

Grant Agreement Number: 825225

Safe-DEED

www.safe-deed.eu

D2.2 Business models for use cases and generic business models

Deliverable number	<i>D2.2</i>
Dissemination level	<i>Public</i>
Delivery date	<i>25 November 2020</i>
Status	<i>Final</i>
Author(s)	<i>Mark de Reuver, Wirawan Agahari, Ricardo Dolci, Gert Breitfuss, Michael Fruewirth</i>



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 825225.

Changes Summary

Date	Author	Summary	Version
23.6.20	Mark de Reuver	Outline created	0.1
29.6.20	Wirawan Agahari	Contents added	0.2
21.7.20	Gert Breitfuss	Contents added	0.3
20.10.20	Mark de Reuver	First complete draft	1.0
26.10.20	Mark de Reuver	Formatted correctly	1.1
4.11.20	Mark de Reuver	Added summary	1.2
4.11.20	Gert Breitfuss	Internal WP review	1.3
10.11.20	Dieter Decraene, Alessandro Bruni	Internal WP review	1.4
4.11.20	Yiannis Markopoulos	Internal review	
25.11.20	Mark de Reuver	Final version	2.0

Executive summary

This report develops business models for Safe-DEED technologies. First, a background is provided on the context of data marketplaces, which is the likely setting where Safe-DEED technologies will be utilized. The background also covers the specific Safe-DEED technology group of multiparty computation.

Next, two qualitative studies are presented. In one of the studies, business model tools from D2.3 are applied to develop generic business models for Safe-DEED technologies. Barriers and incentives for data sharing are found in a workshop. In the other study, qualitative interviews with data marketplace stakeholders and experts are conducted, in order to find affordances and business models for Safe-DEED technologies. New value propositions are found (i.e. privacy, control, no need for trust), new value capturing models (i.e. pay what you want, protect the buyer) and three new value delivery architectures (i.e. peer-to-peer, intermediary and aggregator).

Specific business models for the WP6 use case are developed next. Three scenarios are examined, which involve different ways of data sharing within as well as between organizations.

The exploratory studies reported on in this report provide a rich understanding of what business models for Safe-DEED technologies may look like. Our purpose is to lay a basis for further work. Specifically, the exploratory work on the affordances and value creation of Safe-DEED technologies developed in this report will be further tested in T2.4. The contextualized business models in the report provide a basis for illustrating and testing the business model tools from T2.3. Our understanding of business models will further be utilized as a basis for economic modelling in T2.4 as well as for developing specific exploitation strategies in WP8.

Table of Contents

1	Introduction	7
2	Background.....	7
2.1	Data marketplaces.....	8
2.2	Multi-party computation	9
2.3	Preliminary analysis.....	11
3	Qualitative research	12
3.1	Workshop: Added value of Safe-DEED technologies for data sharing.....	12
3.1.1.	Approach	12
3.1.2.	Results	13
3.2	Affordances of MPC for data marketplace operators: Qualitative interview study.....	15
3.3	How MPC transforms the business model of data marketplaces	16
3.2.1.	Approach	16
3.2.2.	Results	17
4	Business models for the use cases.....	19
4.1	Data sharing between departments.	19
4.2	Data sharing between two firms that have a joint interest.....	20
4.3	Selling data to firms in other industries.	21
5	Conclusions	21
	References	22

List of Figures

Figure 1: Roles in data marketplaces ecosystem, adapted from Spiekermann (2019).....	8
Figure 1: Roles in data marketplaces ecosystem, adapted from Spiekermann (2019).....	10
Figure 3: Designed canvas for interactive session 1 to support collection of data sources.....	13
Figure 4: Data sharing between departments	20
Figure 5: Data sharing between two different firms.....	21
Figure 6: Selling data to firms in other industries	21

Abbreviations

BM :	Business Model
GDPR :	General Data Protection Regulation
MPC :	Multi Party Computation

1 Introduction

The purpose of Task 2.2 is to design and evaluate business models (BMs) for Safe-DEED, with a specific focus on privacy and confidentiality preserving technologies.

