medical informatics 88* (2016), 8–17.

[370] LI, Q., WANG, Z.-Y., LI, W.-H., LI, J., WANG, C., AND DU, R.-Y. Applications integration in a hybrid cloud computing environment: Modelling and platform. *Enterprise Information Systems 7*, 3 (2013), 237–271.

[371] LI, Y. Theories in online information privacy research: A critical review and an integrated framework. *Decision support systems 54*, 1 (2012), 471–481.

[372] LI, Y., CHEN, H., FU, Z., GE, Y., AND ZHANG, Y. User-oriented fairness in recommendation. In *Proceedings of the web conference 2021* (2021), pp. 624–632.

[373] LIEBERMAN, H., PATERNÒ, F., KLANN, M., AND WULF, V. End-user development: An emerging paradigm. In *Human-Computer Interaction Series*. Springer Netherlands, 2006, pp. 1–8.

[374] LIM, C., KIM, K.-H., KIM, M.-J., HEO, J.-Y., KIM, K.-J., AND MAGLIO, P. P. From data to value: A nine-factor framework for data-based value creation in information-intensive services. *International journal of information management 39* (2018), 121–135.

[375] LINARDATOS, P., PAPASTEFANOPOULOS, V., AND KOTSIANTIS, S. Explainable ai: A review of machine learning interpretability methods. *Entropy 23*, 1 (2020), 18.

[376] LINDEN, G., SMITH, B., AND YORK, J. Amazon. com recommendations: Item-to-item collaborative filtering. *IEEE Internet computing 7*, 1 (2003), 76–80.

[377] LINDLEY, J., COULTON, P., AND STURDEE, M. Implications for adoption. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (2017), pp. 265–277.

[378] LIPHOTO, M., DU, C., AND NGWIRA, S. A survey on recommender systems. In *2016 International Conference on Advances in Computing and Communication Engineering (ICACCE)* (2016), IEEE, pp. 276–280.

[379] LIU, D., BRASS, D. J., LU, Y., AND CHEN, D. Friendships in online peer-to-peer lending. *Mis Quarterly 39*, 3 (2015), 729–742.

[380] LIU, L., AND MUNRO, M. Systematic analysis of centralized online reputation systems. *Decision support systems 52*, 2 (2012), 438–449.

[381] LOEBBECKE, C., AND PICOT, A. Reflections on societal and business model transformation arising from digitization and big data analytics: A research agenda. *The Journal of Strategic Information Systems 24*, 3 (2015), 149–157.

[382] LÖNBERG, N., ANDO MANTOVAARA, L., AND JOHANSSON, P. The death of the third-party cookie: A qualitative study regarding new strategies and methods for collecting consumer data to obtain effective digital marketing, 2022.

[383] LOOSE, W. The state of european car-sharing. *Project Momo Final Report D 2* (2010), 1–119.

[384] LOWDERMILK, T. *User-centered design: a developer's guide to building user-friendly applications*. " O'Reilly Media, Inc.", 2013.

[385] LÜDEKE-FREUND, F., CARROUX, S., JOYCE, A., MASSA, L., AND BREUER, H. The sustainable business model pattern taxonomy—45 patterns to support sustainability-oriented business model innovation. *Sustainable Production and Consumption 15* (2018), 145–162.

[386] LUFTMAN, J., PAPP, R., AND BRIER, T. Enablers and inhibitors of business-it alignment. *Communications of the Association for information Systems 1*, 1 (1999), 11.

[387] LUGER, E., AND SELLEN, A. " like having a really bad pa" the gulf between user expectation and experience of conversational agents. In *Proceedings of the 2016 CHI conference on human factors in computing systems* (2016), pp. 5286–5297.

[388] LUGMAYR, A., SUTINEN, E., SUHONEN, J., SEDANO, C. I., HLAVACS, H., AND MONTERO, C. S. Serious storytelling–a first

definition and review. *Multimedia tools and applications 76* (2017), 15707–15733.

[389] LUHMANN, N. Trust: A mechanism for the reduction of social complexity. *Trust and power: Two works by Niklas Luhmann* (1979), 1–103.

[390] LYNN SHOSTACK, G. How to design a service. *European journal of Marketing 16*, 1 (1982), 49–63.

[391] LYU, D., YANG, F., KWON, H., DONG, W., YILMAZ, L., AND LIU, B. Tdm: trustworthy decision-making via interpretability enhancement. *IEEE Transactions on Emerging Topics in Computational Intelligence 6*, 3 (2021), 450–461.

[392] MA, X., HANCOCK, J. T., LIM MINGJIE, K., AND NAAMAN, M. Self-disclosure and perceived trustworthiness of airbnb host profiles. In *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing* (2017), pp. 2397–2409.

[393] MA, Y.-L., ZHU, X., HU, X., AND CHIU, Y.-C. The use of context-sensitive insurance telematics data in auto insurance rate making. *Transportation Research Part A: Policy and Practice 113* (2018), 243–258.

[394] MACK, M., TISCHLER, R., AND VITSENKO, J. Interactive analytical storytelling. *From the Fireplace to the Boardroom. BARC Research Study. BARC, Würzburg* (2017).

[395] MADUMAL, P., SINGH, R., NEWN, J., AND VETERE, F. Interaction design for explainable ai: workshop proposal. In *Proceedings of the 30th Australian Conference on Computer-Human Interaction* (2018), pp. 607–608.

[396] MAGUIRE, M., AND BEVAN, N. User requirements analysis. In *IFIP World Computer Congress, TC 13* (2002), Springer, pp. 133–148.

[397] MALKIN, N., DEATRICK, J., TONG, A., WIJESEKERA, P., EGELMAN, S., AND WAGNER, D. Privacy attitudes of smart speaker users. *Proceedings on Privacy Enhancing Technologies 2019*, 4 (2019).

[398] MANZINI, E. Il design dei servizi. *La progettazione del 54* (1993).

[399] MANZOCCO, L., ALONGI, M., SILLANI, S., AND NICOLI, M. C. Technological and consumer strategies to tackle food wasting. *Food engineering reviews 8* (2016), 457–467.

[400] MAO, J.-Y., VREDENBURG, K., SMITH, P. W., AND CAREY, T. The state of user-centered design practice. *Communications of the ACM 48*, 3 (2005), 105–109.

[401] MARIAN, M. ipaas: Different ways of thinking. *Procedia Economics and Finance 3* (2012), 1093–1098.

[402] MARRON, D. 'lending by numbers': credit scoring and the constitution of risk within american consumer credit. *Economy and society 36*, 1 (2007), 103–133.

[403] MASSA, L., TUCCI, C. L., AND AFUAH, A. A critical assessment of business model research. *Academy of Management annals 11*, 1 (2017), 73–104.

[404] MATHIS, K., AND KÖBLER, F. Data-need fit–towards data-driven business model innovation. In *Service Design Geographies. Proceedings of the ServDes. 2016 Conference* (2016), no. 125, Linköping University Electronic Press, pp. 458–467.

[405] MAYER, R. C., DAVIS, J. H., AND SCHOORMAN, F. D. An integrative model of organizational trust. *Academy of management review 20*, 3 (1995), 709–734.

[406] MAYRING, P. Einführung in die qualitative sozialforschung. beltz, 2002.

[407] MAYRING, P. Qualitative inhaltsanalyse. handbuch qualitative forschung in der psychologie. *Handbuch Qualitative Forschung in der Psychologie; Mey, G., Mruck, K., Eds* (2010), 601–613.

[408] MAYRING, P., ET AL. Qualitative content analysis. *A companion to qualitative research 1*, 2 (2004), 159–176.

[409] MCAFEE, A., BRYNJOLFSSON, E., DAVENPORT, T. H., PATIL, D., AND BARTON, D. Big data: the management revolution. *Harvard business review 90*, 10 (2012), 60–68.

[410] MCEACHERN, A. What is a repeat customer and why are they profitable?, 2021. https://learn.smile.io/blog/repeat-customers-are-profitable-and-we-can-prove-it.

[411] MCINTYRE, D., SRINIVASAN, A., AFUAH, A., GAWER, A., AND KRETSCHMER, T. Multi-sided platforms as new organizational forms. *Academy of Management Perspectives*, ja (2020).