Technologies have no value in and on themselves, but only when supported by viable business models (Chesbrough 2010). Business models help to make explicit how technologies help actors to create and capture value (Bouwman et al 2008). A viable business model not only creates value for the providers of the technologies, but also for its users (Bouwman et al 2008).

In the context of Safe-DEED, there are two basic classes of technologies. The first class is privacy and confidentiality enhancing technologies as developed in WP5. Within that work-package, these are specifically multiparty computation but also deanonymization checks. The second class is technologies for evaluating the value of datasets: data valuation technologies as developed in WP4.

To understand the business models, it is essential to specify the context-of-use. The primary use context for Safe-DEED is data marketplaces. However, the technologies can also be used in different scenarios, such as sharing of data between departments within a company, or sharing of data with known actors such as supply chain partners.

Safe-DEED technologies create direct value. For instance, privacy-preserving technologies can improve the privacy of citizens, which is a source of intangible value. Data valuation technologies can help data owners to charge a fairer price for their datasets, thus increasing their revenues. However, Safe-DEED technologies also indirectly create value. As Safe-DEED technologies (are meant to) incentivize data owners to expose datasets on data marketplaces, data users will gain by having more relevant datasets and information. The value of these datasets, in turn, depends on the use context by the data owner. The use of datasets to improve business is referred to as data-driven BMs. In this way, there are two crucial interactions between data-driven BMs and Safe-DEED:

- Safe-DEED technologies (are assumed to) lead to increased availability of datasets through, e.g. data marketplaces, thus enabling data users to implement data-driven BMs
- Data users will only use data marketplaces or pay for datasets if they derive value from the data; hence data-driven BMs are conditional for any viable BM of Safe-DEED technologies

For these reasons, a basic understanding of data-driven BMs is required in T2.2, and a thorough analysis is required of how and why Safe-DEED technologies affect the feasibility and viability of data-driven BMs.

In this deliverable, we first provide a broad background, based on extant literature, on data-driven business models and data marketplaces (Section 2). Based on this understanding, Section 2 concludes with an analysis of how Safe-DEED technologies may create value for different actors involved. Next, we provide an overview of the qualitative research done to develop business models. Specifically, we report the results from a workshop and a series of qualitative interviews (Section 3). Next, we design business models for three business scenarios derived from the use case in WP6, using the business model tools being developed in T2.3 (Section 4). Section 5 discusses and concludes the results and attempts to define generic business models for Safe-DEED. We utilize primarily the business model tools that T2.3 produced, rather than more generic business model tools, given that these tools are specifically tailored for our purposes.

2 Background

This section provides a background on the primary use context for Safe-DEED technologies: data marketplaces and data-driven business. We consider that Safe-DEED technologies directly adds most

value within a data marketplace context. Here, privacy-preserving technologies may make data owners more willing to trade data, and data valuation technologies can help to create better incentives for data sharing. Further, we consider that Safe-DEED technologies indirectly create value mostly by enabling new data-driven business models. If businesses start to trade more data, data-driven business models become possible that were not feasible before.

2.1 Data marketplaces

Data marketplaces are digital platforms that enable organizations to share and sell datasets (Koutroumpis et al., 2017; Richter & Slowinski, 2019; Spiekermann, 2019). Access to the data, manipulation and the use of the data by other entities is commonly governed by the data marketplace using a range of standardized or negotiated licensing models (Schomm, Stahl, & Vossen, 2013; Stahl, Schomm, Vossen, & Vomfell, 2016). Both static and dynamic data streams can be shared and traded in data marketplaces, in which it is accessible via individual file downloads, Application Programming Interfaces (APIs) or customized web interfaces (Fricker & Maksimov, 2017; Spiekermann, 2019). On top of that, data marketplaces also offer complementary applications and services such as data visualizations, data valuation and data analytics (Schrieck, Hein, Wiesche, & Krcmar, 2018; Spiekermann, 2019; van den Broek & van Veenstra, 2018). Hence, such platforms would create value to its participants by lowering transaction costs, stimulating innovation by third-party developers and generating network effects.