[412] MCKNIGHT, D. H., AND CHERVANY, N. L. What trust means in e-commerce customer relationships: An interdisciplinary conceptual typology. *International journal of electronic commerce 6*, 2 (2001), 35–59.

[413] MEAD, G. H., ET AL. *Mind, self, and society*, vol. 111. University of Chicago press Chicago, 1934.

[414] MEIERHOFER, J., AND HERRMANN, A. End-to-end methodological approach for the data-driven design of customer-centered digital services. In *Exploring Service Science: 9th International Conference, IESS 2018, Karlsruhe, Germany, September 19–21, 2018, Proceedings 9* (2018), Springer, pp. 208–218.

[415] MERKEL, D., SANTAS, F., HEBERLE, A., AND PLOOM, T. Cloud integration patterns. In *European Conference on Service-Oriented and Cloud Computing* (2015), Springer, pp. 199–213.

[416] MERTLER, C. A., AND REINHART, R. V. *Advanced and multivariate statistical methods: Practical application and interpretation*. Taylor & Francis, 2016.

[417] MERZINGER, M., AND ULBRICH, J. Analyse und optimierung der privatsphäre und des datenschutzes bei telematik-systemen. *informatikJournal 18* (2017), 83–92.

[418] METCALF, D., MILLIARD, S. T., GOMEZ, M., AND SCHWARTZ, M. Wearables and the internet of things for health: Wearable, interconnected devices promise more efficient and comprehensive health care. *IEEE pulse 7*, 5 (2016), 35–39.

[419] MEURER, J., LAWO, D., PAKUSCH, C., TOLMIE, P., AND WULF, V. Opportunities for sustainable mobility. In *Proceedings of the 9th International Conference on Communities & Technologies-Transforming Communities* (2019), pp. 102–113.

[420] MEURER, J., STEIN, M., RANDALL, D., ROHDE, M., AND WULF, V. Social dependency and mobile autonomy: supporting older adults' mobility with ridesharing ict. In *Proceedings of the SIGCHI conference on human factors in computing systems* (2014), pp. 1923–1932.

[421] MEUSER, M., AND NAGEL, U. Das experteninterview—konzeptionelle grundlagen und methodische anlage. *Methoden der vergleichenden Politik-und Sozialwissenschaft: neue Entwicklungen und Anwendungen* (2009), 465–479.

[422] MICHELETTI, M. Political virtue and shopping: Individuals, consumerism and collective action: Pelgrave macmillan.

[423] MIGLAUTSCH, J. R. Thoughts on rfm scoring. *Journal of Database Marketing & Customer Strategy Management 8*, 1 (2000), 67–72.

[424] MIKUSZ, M., JUD, C., AND SCHÄFER, T. Business model patterns for the connected car and the example of data orchestrator. In *Software Business: 6th International Conference, ICSOB 2015, Braga, Portugal, June 10-12, 2015, Proceedings 6* (2015), Springer, pp. 167–173.

[425] MIRANDA, T., CLAYPOOL, M., GOKHALE, A., MIR, T., MURNIKOV, P., NETES, D., AND SARTIN, M. Combining content-based and collaborative filters in an online newspaper. In *In Proceedings of ACM SIGIR Workshop on Recommender Systems* (1999), Citeseer.

[426] MÖHLMANN, M. Collaborative consumption: determinants of satisfaction and the likelihood of using a sharing economy option again. *Journal of Consumer Behaviour 14*, 3 (2015), 193–207.

[427] MOLM, L. D. Power, trust, and fairness: Comparisons of negotiated and reciprocal exchange. In *Power and status*, vol. 20. Emerald Group Publishing Limited, 2003, pp. 31–65.

[428] MORGAN, D. L. *Focus groups as qualitative research*, vol. 16. Sage publications, 1996.

[429] MORGAN, L. Data storytelling: what it is, why it matters. *informationweek. https://www.informationweek.com/big-data/big-data-analytics/data-storytelling-what-it-is-why-it-matters/a/d-id/1325544. Zugegriffen 25* (2016).

[430] MORRIS, M., SCHINDEHUTTE, M., AND ALLEN, J. The entrepreneur's business model: toward a unified perspective. *Journal of business research 58*, 6 (2005), 726–735.

[431] MORRISON, D. G., AND SCHMITTLEIN, D. C. Generalizing the nbd model for customer purchases: What are the implications and is it worth the effort? *Journal of Business & Economic Statistics 6*, 2 (1988), 145–159.

[432] MUERMANN, A., AND STRAKA, D. Asymmetric information in automobile insurance: new evidence from telematic data. Tech. rep., Citeseer, 2011.

[433] MÜLLER, R. M., AND THORING, K. Design thinking vs. lean startup: A comparison of two user-driven innovation strategies. *Leading through design 151* (2012), 91–106.

[434] MÜNZEL, K., BOON, W., FRENKEN, K., AND VASKELAINEN, T. Carsharing business models in germany: characteristics, success and future prospects. *Information Systems and e-Business Management 16*, 2 (2018), 271–291.

[435] MURDOCH, W. J., SINGH, C., KUMBIER, K., ABBASI-ASL, R., AND YU, B. Interpretable machine learning: definitions, methods, and applications. *arXiv preprint arXiv:1901.04592* (2019).

[436] MUSTO, C., TRATTNER, C., STARKE, A., AND SEMERARO, G. Towards a knowledge-aware food recommender system exploiting holistic user models. In *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization* (2020), pp. 333–337.

[437] NAGULENDRA, S., AND VASSILEVA, J. Providing awareness, explanation and control of personalized filtering in a social networking site. *Information Systems Frontiers 18*, 1 (2016), 145–158.

[438] NATHANSON, T., BITTON, E., AND GOLDBERG, K. Eigentaste 5.0: constant-time adaptability in a recommender system using item clustering. In *Proceedings of the 2007 ACM conference on Recommender systems* (2007), pp. 149–152.

[439] NATROP, J. *Grundzüge der Angewandten Mikroökonomie 2. A.* Oldenbourg Wissenschaftsverlag, 2012.

[440] NEIFER, T., BOSSAUER, P., LAWO, D., VOLKENING, R., AND GADATSCH, A. The role of marketplaces for the transformation from robotic process automation to intelligent process automation. In *ICSBT* (2022), pp. 15–25.

[441] NEIFER, T., BOSSAUER, P., PAKUSCH, C., BOEHM, L., AND LAWO, D. Trust-building in peer-to-peer carsharing: Design case study for algorithm-based reputation systems. *Computer Supported Cooperative Work (CSCW)* (2023), 1–35.

[442] NEIFER, T., SCHMIDT, A., BOSSAUER, P., AND GADATSCH, A. Data science management: Planung, steuerung und kontrolle von data science. *Rethinking Finance*, 03 (2019), 15–18.

[443] NEIFER, T., SCHMIDT, A., BOSSAUER, P., AND GADATSCH, A. Data science canvas: Ein instrument zur operationalisierung von daten. *Big Data: Anwendung und Nutzungspotenziale in der Produktion* (2020), 37.

[444] NEWKIRK, H. E., LEDERER, A. L., AND JOHNSON, A. M. Rapid business and it change: drivers for strategic information systems planning? *European Journal of Information Systems 17*, 3 (2008), 198–218.

[445] NG, K. H., SHIPP, V., MORTIER, R., BENFORD, S., FLINTHAM, M., AND RODDEN, T. Understanding food consumption lifecycles using wearable cameras. *Personal and Ubiquitous Computing 19* (2015), 1183–1195.

[446] NICOLA, S., FERREIRA, E. P., AND FERREIRA, J. P. A novel framework for modeling value for the customer, an essay on negotiation. *International Journal of Information Technology & Decision Making 11*, 03 (2012), 661–703.

[447] NICOLINI, D. Zooming in and zooming out: A package of method and theory to study work practices. *Organizational ethnography: Studying the complexities of everyday life* (2009), 120–138.

[448] NIELSEN, J., BØDKER, M., AND VATRAPU, R. Culture and (i) literacy as challenges to scandinavian cooperative design. In *Proceedings of the 3rd international conference on Intercultural collaboration* (2010), pp. 271–274.

[449] NIELSEN, L. Personas in co-creation and co-design. In *Proceedings of the 11th Human-Computer Interaction Research Symposium* (2011), pp. 38–40.