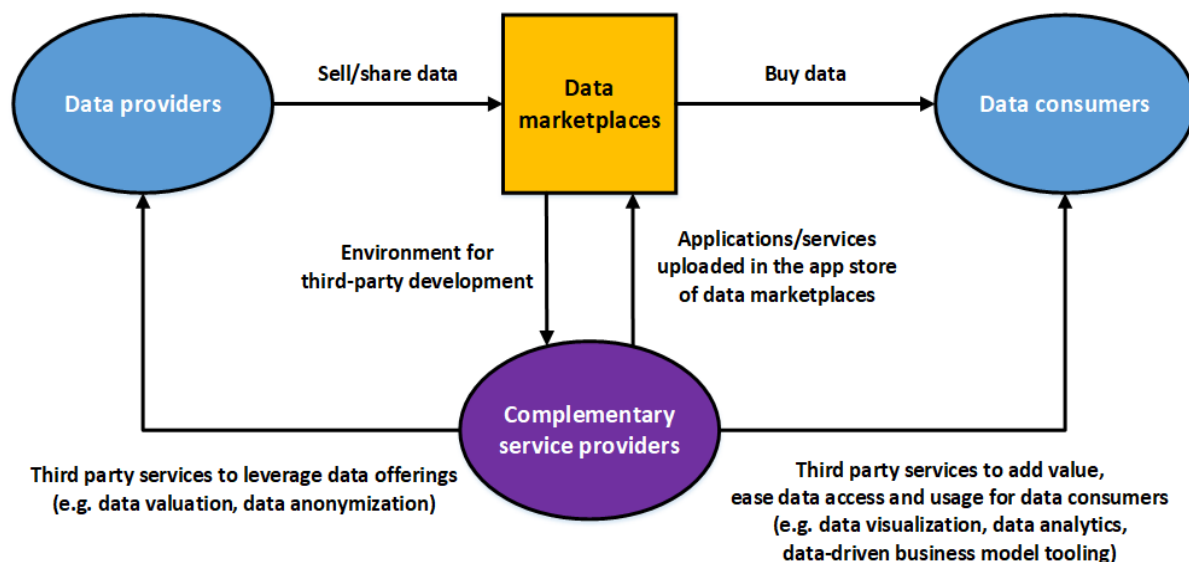


Figure 1: Roles in data marketplaces ecosystem, adapted from Spiekermann (2019)

Figure 1 illustrates the actors involved in the data marketplaces ecosystem. The core function of data marketplaces is to match between two sides of the market. On the one side, data providers want to monetize their data by sharing/selling it via data marketplaces. Then, there are data consumers on the other side of the market who want to buy the data products offered by data providers, and therefore access data marketplaces and look for available data. Other than that, data marketplaces also provide an environment for complementary service providers so that they can join the platform to develop data-driven applications and services. Examples include data anonymization, data valuation, data visualizations and data analytics. These applications and services are uploaded in the app store provided by data marketplaces, which can be used by data providers to leverage data offerings or by data consumers to add value to the data that they bought.

Value proposition	Transaction-centric			Data-centric	
Market positioning	Owned by the data provider			Neutral	
Market access	Closed		Hybrid	Open	
Integration	Domain-specific			General	
Transformation	Raw data	Normalization	Aggregation	Quality assurance	
Architecture	Centralized		Hybrid	Decentralized	
Price model	Free	Fixed price/ Subscription	Package	Pay-per-use	Progressive price
Revenue model	Free	Freemium	Flat rate	Fee	
	Listing fee	Transaction fee/ commission	Service fee	Storage fee	

Table 1: Taxonomy of data marketplaces (Spiekermann, 2019)

Spiekermann (2019) developed a taxonomy of data marketplaces based on eight attributes, as can be seen in Table 1. According to this taxonomy, data marketplaces can focus only on the direct switching of data (transaction-centric) or provides complimentary services (data-centric). Regarding ownership, some platforms are owned by data providers, while others are provided by independent third-party (neutral). Platform owners can also choose to target broad and unknown participants (open), limited to certain partners (closed) or somewhere in between (hybrid). Meanwhile, in terms of architecture, data marketplaces can have a centralized (i.e. central location for data storage), decentralized (i.e. data providers keep their data) or a hybrid approach. Moreover, data marketplaces can have a broad domain spectrum, either general (i.e. not focusing on specific areas) or domain-specific. Besides, the data traded in data marketplaces can simply be raw/unprocessed data, standardized/normalized data, aggregated data, or high-quality data with quality assurance checks. Furthermore, different pricing and revenue model can be distinguished, including (but not limited to) free of charge, fixed price/subscription and pay-per-use.