[450] NOBIS, C. Carsharing as key contribution to multimodal and sustainable mobility behavior: Carsharing in germany. *Transportation Research Record 1986*, 1 (2006), 89–97.

[451] NORBERG, P. A., HORNE, D. R., AND HORNE, D. A. The privacy paradox: Personal information disclosure intentions versus behaviors. *Journal of consumer affairs 41*, 1 (2007), 100–126.

[452] NORMAN, D. *The design of everyday things: Revised and expanded edition*. Basic books, 2013.

[453] NORMAN, D. A., AND DRAPER, S. W. *User centered system design; new perspectives on human-computer interaction*. L. Erlbaum Associates Inc., 1986.

[454] NORMANN, R., AND RAMIREZ, R. Designing interactive strategy: From value chain to value constellation, chichester-new york, wiley.

[455] OAKIL, A. T. M., MANTING, D., AND NIJLAND, H. Determinants of car ownership among young households in the netherlands: The role of urbanisation and demographic and economic characteristics. *Journal of transport geography 51* (2016), 229–235.

[456] OJO, A., AND HERAVI, B. Patterns in award winning data storytelling: Story types, enabling tools and competences. *Digital journalism 6*, 6 (2018), 693–718.

[457] OLAKANMI, O., AND OLUWASEUN, S. A trust based secure and privacy aware framework for efficient taxi and car sharing system. *International Journal of Vehicular Telematics and Infotainment Systems (IJVTIS) 2*, 1 (2018), 34–47.

[458] OLATEJU, O., OKON, S. U., OLANIYI, O. O., SAMUEL-OKON, A. D., AND ASONZE, C. U. Exploring the concept of explainable ai and developing information governance standards for enhancing trust and transparency in handling customer data. *Available at SSRN* (2024).

[459] O'ROURKE, D., AND RINGER, A. The impact of sustainability information on consumer decision making. *Journal of Industrial Ecology 20*, 4 (aug 2016), 882–892.

[460] OSTERWALDER, A., AND PIGNEUR, Y. Modeling value propositions in e-business. In *Proceedings of the 5th international conference on Electronic commerce* (2003), pp. 429–436.

[461] OSTERWALDER, A., AND PIGNEUR, Y. Towards business and information systems fit through a business model ontology. In *Proceedings of the Strategic Management Society Conference* (2003), Citeseer, pp. 1–8.

[462] OSTERWALDER, A., AND PIGNEUR, Y. *Business model generation: a handbook for visionaries, game changers, and challengers*, vol. 1. John Wiley & Sons, 2010.

[463] OSTERWALDER, A., AND PIGNEUR, Y. Aligning profit and purpose through business model innovation. *Responsible management practices for the 21st century* (2011), 61–76.

[464] OSTERWALDER, A., PIGNEUR, Y., BERNARDA, G., AND SMITH, A. *Value proposition design: How to create products and services customers want*. John Wiley & Sons, 2014.

[465] OULASVIRTA, A., AND HORNBÆK, K. Hci research as problem-solving. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (2016), pp. 4956–4967.

[466] OWYANG, J., TRAN, C., AND SILVA, C. The collaborative economy. *Altimeter, United States* (2013).

[467] OZDEMIR, S. *Principles of data science*. Packt Publishing Ltd, 2016.

[468] ÖZEKICI, S., AND SOYER, R. Bayesian analysis of markov modulated bernoulli processes. *Mathematical methods of operations research 57*, 1 (2003), 125–140.

[469] PAKUSCH, C., NEIFER, T., BOSSAUER, P., AND STEVENS, G. P2p-carsharing. motive, ängste und barrieren bei der teilnahme-eine explorative studie. *Internationales Verkehrswesen 70*, 4 (2018).

[470] PALEN, L., AND DOURISH, P. Unpacking" privacy" for a networked world. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (2003), pp. 129–136.

[471] PALLANT, J. I., PALLANT, J. L., SANDS, S. J., FERRARO, C. R., AND AFIFI, E. When and how consumers are willing to exchange data with retailers: An exploratory segmentation. *Journal of Retailing and Consumer Services 64* (2022), 102774.

[472] PANETTA, K. Analyst Answers: The Biggest Challenges for Data & Analytics Leaders Today.

[473] PAREKH, J., MOZHAROVSKYI, P., AND D'ALCHÉ BUC, F. A framework to learn with interpretation. *Advances in Neural Information Processing Systems 34* (2021), 24273–24285.

[474] PARK, Y., AND LEE, T.-Q. Using temporal information in collaborative filtering: An empirical study. In *CSREA EEE* (2006), Citeseer, p. 316.

[475] PARKER, A. G., AND GRINTER, R. E. Collectivistic health promotion tools: Accounting for the relationship between culture, food and nutrition. *International Journal of Human-Computer Studies 72*, 2 (2014), 185–206.

[476] PAUKSTADT, U., STROBEL, G., AND EICKER, S. Understanding services in the era of the internet of things: a smart service taxonomy.

[477] PAZZANI, M. J. A framework for collaborative, content-based and demographic filtering. *Artificial intelligence review 13*, 5-6 (1999), 393–408.

[478] PEFFERS, K., ROTHENBERGER, M., TUUNANEN, T., AND VAEZI, R. Design science research evaluation. In *International Conference*

on Design Science Research in Information Systems (2012), Springer, pp. 398–410.

[479] PEFFERS, K., TUUNANEN, T., ROTHENBERGER, M. A., AND CHATTERJEE, S. A design science research methodology for information systems research. *Journal of management information systems 24*, 3 (2007), 45–77.

[480] PEPPARD, J., AND WARD, J. *The strategic management of information systems: Building a digital strategy*. John Wiley & Sons, 2016.

[481] PEREZ, C. E. *The deep learning AI playbook: Strategy for disruptive artificial intelligence*. CreateSpace, 2017.

[482] PEZZINI, M., AND LHEUREUX, B. Integration platform as a service: moving integration to the cloud. *Gartner RAS Core Research Note G 210747* (2011), 7.

[483] PFADENHAUER, M. Das experteninterview: Ein gespräch auf gleicher augenhöhe. *Qualitative Marktforschung: Konzepte–Methoden–Analysen* (2009), 449–461.

[484] PFITZMANN, A. Multilateral security: Enabling technologies and their evaluation. In *Informatics: 10 Years Back, 10 Years Ahead*. Springer, 2001, pp. 50–62.

[485] PFITZMANN, A., SCHILL, A., WESTFELD, A., AND WOLF, G. Mehrseitige sicherheit in offenen netzen. *Grundlagen, praktische Umsetzung und in Java implementierte Demonstrations-Software* (2000).

[486] PIATETSKY, G. Python eats away at r: Top software for analytics, data science, machine learning in 2018: Trends and analysis.

[487] PIGNI, F., PICCOLI, G., AND WATSON, R. Digital data streams: Creating value from the real-time flow of big data. *California Management Review 58*, 3 (2016), 5–25.

[488] PINS, D., JAKOBI, T., BODEN, A., ALIZADEH, F., AND WULF, V. Alexa, we need to talk: A data literacy approach on voice assistants. In

Proceedings of the 2021 ACM Designing Interactive Systems Conference (2021), pp. 495–507.

[489] PITTS, S. B. J., NG, S. W., BLITSTEIN, J. L., GUSTAFSON, A., AND NICULESCU, M. Online grocery shopping: promise and pitfalls for healthier food and beverage purchases. *Public health nutrition 21*, 18 (2018), 3360–3376.

[490] POGGI, N., MUTHUSAMY, V., CARRERA, D., AND KHALAF, R. Business process mining from e-commerce web logs. In *Business process management*. Springer, 2013, pp. 65–80.

[491] PORTAL, E. G. Key changes with the general data protection regulation. *EU GDPR Portal* (2017).

[492] PORTER, M. E., HEPPELMANN, J. E., ET AL. How smart, connected products are transforming competition. *Harvard business review 92*, 11 (2014), 64–88.

[493] POTOČNIK, M., AND JURIC, M. B. Integration of saas using ipaas. In *The 1st international conference on CLoud Assisted ServiceS* (2012), p. 35.

[494] PRICEWATERHOUSECOOPERS. Share economy in deutschland wächst weiter.

[495] PROST, S., CRIVELLARO, C., HADDON, A., AND COMBER, R. Food Democracy in the Making: Designing with Local Food Networks. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (apr 2018), CHI '18, ACM, pp. 333:1–333:14.

[496] PROST, S., VLACHOKYRIAKOS, V., MIDGLEY, J., HERON, G., MEZIANT, K., AND CRIVELLARO, C. Infrastructuring Food Democracy: The Formation of a Local Food Hub in the Context of Socio-Economic Deprivation. *Proceedings of the ACM on Human-Computer Interaction 3*, CSCW (nov 2019), 1–27.

[497] PROVOST, F., AND FAWCETT, T. Data science and its relationship to big data and data-driven decision making. *Big data 1*, 1 (2013), 51–59.

[498] PYCZAK, T. Data storytelling: Wie geschichten gute entscheidungen ermöglichen. *Strategisches Storytelling. https://www.strategisches-storytelling.de/data-storytelling/ 25* (2017).

[499] QIU, J., WU, Q., DING, G., XU, Y., AND FENG, S. A survey of machine learning for big data processing. *EURASIP Journal on Advances in Signal Processing 2016* (2016), 1–16.

[500] QIU, W., PARIGI, P., AND ABRAHAO, B. More stars or more reviews? In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (2018), pp. 1–11.

[501] QLIK. How to Drive Data Literacy With the Enterprise.

[502] RABIEE, F. Focus-group interview and data analysis. *Proceedings of the nutrition society 63*, 4 (2004), 655–660.

[503] RACHINGER, M., RAUTER, R., MÜLLER, C., VORRABER, W., AND SCHIRGI, E. Digitalization and its influence on business model innovation. *Journal of manufacturing technology management 30*, 8 (2018), 1143–1160.

[504] RADER, E., COTTER, K., AND CHO, J. Explanations as mechanisms for supporting algorithmic transparency. In *Proceedings of the 2018 CHI conference on human factors in computing systems* (2018), pp. 1–13.

[505] RAGHUNATH, V., KUNKULAGUNTA, M., AND NADELLA, G. S. Ai-driven business analytics framework for data integration across hybrid cloud systems. *Transactions on Latest Trends in Artificial Intelligence 4*, 4 (2023).

[506] RAINER, R. K., AND PRINCE, B. *Introduction to information systems*. John Wiley & Sons, 2021.

[507] RANDALL, D. Consumer strategies for the internet: four scenarios. *Long Range Planning 30*, 2 (1997), 157–147.

[508] REICH, B. H., AND BENBASAT, I. Factors that influence the social dimension of alignment between business and information technology objectives. *MIS quarterly* (2000), 81–113.

[509] REITBERGER, W. H., SPREICER, W., AND FITZPATRICK, G. Nutriflect: reflecting collective shopping behavior and nutrition. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2014), pp. 3309–3318.

[510] REKSOATMODJO, W., HARTONO, J., DJUNAEDI, A., AND UTOMO, H. Exploratory study on alignment between it and business strategies. *Gadjah Mada International Journal of Business 14*, 2 (2012), 139–162.

[511] REPSCHLÄGER, J., ZARNEKOW, R., MEINHARDT, N., RÖDER, C., AND PRÖHL, T. *Vertrauen in der Share Economy: Studie: Analyse von Vertrauensfaktoren für Online-Profile*, vol. 15. Universitätsverlag der TU Berlin, 2015.

[512] RESNICK, P., KUWABARA, K., ZECKHAUSER, R., AND FRIEDMAN, E. Reputation systems. *Communications of the ACM 43*, 12 (2000), 45–48.

[513] RESNICK, P., AND ZECKHAUSER, R. Trust among strangers in internet transactions: Empirical analysis of ebay's reputation system. In *The Economics of the Internet and E-commerce*. Emerald Group Publishing Limited, 2002, pp. 127–157.

[514] RIBEIRO, M. T., SINGH, S., AND GUESTRIN, C. " why should i trust you?" explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining* (2016), pp. 1135–1144.

[515] RICCI, F., ROKACH, L., AND SHAPIRA, B. Introduction to recommender systems handbook. In *Recommender systems handbook*. Springer, 2010, pp. 1–35.

[516] RICHTER, C., KRAUS, S., AND SYRJÄ, P. The shareconomy as a precursor for digital entrepreneurship business models. *International Journal of Entrepreneurship and Small Business 25*, 1 (2015), 18–35.

[517] RIDINGS, C. M., GEFEN, D., AND ARINZE, B. Some antecedents and effects of trust in virtual communities. *The journal of strategic information systems 11*, 3-4 (2002), 271–295.

[518] RIDLEY, D. The literature review: A step-by-step guide for students.

[519] RIES, E. *The lean startup: How today's entrepreneurs use continuous innovation to create radically successful businesses*. Currency, 2011.

[520] RITTER, E. F., BAXTER, D. G., AND CHURCHILL, F. E. *Foundations for designing user-centered systems: What system designers need to know about people*. Springer, 2014.

[521] ROEL, V., ANTONIO, K., AND CLAESKENS, G. Unraveling the predictive power of telematics data in car insurance pricing. *Available at SSRN 2872112* (2017).

[522] ROGERS, B. The social costs of uber. *U. Chi. L. Rev. Dialogue 82* (2015), 85.

[523] ROGERS, Y., SHARP, H., AND PREECE, J. *Interaction design: Beyond human-computer interaction*, 6 ed. John Wiley & Sons, Nashville, TN, 2023.

[524] ROTELLA, P. Is data the new oil? *Forbes Magazine* (2012).

[525] ROTT, P. Data protection law as consumer law–how consumer organisations can contribute to the enforcement of data protection law. *Journal of European Consumer and Market Law 6*, 3 (2017).

[526] ROTTER, J. B. A new scale for the measurement of interpersonal trust. *Journal of personality* (1967).

[527] ROUVROY, A. The end (s) of critique: Data behaviourism versus due process. In *Privacy, due process and the computational turn*. Routledge, 2013, pp. 157–182.

[528] ROWLEY, J., AND SLACK, F. Conducting a literature review. *Management research news 27*, 6 (2004), 31–39.

[529] ROZONY, F. Z., AKTAR, M., ASHRAFUZZAMAN, M., AND ISLAM, A. A systematic review of big data integration challenges and solutions for heterogeneous data sources. *Academic Journal on Business Administration, Innovation & Sustainability 4*, 04 (2024), 1–18.

[530] RUBY, M. B., HEINE, S. J., KAMBLE, S., CHENG, T. K., AND WADDAR, M. Compassion and contamination. cultural differences in vegetarianism. *Appetite 71* (2013), 340–348.

[531] RUOHOMAA, S., KUTVONEN, L., AND KOUTROULI, E. Reputation management survey. In *The Second International Conference on Availability, Reliability and Security (ARES'07)* (2007), IEEE, pp. 103–111.

[532] RUSSOM, P. Managing big data. *TDWI Best Practices Report, TDWI Research* (2013), 1–40.

[533] RUSSOM, P., ET AL. Big data analytics. *TDWI best practices report, fourth quarter 19*, 4 (2011), 1–34.

[534] RUST, R. T., AND HUANG, M.-H. The service revolution and the transformation of marketing science. *Marketing Science 33*, 2 (2014), 206–221.

[535] RYAN, L. *The visual imperative: Creating a visual culture of data discovery*. Morgan kaufmann, 2016.

[536] RYAN, S. D., HARRISON, D. A., AND SCHKADE, L. L. Information-technology investment decisions: when do costs and benefits in the social subsystem matter? *Journal of Management Information Systems 19*, 2 (2002), 85–127.

[537] SAAD-SULONEN, J., DE GÖTZEN, A., MORELLI, N., AND SIMEONE, L. Service design and participatory design: time to join forces? In *Proceedings of the 16th Participatory Design Conference 2020-Participation (s) Otherwise-Volume 2* (2020), pp. 76–81.

[538] SAARIJÄRVI, H., MITRONEN, L., AND YRJÖLÄ, M. From selling to supporting–leveraging mobile services in the context of food retailing. *Journal of Retailing and Consumer Services 21*, 1 (2014), 26–36.

[539] SABATER, J., AND SIERRA, C. Review on computational trust and reputation models. *Artificial intelligence review 24* (2005), 33–60.