Fruewirth et al (2019) also create a taxonomy of data marketplaces, and find the following archetypes:

- **Centralized data trading:** Marketplaces for selling and buying data from any origin, domain, type or price
- **Centralized data trading with smart contract:** A centralized data trading infrastructure, amended with smart contracts between the buyers and sellers
- **Decentralized data trading:** Decentralized buying and selling of data
- **Personal data trading:** Data marketplaces on which consumers sell their data to businesses

2.2 Multiparty computation

MPC is a cryptographic technique where two or more parties perform a joint computation, which results in a meaningful output without disclosing the input provided by either party (Bestavros et al., 2017; J. I. Choi & Butler, 2019; Zhao et al., 2019). Conceptually, MPC makes it possible to balance the interest between different actors. On the one hand, data consumers (i.e. businesses that use data or insights) can gain insights from the data shared by data providers in a secure manner. On the other hand, data providers can also get security assurance because they can retain the secrecy of the data.

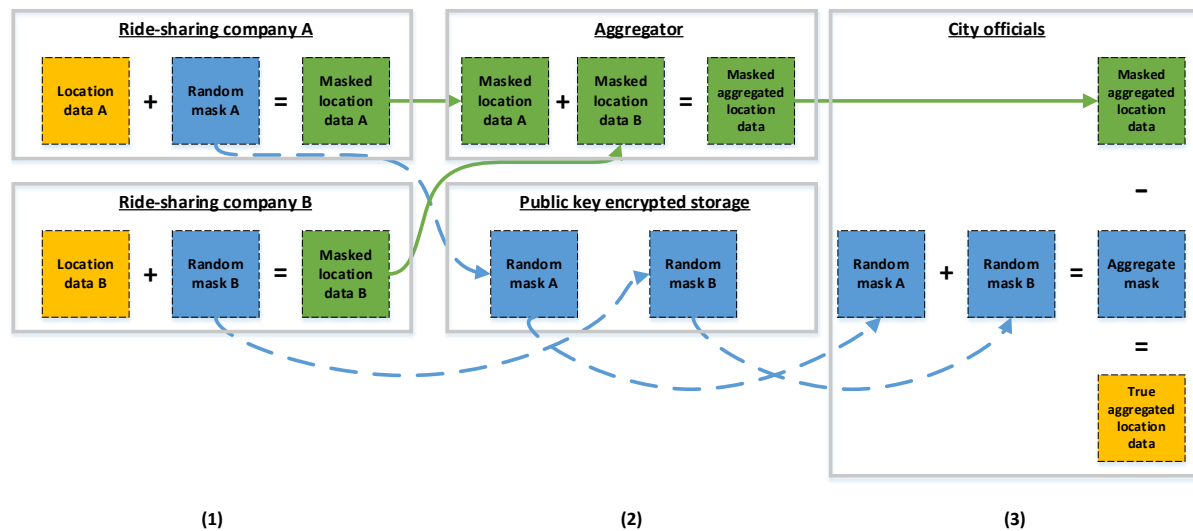


Figure 2 illustration of MPC, adapted from Bestavros et al. (2017)

Figure 2 shows the illustration of how MPC works. To contextualize this illustration in a real-life setting, consider an example use case where city officials (column (3) in Figure 2) are trying to understand the influence of ride-sharing vehicles on traffic congestion. Therefore, some essential data held by ride-sharing companies (column (1) in Figure 2) are needed. This data includes, for example, popular pick-up spots and the number of cars in service during rush hour. However, this is confidential and sensitive data, meaning that releasing such information may result in adverse effects such as losing a competitive advantage over rivals. In this case, we can then use the MPC-based solution to allow the aggregation of ride-sharing data from companies without actually disclosing the individual data point.