[540] SABHERWAL, R., AND CHAN, Y. E. Alignment between business and is strategies: A study of prospectors, analyzers, and defenders. *Information systems research 12*, 1 (2001), 11–33.

[541] SALAM, A. F., IYER, L., PALVIA, P., AND SINGH, R. Trust in e-commerce. *Communications of the ACM 48*, 2 (2005), 72–77.

[542] SALIPANTE, P., NOTZ, W., AND BIGELOW, J. A matrix approach to literature reviews. *Research in organizational behavior 4* (1982), 321–348.

[543] SALOVAARA, A., OULASVIRTA, A., AND JACUCCI, G. Evaluation of prototypes and the problem of possible futures. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (2017), pp. 2064–2077.

[544] SAMUEL, A. Data is the next big thing in content marketing. *Harvard Business Review 9* (2015).

[545] SANGIORGI, D., PACENTI, E., ET AL. Service design research pioneers. an overview of service design research developed in italy since the '90s. *Design Research Journal 1*, 10 (2010), 26–33.

[546] SANKARAN, K. Data science principles for interpretable and explainable ai. *arXiv preprint arXiv:2405.10552* (2024).

[547] SARAIVA, A., FERNANDES, E., AND VON SCHWEDLER, M. The green identity formation process in organic consumer communities. *Qualitative Market Research: An International Journal 23* (2020), 69–86.

[548] SARKER, S., CHATTERJEE, S., AND XIAO, X. How" sociotechnical" is our is research?: An assessment and possible ways forward. In *Proceedings of the 34th International Conference on Information Systems. ICIS 2013* (2013), Association for Information Systems. AIS Electronic Library (AISeL), p. 1185.

[549] SARWAR, B., KARYPIS, G., KONSTAN, J., AND RIEDL, J. Item-based collaborative filtering recommendation algorithms. In *Proceedings of the 10th international conference on World Wide Web* (2001), pp. 285–295.

[550] SARWAR, B., KARYPIS, G., KONSTAN, J., AND RIEDL, J. Incremental singular value decomposition algorithms for highly scalable recommender systems. In *Fifth international conference on computer and information science* (2002), vol. 1, Citeseer, pp. 27–8.

[551] SATHANANTHAN, S., HOETKER, P., GAMRAD, D., KATTERBACH, D., AND MYRZIK, J. Realizing digital transformation through a digital business model design process. In *2017 Internet of Things Business Models, Users, and Networks* (2017), IEEE, pp. 1–8.

[552] SAVAGE, A. Dear Companies: Your Data Science Job Descriptions are Awful. . . , Apr. 2019.

[553] SAVOLAINEN, K. User-centred design without involving users: A longitudinal case study in a human-centred-design–mature company. *The Design Journal 24*, 6 (2021), 887–905.

[554] SCHAFER, J. B., KONSTAN, J., AND RIEDL, J. Recommender systems in e-commerce. In *Proceedings of the 1st ACM conference on Electronic commerce* (1999), pp. 158–166.

[555] SCHMIDT, M., EARNSHAW, Y., TAWFIK, A. A., AND JAHNKE, I. Methods of user centered design and evaluation for learning designers. *Learner and User Experience Research* (2020).

[556] SCHMIT, S., AND RIQUELME, C. Human interaction with recommendation systems. In *International Conference on Artificial Intelligence and Statistics* (2018), pp. 862–870.

[557] SCHREIER, H., BECKER, I. U. J., AND HELLER, J. Endbericht evaluation carsharing (eva-cs).

[558] SCHUBERT, C. Exploring the (behavioural) political economy of nudging. *Journal of Institutional Economics 13*, 3 (2017), 499–522.

[559] SCHUBERT, C. Green nudges: Do they work? are they ethical? *Ecological economics 132* (2017), 329–342.

[560] SCHÜRITZ, R., FARRELL, K., WIXOM, B. W., AND SATZGER, G. Value co-creation in data-driven services: towards a deeper understanding of the joint sphere.

[561] SCHÜRITZ, R., AND SATZGER, G. Patterns of data-infused business model innovation. In *2016 IEEE 18th Conference on Business Informatics (CBI)* (2016), vol. 1, IEEE, pp. 133–142.

[562] SCHÜRITZ, R., SEEBACHER, S., AND DORNER, R. Capturing value from data: Revenue models for data-driven services. In *Proceedings of the 50th Hawaii International Conference on System Sciences* (2017).

[563] SCHYMANIETZ, M., JONAS, J. M., AND MÖSLEIN, K. M. Exploring data-driven service innovation—aligning perspectives in research and practice. *Journal of Business Economics 92*, 7 (2022), 1167–1205.

[564] SEGEL, E., AND HEER, J. Narrative visualization: Telling stories with data. *IEEE transactions on visualization and computer graphics 16*, 6 (2010), 1139–1148.

[565] SEIDELIN, C. *Towards a co-design perspective on data: Foregrounding data in the design and innovation of data-based services*. IT-Universitetet i København, 2020.

[566] SEIN, M. K., HENFRIDSSON, O., PURAO, S., ROSSI, M., AND LINDGREN, R. Action design research. *MIS quarterly* (2011), 37–56.

[567] SEN, S., AND LERMAN, D. Why are you telling me this? an examination into negative consumer reviews on the web. *Journal of interactive marketing 21*, 4 (2007), 76–94.

[568] SEN, S., AND SEN, S. ipaas–cloud solution for the cloud problem.

[569] SETNES, M., BABUSKA, R., AND VERBRUGGEN, H. B. Rule-based modeling: Precision and transparency. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) 28*, 1 (1998), 165–169.

[570] SHAHEEN, S., AND COHEN, A. Innovative mobility: Carsharing outlook carsharing market overview, analysis, and trends. Tech. rep., Institute of Transportation Studies, UC Berkeley, 2020.

[571] SHAHEEN, S. A., AND COHEN, A. P. Carsharing and personal vehicle services: worldwide market developments and emerging trends. *International journal of sustainable transportation 7*, 1 (2013), 5–34.

[572] SHAW, S. D., AND BAGOZZI, R. P. The neuropsychology of consumer behavior and marketing. *Consumer Psychology Review 1*, 1 (2018), 22–40.

[573] SHEERAN, P., AND WEBB, T. L. The intention-behavior gap. *Social and personality psychology compass 10*, 9 (2016), 503–518.

[574] SHIN, D. The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable ai. *International Journal of Human-Computer Studies 146* (2021), 102551.

[575] SHMUELI, G. To explain or to predict?

[576] SHOSTACK, G. Designing services that deliver. *Harvard Business Review* (1984).

[577] SHOVE, E., WATSON, M., AND PANTZAR, M. The dynamics of social practice: Everyday life and how it changes.

[578] SICHEL, H. Repeat-buying and the generalized inverse gaussian–poisson distribution. *Journal of the Royal Statistical Society: Series C (Applied Statistics) 31*, 3 (1982), 193–204.

[579] SILVER, M. Repeat-buying: Facts, theory and applications. *Journal of the Royal Statistical Society: Series A (Statistics in Society) 152*, 3 (1989), 420–421.

[580] SILVIUS, A. G. Business & it alignment in theory and practice. In *2007 40th Annual Hawaii International Conference on System Sciences (HICSS'07)* (2007), IEEE, pp. 211b–211b.

[581] SIMONSEN, J., AND ROBERTSON, T. *Routledge international handbook of participatory design*, vol. 711. Routledge New York, 2013.

[582] ŞIMŞEKOĞLU, Ö., NORDFJÆRN, T., AND RUNDMO, T. The role of attitudes, transport priorities, and car use habit for travel mode use and intentions to use public transportation in an urban norwegian public. *Transport policy 42* (2015), 113–120.

[583] SKAAR, J., BØLVIKEN, T., KOSKELA, L., AND KALSAAS, B. T. Principles as a bridge between theory and practice. In *28th Annual Conference of the International Group for Lean Construction* (2020), The International Group for Lean Construction.

[584] SKUBIC, M., GUEVARA, R. D., AND RANTZ, M. Automated health alerts using in-home sensor data for embedded health assessment. *IEEE journal of translational engineering in health and medicine 3* (2015), 1–11.

[585] SMEENK, W. The empathic co-design canvas: A tool for supporting multi-stakeholder co-design processes. *International Journal of Design 17*, 2 (2023), 81–98.

[586] SMITH, B., AND LINDEN, G. Two decades of recommender systems at amazon.com. *Ieee internet computing 21*, 3 (2017), 12–18.

[587] SMITH, H. J., DINEV, T., AND XU, H. Information privacy research: an interdisciplinary review. *MIS quarterly* (2011), 989–1015.

[588] SMITH, R. C., BOSSEN, C., AND KANSTRUP, A. M. Participatory design in an era of participation, 2017.

[589] SMITHSON, J. Using and analysing focus groups: limitations and possibilities. *International journal of social research methodology 3*, 2 (2000), 103–119.

[590] SOKOL, K., AND FLACH, P. One explanation does not fit all: The promise of interactive explanations for machine learning transparency. *KI-Künstliche Intelligenz 34*, 2 (2020), 235–250.

[591] SOLOMON, M. R. *Consumer behavior: Buying, having, and being*. Pearson, 2020.

[592] SOLOVE, D. J. Understanding privacy.

[593] SORESCU, A. Data-driven business model innovation. *Journal of Product Innovation Management 34*, 5 (2017), 691–696.

[594] SPADA, E. *Data science for connected car insurance: use of trips raw telematics data for knowledge discovery and customers profiling*. PhD thesis, Universidade Nova de Lisboa, 2018.

[595] SPREMANN, K. Agent and principal, in 'agency theory, information, and incentives'. *Berlin-Heidelberg-New York: Spring Valley* (1987).

[596] SQUICCIARINI, A., BERTINO, E., FERRARI, E., PACI, F., AND THURAISINGHAM, B. Pp-trust-x: A system for privacy preserving trust negotiations. *ACM Transactions on Information and System Security (TISSEC) 10*, 3 (2007), 12–es.

[597] STANDING, C., STANDING, S., AND BIERMANN, S. The implications of the sharing economy for transport. *Transport Reviews 39*, 2 (2019), 226–242.

[598] STAR, S. L. The structure of ill-structured solutions: Boundary objects and heterogeneous distributed problem solving. In *Distributed artificial intelligence*. Elsevier, 1989, pp. 37–54.

[599] STARKE, A. Recsys challenges in achieving sustainable eating habits. In *HealthRecSys* (2019), pp. 29–30.

[600] STEVENS, G., BODEN, A., WINTERBERG, L., GÓMEZ, J. M., AND BALA, C. Digitaler konsum: Herausforderungen und chancen der verbraucherinformatik.

[601] STEVENS, G., AND BOSSAUER, P. Dealing with personal data in the age of big data economies. *Zeitschrift fuer Geistiges Eigentum-Intellectual Property Journal 9*, 3 (2017), 266–278.

[602] STEVENS, G., AND BOSSAUER, P. Who do you trust: Peers or technology? a conjoint analysis about computational reputation mechanisms. In *Proceedings of 18th European Conference on Computer-Supported Cooperative Work* (2020), European Society for Socially Embedded Technologies (EUSSET).

[603] STEVENS, G., BOSSAUER, P., JAKOBI, T., AND PAKUSCH, C. Second dashboard: Information demands in a connected car. *Mensch und Computer 2017-Tagungsband* (2017).

[604] STEVENS, G., BOSSAUER, P., JAKOBI, T., AND PAKUSCH, C. Mehrseitiges vertrauen bei iot-basierten reputationssystemen. *Mensch und Computer 2018-Workshopband* (2018).

[605] STEVENS, G., BOSSAUER, P., NEIFER, T., AND HANSCHKE, S. Using shopping data to design sustainable consumer apps. In *2017 Sustainable Internet and ICT for Sustainability (SustainIT)* (2017), IEEE, pp. 1–3.

[606] STEVENS, G., JAKOBI, T., AND DETKEN, K.-O. Mehrseitige, barrierefreie sicherheit intelligenter messsysteme. *Datenschutz und Datensicherheit-DuD 38*, 8 (2014), 536–544.

[607] STEVENS, G., AND PIPEK, V. Making use: understanding, studying, and supporting appropriation. In *Wulf, Pipek et al.(Eds.): Socio-Informatics: A Practice-Based Perspective on the Design and Use of IT Artifacts*. Oxford University Press, 2018, pp. 139–176.

[608] STEVENS, G., AND WULF, V. A new dimension in access control: Studying maintenance engineering across organizational boundaries. In *Proceedings of the 2002 ACM conference on Computer supported cooperative work* (2002), pp. 196–205.

[609] STEVENS, G., AND WULF, V. Computer-supported access control. *ACM Transactions on Computer-Human Interaction (TOCHI) 16*, 3 (2009), 1–26.

[610] STEWART, D. W., AND ZHAO, Q. Internet marketing, business models, and public policy. *Journal of public policy & marketing 19*, 2 (2000), 287–296.

[611] STICKDORN, M., AND SCHNEIDER, J. *This is service design thinking: Basics, tools, cases*. John Wiley & Sons, 2012.

[612] STOLLE, D., AND HOOGHE, M. Consumers as political participants? shifts in political action repertoires in western societies. *Politics, products, and markets: Exploring political consumerism past and present* (2004), 265–288.

[613] STOLLE, D., AND MICHELETTI, M. *Political Consumerism*. Cambridge University Press, 2013.

[614] STONE, E. Podcast: The power of persuasive storytelling. in: Kellogg insight. `https://insight.kellogg.northwestern.edu/article/the-power-of-persuasive-storytelling`, 2015. (Accessed on 19/05/2023).

[615] STREECK, H., HARTMANN, G., EXNER, M., AND SCHMID, M. Vorläufiges ergebnis und schlussfolgerungen der covid-19 case-cluster-study (gemeinde gangelt). *Preprint published online on* (2020), 04–09.

[616] STROBEL, M. An axiomatic approach to explain computer generated decisions. In *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society* (2018), pp. 380–381.

[617] SUTCLIFFE, A. G., AND MINOCHA, S. Linking business modelling to socio-technical system design. In *Advanced Information Systems Engineering: 11th International Conference, CAiSE" 99 Heidelberg, Germany, June 14—18, 1999 Proceedings 11* (1999), Springer, pp. 73–87.

[618] SWAN, M. Connected car: quantified self becomes quantified car. *Journal of Sensor and Actuator Networks 4*, 1 (2015), 2–29.

[619] TALASU, N., JONNALAGADDA, A., PILLAI, S. S. A., AND RAHUL, J. A link prediction based approach for recommendation systems. In *2017 international conference on advances in computing, communications and informatics (ICACCI)* (2017), IEEE, pp. 2059–2062.

[620] TAN, F. B., AND GALLUPE, R. B. Aligning business and information systems thinking: A cognitive approach. *IEEE transactions on Engineering Management 53*, 2 (2006), 223–237.

[621] TANG, T. Y., WINOTO, P., AND CHAN, K. C. Scaling down candidate sets based on the temporal feature of items for improved hybrid recommendations. In *IJCAI Workshop on Intelligent Techniques for Web Personalization* (2003), Springer, pp. 169–186.

[622] TANNER JR, J. F. *Analytics and dynamic customer strategy: Big profits from big data*. John Wiley & Sons, 2014.

[623] TAPSELL, J., AKRAM, R. N., AND MARKANTONAKIS, K. Consumer centric data control, tracking and transparency–a position paper. In *2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE)* (2018), IEEE, pp. 1380–1385.

[624] TASCHNER, A. *Business Cases*. Springer Fachmedien Wiesbaden, Wiesbaden, 2017.

[625] TASHAKKORI, A., AND TEDDLIE, C. *Sage handbook of mixed methods in social & behavioral research*. SAGE publications, 2010.

[626] TATIANA, K., AND MIKHAIL, M. Market basket analysis of heterogeneous data sources for recommendation system improvement. *Procedia Computer Science 136* (2018), 246–254.

[627] TEAM, I. G. P. *Eu general data protection regulation (gdpr)–an implementation and compliance guide*. IT Governance Ltd, 2020.

[628] TEECE, D. J. Business models, business strategy and innovation. *Long range planning 43*, 2-3 (2010), 172–194.

[629] TEIGLAND, R., HOLMBERG, H., AND FELLÄNDER, A. The importance of trust in a digital europe: Reflections on the sharing economy and blockchains. *Trust in the European Union in Challenging Times: Interdisciplinary European Studies* (2019), 181–209.