For the MPC-based solution, the ride-sharing companies will apply a random number to mask/protect their data. In this way, we ensure that the actual value of the data cannot be read anymore. An aggregator then aggregates this masked data. At the same time, the public key encrypted storage aggregates only the different random masks, which do not hold any data, used by the companies (see column (2) in Figure 2). Finally, the requester party (in this case city officials) then receives the aggregated masked data and the aggregated mask. They can then use the aggregated mask to transform the masked aggregated results into the plain-text aggregated results (see column (3) in Figure 2). In this stage, the city officials now hold the plain text aggregated data, for example, to build heat-maps, without any party involved in the computation having access to other parties' plain text data.

MPC could overcome barriers of data sharing in the business-to-business context. By using MPC, data providers could regain control over their data since it is not necessary to exchange data. Instead, data consumers will only receive insights from the computation of multiple datasets. This is a value proposition that MPC offers: allowing data sharing safely and securely. In such a way, MPC can also help to deal with compliance, depending on the way it is implemented (Archer et al., 2018)¹. Ultimately, MPC could potentially increase trust in sharing data via data marketplaces.

It is important to be aware that massive implementations of MPC in real-life settings are yet to happen and still limited to only a few applications, such as auction-based pricing (Bogetoft et al., 2009), tax fraud detection (Bogdanov, Jõemets, Siim, & Vaht, 2015) and satellite collision prevention (Hemenway, Lu, Ostrovsky, & Welser Iv, 2016). There are multiple barriers to this lack of implementation, such as usability issues (e.g. too complex to understand by non-experts, suspicion in the computation results), technical issues (i.e. performance limitations and scalability) and legal aspects (i.e. current regulations discourage cooperation) (Choi & Butler, 2019). Furthermore, to the best of our knowledge, the

¹ More on the legal implications of MPC in Deliverable D3.4

application of MPC within the data marketplaces setting is still scarce. A notable exception is Roman & Vu (2018), who developed a proof-of-concept of data marketplaces based on smart contracts and MPC.

2.3 Preliminary analysis

In Table 2, we provide an initial analysis of how Safe-DEED technologies create value for the stakeholders in a data marketplace ecosystem. We distinguish value from privacy/confidentiality preserving technologies (such as MPC and de-anonymization checks that feature in WP5) and data valuation technologies (such as the pricing models that feature in WP4).

Centrality	Stakeholder type	Value derived from privacy/confidentiality preserving technologies	Value derived from data valuation technologies
Outside	Citizens, businesses represented by the last two rows?	Improved privacy / confidentiality	N/A? Compensation for allowing third party access to data?
Second	Data owners / Data Provider (provides data for third party to use, sets pricing and/or usage conditions)	Reduced barriers to expose data through marketplace Reduced risk (i.e. compliance with regulation) Revenues, other intangible benefits	Opportunities to charge for data
Central	Data marketplace provider/operator (offers infrastructure and WP4/5 technologies)	More datasets being exposed Increased use of data marketplace Revenues from e.g. licensing privacy-preserving technologies	More datasets being exposed Increased use of data marketplace
Second	Data users /Data Market Customer (uses data services and infrastructure for commercial purposes)	Increased access to relevant datasets / information à Enables data-driven business models	Increased access to relevant datasets / information à Enables data-driven business models
Third	End User (buys data services/applications from data market customer or non-commercial usage of data)	Improved privacy / confidentiality	

Table 2 Preliminary analysis: value created by Safe-DEED technologies for data marketplace stakeholders

3 Qualitative research

In this section, we describe three qualitative studies done on the business model implications of Safe-DEED technologies. First, a workshop was conducted with practitioners and scholars (Section 3.1). Next, qualitative interviews were done with data marketplace operators about the affordances of MPC for their business models (3.2). Finally, qualitative interviews were conducted with data owners to explore how MPC would change their perspectives on data sharing (3.3).

3.1 Workshop: Added value of Safe-DEED technologies for data sharing

3.1.1. Approach

We conducted a workshop with business actors as an exploratory to better understand the barriers and incentives of business-to-business data sharing from the firms' perspective. Understanding barriers and incentives are beneficial to clarify what kind of risks and trust issues that exists, why they affect intention to share data and exploring alternative explanations that are needed to control for in the experiment.