[630] TER HUURNE, M., RONTELTAP, A., CORTEN, R., AND BUSKENS, V. Antecedents of trust in the sharing economy: A systematic review. *Journal of Consumer Behaviour 16*, 6 (2017), 485–498.

[631] TEUBNER, T. Thoughts on the sharing economy. In *Proceedings of the International Conference on e-Commerce* (2014), vol. 11, pp. 322–326.

[632] TEUBNER, T., AND FLATH, C. M. Privacy in the sharing economy. *Journal of the Association for Information Systems 20*, 3 (2019), 2.

[633] TEUBNER, T., SAADE, N., HAWLITSCHEK, F., AND WEINHARDT, C. It's only pixels, badges, and stars: On the economic value of reputation on airbnb. In *ACIS 2016 Proceedings* (2016).

[634] THEMISTOCLEOUS, M., AND MORABITO, V. How can user-centred design affect the acceptance and adoption of service oriented health-

care information systems? *International journal of healthcare technology and management 13*, 5-6 (2012), 321–344.

[635] THOMAS, L., MACMILLAN, J., MCCOLL, E., HALE, C., AND BOND, S. Comparison of focus group and individual interview methodology in examining patient satisfaction with nursing care. *Social Sciences in Health 1*, 4 (1995), 206–220.

[636] THORAT, P. B., GOUDAR, R., AND BARVE, S. Survey on collaborative filtering, content-based filtering and hybrid recommendation system. *International Journal of Computer Applications 110*, 4 (2015), 31–36.

[637] TINTAREV, N., AND MASTHOFF, J. Similarity for news recommender systems. In *Proceedings of the AH'06 Workshop on Recommender Systems and Intelligent User Interfaces* (2006), Citeseer.

[638] TJADEN, J. D., SCHWEMMER, C., AND KHADJAVI, M. Ride with me—ethnic discrimination, social markets, and the sharing economy. *European Sociological Review 34*, 4 (2018), 418–432.

[639] TOCH, E., WANG, Y., AND CRANOR, L. F. Personalization and privacy: a survey of privacy risks and remedies in personalization-based systems. *User Modeling and User-Adapted Interaction 22* (2012), 203–220.

[640] TOLKIEN, J. R. R. *The lord of the rings*. HarperCollins, London, England, 1991.

[641] TRAN, T. N. T., ATAS, M., FELFERNIG, A., AND STETTINGER, M. An overview of recommender systems in the healthy food domain. *Journal of Intelligent Information Systems 50*, 3 (jun 2017), 501–526.

[642] TRANFIELD, D., DENYER, D., AND SMART, P. Towards a methodology for developing evidence-informed management knowledge by means of systematic review. *British journal of management 14*, 3 (2003), 207–222.

[643] TRANG TRAN, T. N., ATAS, M., FELFERNIG, A., AND STETTINGER, M. An overview of recommender systems in the healthy food domain. *Journal of Intelligent Information Systems 50* (2018), 501–526.

[644] TRATTNER, C., AND ELSWEILER, D. Food Recommender Systems: Important Contributions, Challenges and Future Research Directions.

[645] TRATTNER, C., AND ELSWEILER, D. An evaluation of recommendation algorithms for online recipe portals. In *HealthRecSys@ RecSys* (2019), pp. 24–28.

[646] TRATTNER, C., MOESSLANG, D., AND ELSWEILER, D. On the predictability of the popularity of online recipes. *EPJ Data Science 7*, 1 (2018), 20.

[647] TRENTMANN, F. Citizenship and consumption. *Journal of Consumer Culture 7*, 2 (jul 2007), 147–158.

[648] TREVISIOL, M., CHIARANDINI, L., AND BAEZA-YATES, R. Buon appetito: recommending personalized menus. In *Proceedings of the 25th ACM conference on Hypertext and social media* (2014), pp. 327–329.

[649] TRINH, G., RUNGIE, C., WRIGHT, M., DRIESENER, C., AND DAWES, J. Predicting future purchases with the poisson log-normal model. *Marketing Letters 25*, 2 (2014), 219–234.

[650] TSAI, J. Y., EGELMAN, S., CRANOR, L., AND ACQUISTI, A. The effect of online privacy information on purchasing behavior: An experimental study. *Information systems research 22*, 2 (2011), 254–268.

[651] TURBAN, E., RAINER, R. K., AND POTTER, R. E. *Introduction to Information Technology: International Edition*. John Wiley Sons; 3. Edition, 2004.

[652] TWINE, R. A practice theory framework for understanding vegan transition. *Animal Studies Journal 6*, 2 (2017), 192–224.

[653] UETA, T., IWAKAMI, M., AND ITO, T. A recipe recommendation system based on automatic nutrition information extraction. In *International Conference on Knowledge Science, Engineering and Management* (2011), Springer, pp. 79–90.

[654] UNESCO-KOMMISION, D. Bundesweites verzeichnis immaterielles kulturerbe.

[655] URRY, J. The 'system'of automobility. *Theory, culture & society 21*, 4-5 (2004), 25–39.

[656] VAN METEREN, R., AND VAN SOMEREN, M. Using content-based filtering for recommendation. In *Proceedings of the Machine Learning in the New Information Age: MLnet/ECML2000 Workshop* (2000), vol. 30, pp. 47–56.

[657] VARGO, S. L., AND LUSCH, R. F. Evolving to a new dominant logic for marketing. *Journal of marketing 68*, 1 (2004), 1–17.

[658] VEIT, D., CLEMONS, E., BENLIAN, A., BUXMANN, P., HESS, T., KUNDISCH, D., LEIMEISTER, J. M., LOOS, P., AND SPANN, M. Business models. *Business & Information Systems Engineering 6*, 1 (2014), 45–53.

[659] VENABLE, J., PRIES-HEJE, J., AND BASKERVILLE, R. A comprehensive framework for evaluation in design science research. In *International Conference on Design Science Research in Information Systems* (2012), Springer, pp. 423–438.

[660] VENABLE, J., PRIES-HEJE, J., AND BASKERVILLE, R. Feds: a framework for evaluation in design science research. *European journal of information systems 25* (2016), 77–89.

[661] VENDRELL-HERRERO, F., PARRY, G., BUSTINZA, O. F., AND GOMES, E. Digital business models: Taxonomy and future research avenues. *Strategic Change 27*, 2 (2018), 87–90.

[662] VERMEULEN, A. F. *Practical Data Science: A Guide to Building the Technology Stack for Turning Data Lakes into Business Assets*. Apress, 2018.

[663] VON BRIEL, F. The future of omnichannel retail: A four-stage delphi study. *Technological Forecasting and Social Change 132* (2018), 217–229.

[664] VON GAGERN, S. User experience, multiscreen-design und virtual reality: 6 kreativitätstrends 2017–und wie sie damit umgehen, 2017.

[665] VORA, S. *The power of data storytelling*. SAGE Publications Pvt. Limited, 2019.

[666] VREDENBERG, K., ISENSEE, S., AND RIGHI, C. *User-Centered Design: An Integrated Approach with Cdrom*. Prentice Hall PTR, 2001.

[667] VREDENBURG, K., MAO, J.-Y., SMITH, P. W., AND CAREY, T. A survey of user-centered design practice. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (2002), pp. 471–478.

[668] VUORENMAA, R. Monolith erp and current it-trends: Creating a step by step development model (ssdm) for existing monolith erp system to adapt to the current it-trends.

[669] WAHLSTRÖM, J. *Sensor Fusion for Smartphone-based Vehicle Telematics*. PhD thesis, KTH Royal Institute of Technology, 2017.

[670] WAHLSTRÖM, J., SKOG, I., AND HÄNDEL, P. Smartphone-based vehicle telematics: A ten-year anniversary. *IEEE Transactions on Intelligent Transportation Systems 18*, 10 (2017), 2802–2825.

[671] WALTER, J., ABENDROTH, B., AND AGARWAL, N. Pricon: self-determined privacy in the connected car motivated by the privacy calculus model. In *Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia* (2017), pp. 421–427.

[672] WARDE, A. *Consumption: A sociological analysis*. Springer, 2016.

[673] WÄRNESTÅL, P. Multi-disciplinary learning and innovation for professional design of ai-powered services. In *International Conference on Design, Learning, and Innovation* (2021), Springer, pp. 21–36.

[674] WATERMAN, R. W., AND MEIER, K. J. Principal-agent models: an expansion? *Journal of public administration research and theory 8*, 2 (1998), 173–202.

[675] WATSON, H. J. Tutorial: business intelligence–past, present, and future. *Communications of the Association for Information systems 25*, 1 (2009), 39.

[676] WEBSTER, J., AND WATSON, R. T. Analyzing the past to prepare for the future: Writing a literature review. *MIS quarterly* (2002), xiii–xxiii.

[677] WEIDNER, W., TRANSCHEL, F. W., AND WEIDNER, R. Telematic driving profile classification in car insurance pricing. *Annals of actuarial science 11*, 2 (2017), 213–236.

[678] WEITZEL, T., WENDT, O., AND WESTARP, F. V. Reconsidering network effect theory. *ECIS 2000 Proceedings* (2000), 91.

[679] WELLINGTON, B. Making data mean more through storytelling. *TEDxBroadway. https://www. tedxbroadway. com/talks/2018/3/28/making-data-mean-more-through-storytelling-ben-wellington. Zugegriffen 25* (2020).

[680] WIEGAND, G., SCHMIDMAIER, M., WEBER, T., LIU, Y., AND HUSSMANN, H. I drive-you trust: Explaining driving behavior of autonomous cars. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (2019), pp. 1–6.

[681] WIEGARD, R.-B., AND BREITNER, M. H. Smart services in healthcare: A risk-benefit-analysis of pay-as-you-live services from customer perspective in germany. *Electronic Markets 29* (2019), 107–123.

[682] WIENER, M., SAUNDERS, C., AND MARABELLI, M. Big-data business models: A critical literature review and multiperspective research framework. *Journal of Information Technology 35*, 1 (2020), 66–91.

[683] WIKIPEDIA. Verbraucherinformatik — Wikipedia, the free encyclopedia. `http://de.wikipedia.org/w/index.php?title=Verbraucherinformatik&oldid=236616741`, 2024. [Online; accessed 12-May-2024].

[684] WILHELMS, M.-P., HENKEL, S., AND MERFELD, K. You are what you share: understanding participation motives in peer-to-peer carsharing. In *Disrupting Mobility*. Springer, 2017, pp. 105–119.

[685] WIRTH, R., AND HIPP, J. Crisp-dm: Towards a standard process model for data mining. In *Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining* (2000), Springer-Verlag London, UK, pp. 29–39.

[686] WOLF, C., AND BLOMBERG, J. Evaluating the promise of human-algorithm collaborations in everyday work practices. *Proceedings of the ACM on Human-Computer Interaction 3*, CSCW (2019), 1–23.

[687] WOMACK, J. P., JONES, D. T., AND ROOS, D. *The machine that changed the world: The story of lean production–Toyota's secret weapon in the global car wars that is now revolutionizing world industry*. Simon and Schuster, 2007.

[688] WR, S. Institutions and organizations: Ideas and interests, 2008.

[689] WRIGHT, L., AND FULTON, L. Climate change mitigation and transport in developing nations. *Transport Reviews 25*, 6 (2005), 691–717.

[690] WULF, V., MÜLLER, C., PIPEK, V., RANDALL, D., ROHDE, M., AND STEVENS, G. Practice-based computing: Empirically grounded conceptualizations derived from design case studies. *Designing socially embedded technologies in the real-world* (2015), 111–150.

[691] WULF, V., ROHDE, M., PIPEK, V., AND STEVENS, G. Engaging with practices: Design case studies as a research framework in CSCW. In *Proceedings of the ACM 2011 conference on Computer supported cooperative work - CSCW '11* (2011), ACM Press, p. 505.

[692] XIE, K., WU, Y., XIAO, J., AND HU, Q. Value co-creation between firms and customers: The role of big data-based cooperative assets. *Information & Management 53*, 8 (2016), 1034–1048.

[693] XU, G., LI, L., ZHANG, Y., YI, X., AND KITSUREGAWA, M. Modeling user hidden navigational behavior for web recommendation. *Web Intelligence and Agent Systems: An International Journal 9*, 3 (2011), 239–255.

[694] XU, H., TEO, H.-H., TAN, B. C., AND AGARWAL, R. The role of push-pull technology in privacy calculus: the case of location-based services. *Journal of management information systems 26*, 3 (2009), 135–174.

[695] YAMAMOTO, Y. It platforms for utilization of big data. *Hitachi Review 63*, 1 (2014), 46.

[696] YANG, S. *Turning data into strategic assets*. 2020.

[697] YE, Q., LAW, R., AND GU, B. The impact of online user reviews on hotel room sales. *International Journal of Hospitality Management 28*, 1 (2009), 180–182.

[698] YELLA, A., AND KONDAM, A. Big data integration and interoperability: Overcoming barriers to comprehensive insights. *Advances in Computer Sciences 5*, 1 (2022).

[699] YEOW, A., SOH, C., AND HANSEN, R. Aligning with new digital strategy: A dynamic capabilities approach. *The Journal of Strategic Information Systems 27*, 1 (2018), 43–58.

[700] YLIJOKI, O., AND PORRAS, J. Conceptualizing big data: Analysis of case studies. *Intelligent systems in accounting, finance and management 23*, 4 (2016), 295–310.

[701] YOON, D., CHOI, J., KIM, H., AND KIM, J. Future automotive insurance system based on telematics technology. In *2008 10th International Conference on Advanced Communication Technology* (2008), vol. 1, IEEE, pp. 679–681.

[702] ZAPICO, J. L., KATZEFF, C., BOHNÉ, U., AND MILESTAD, R. Eco-feedback visualization for closing the gap of organic food consumption. In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction* (oct 2016), ACM, pp. 1–9.

[703] ZAWADZKI, J. Introducing the AI Project Canvas, Jan. 2019.

[704] ZERVAS, G., PROSERPIO, D., AND BYERS, J. W. The rise of the sharing economy: Estimating the impact of airbnb on the hotel industry. *Journal of marketing research 54*, 5 (2017), 687–705.

[705] ZHANG, Y. Converging data storytelling and visualisation. In *Entertainment Computing–ICEC 2018: 17th IFIP TC 14 International Conference, Held at the 24th IFIP World Computer Congress, WCC 2018, Poznan, Poland, September 17–20, 2018, Proceedings 17* (2018), Springer, pp. 310–316.

[706] ZHANG, Y., LAI, G., ZHANG, M., ZHANG, Y., LIU, Y., AND MA, S. Explicit factor models for explainable recommendation based on phrase-level sentiment analysis. In *Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval* (2014), pp. 83–92.

[707] ZHANG, Z., KUDO, Y., AND MURAI, T. Neighbor selection for user-based collaborative filtering using covering-based rough sets. *Annals of Operations Research 256*, 2 (2017), 359–374.

[708] ZHOU, B. Data integration as a service for applications deployment on the saas platform. In *2013 6th International Conference on Biomedical Engineering and Informatics* (2013), IEEE, pp. 672–676.

[709] ZHU, F., AND ZHANG, X. Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics. *Journal of marketing 74*, 2 (2010), 133–148.

[710] ZHUANG, H., WANG, X., BENDERSKY, M., GRUSHETSKY, A., WU, Y., MITRICHEV, P., STERLING, E., BELL, N., RAVINA, W., AND QIAN, H. Interpretable ranking with generalized additive models. In

Proceedings of the 14th ACM International Conference on Web Search and Data Mining (2021), pp. 499–507.

[711] ZOLNOWSKI, A., ANKE, J., AND GUDAT, J. Towards a cost-benefit-analysis of data-driven business models.

[712] ZOTT, C., AND AMIT, R. Business model innovation: How to create value in a digital world. *NIM Marketing Intelligence Review 9*, 1 (2017), 18–23.

[713] ZOTT, C., AMIT, R., AND MASSA, L. The business model: recent developments and future research. *Journal of management 37*, 4 (2011), 1019–1042.

[714] ZSCHECH, P., FLEISSNER, V., BAUMGÄRTEL, N., AND HILBERT, A. Data science skills and enabling enterprise systems. *HMD Praxis der Wirtschaftsinformatik 55*, 1 (2018), 163–181.