The workshop was conducted in Graz, Austria, in November 2019 as a part of a larger European project. In total, 27 experts and representatives of firms that are interested in the data economy took part in this workshop.

In the workshop, participants were first asked to think about their company or personal data that could be valuable for others, but they were not sure about sharing with other parties. They were asked to list types of data, and to discuss why they were unsure about sharing, and what would incentivize them to share. The canvases shown in Figure 3 were used to facilitate the discussion. The goal of this first round was to get participants in the frame of mind to think about what makes them reluctant to share potentially valuable data.




PRIVATE (sensitive personal data)	COMPANY (confidential/not used data)	PUBLIC INSTITUTIONS (confidential data)
<p>What kind of private and/or sensitive data do you think is valuable for others, but you are not sure about to share to other parties?</p> <p>e.g. health, income, consumption data etc...</p> <ul style="list-style-type: none"> • ... • ... • ... • ... 	<p>What kind of confidential/not used company data do you think is valuable for others, but you are not sure about to share to other parties?</p> <p>e.g. customer/supplier data, process or machine data, etc...</p> <ul style="list-style-type: none"> • ... • ... • ... • ... 	<p>What kind of confidential data from public institution do you think is valuable for others, but is currently not available?</p> <p>e.g. funding data, budget data, taxes, social insurance etc...</p> <ul style="list-style-type: none"> • ... • ... • ... • ... 
<p>Why are you unsure to share/make your data available to other parties?</p> <p>I am unsure to share/make my data available to other parties because</p> <ul style="list-style-type: none"> • ... • ... • ... 	<p>Why are you unsure to share/make company data available to other parties?</p> <p>I am unsure to share/make company data available to other parties because</p> <ul style="list-style-type: none"> • ... • ... • ... 	
<p>What would incentivize you to share/make your data available to other parties?</p> <p>I want to share/make my data available to other parties if....</p> <ul style="list-style-type: none"> • ... • ... • ... 	<p>What would incentivize you to share/make company data available to other parties?</p> <p>My company is willing to share/make the data available to other parties if....</p> <ul style="list-style-type: none"> • ... • ... • ... 	

Figure 3: Designed canvas for interactive session 1 to support collection of data sources

In the second round, participants were asked to imagine that the data could be shared entirely securely while preserving privacy, thanks to Safe-DEED technologies. They were asked to think what

information could be extracted if the "private sensitive, confidential" data were available, what services a third party could create with the data, and who would benefit from such a service. Again, a canvas was used to facilitate the discussion.

3.1.2. Results

In the first round of the workshop, we found that participants were afraid that sharing data with other parties would create knowledge spillovers resulting in competitive disadvantages over rivals. Legal concerns were also dominantly discussed since there is a lack of clarity in terms of process and consequences. Other barriers discussed include the absence of an internal process to support data sharing and the difficulty in quantifying the value of the data.

In terms of incentives for data sharing, participant suggestions are relatively straightforward. They demand a clear benefit, either tangible (e.g. money/revenue stream) or intangible (e.g. benchmarking, value-added services). Other participants suggest a clear and established regulation in data sharing as an essential incentive for them. Finally, there is a need for a guarantee and protection of the data to make sure that firms that provide data will maintain their competitive advantage.

Table 3 provides an overview of the mentioned barriers and incentives.

Barriers	Competitive (dis)advantage, knowledge spillovers, industrial spying
	Legal concerns (GDPR), unclear process and consequences
	Internal resistance, absence of internal process
	Difficult to quantify the value of data
Incentives	Clear benefits: money/revenue stream, image/reputation, benchmarking, value-added services
	Established regulation and protection

Table 3: Results from the first round of workshop: barriers and incentives for data sharing

The second round of the workshop provided results, as displayed in Table 4. We find a wide range of purposes for sharing data that is currently kept private. Some purposes are related to public values (e.g. health) whereas others are related to private values (e.g. benchmarking, marketing).

	PRIVATE (sensitive personal data)	COMPANY (confidential/ not used data)	PUBLIC INSTITUTIONS (confidential data)
Group 1	early cancer detection	any data as long as it doesn't harm my market position	n/a
	early diabetes risk detection	benchmarking service (KPI comparison)	
	early donor suitability	vertical (=no competition) vs. horizontal (=competition) clustering	
	anti terrorism actions		

	digital personal assistants any personal data is sensitive always depends on the context		
Group 2	health recommendation / consultancy / smart living customized products diet management recommendations for life style	smart company / improve value chain know the client better ?? Understanding / new products planning smart hiring / qualification analysis and check suitability of the employee to the environment synergies between companies centralize / federate data processing lower costs if we can upload company data to a private cloud	aligning education with business needs classification of people, cities, ... optimize public transportation, health care system security economy planning based on companies data, projections, etc. use private clouds
Group 3	dynamic scheduling / better planning of public transport fair salary service retirement planner fair loan service nutrition planner, adviser energy consumption benchmarking service (w. gamification)	"not reinventing the wheel service"; balanced innovation production benchmarking procurement benchmarking predictive maintenance energy saving	career advice dynamic curriculum "find an affordable flat" service urban planning
Group 4	genetic data diagnostics --> predisposition --> treatment	sharing information as ML models energy consumption of the household --> (bad) what device is used (TV) --> (good) energy provider balances production --> company	efficiency (e-Government) crime detection

	--> location of health ??	--> clients (cheaper, stable energy)	
	imaging data diagnostics	company = banks --> third party = MPC company --> fraud detection --> government / society	
	--> better ML models for diagnostics	Production data --> supply chain improvement (external data eg. GPS data) --> maintenance scheduling	

Table 4 : Results from second round of workshop: value created by sharing data through Safe-DEED technologies

3.2 Affordances of MPC for data marketplace operators: Qualitative interview study

The workshop reported in Section 3.1 shows that data sharing is hindered mainly by concerns over legal as well as economic issues. For instance, workshop participants reported their concerns over sharing data that would ultimately help their competitors or lead to knowledge spill-overs.

Safe-DEED technologies, and specifically MPC, are expected to enhance the control that data marketplace operators have over data sharing. Theoretically, MPC enhances control over data sharing, as data is not fully disclosed, but only the answers to queries. However, there has been no research to date over whether MPC indeed increases the control that data marketplace operators have.

In this section, we summarise the results of an MSc thesis that was part of WP2², in which qualitative interviews were done with data marketplace owners and MPC experts. The study aimed to investigate the potential adoption of secure MPC by a data marketplace provider for realizing control over data sharing. Semi-structured interviews were conducted among data marketplace providers operating in the mobility domain, data marketplace experts and MPC developers and experts. To guide the study, affordance theory was applied.

The study shows that adoption of MPC could generate three main affordances for a data marketplace provider in terms of platform control: (1) preserving the data, (2) enabling data ownership and (3) preserving the result of the computation. These affordances are generated by the relationship between the data marketplace provider's goals in terms of platform control and the features of the MPC technology. Regarding the former, the following goals were identified: (1) ensure the security and the privacy of the data; (2) guarantee that a data provider has complete control over its data; (3) ensure the correct execution of the computation. Concerning the latter, three critical features offered by the MPC technology could enable platform control: (1) information-theoretic security or computational security, (2) agreement protocols before starting the computation and identification mechanisms if someone deviates from it, (3) and correct execution of the computation. Three factors could influence the realization of the affordances: (1) perception of the technology, (2) need for the technology, and (3) degree of effort required. The results showed that secure MPC could satisfy several different needs of a data marketplace provider.

Some constraints were found that can influence the adoption of MPC among data marketplace providers. Firstly, a data marketplace provider may perceive the MPC as unsafe because of the difficulty to understand the technology. Secondly, a data marketplace provider could consider that secure MPC does

² <https://repository.tudelft.nl/islandora/object/uuid%3A1d568346-86d5-402b-babe-26d2ba46809b>

not currently present an adequate maturity level to adopt the technology in its platform. Finally, a data marketplace provider could prefer to maintain its current situation to avoid a radical change. The adoption of MPC technology by a data marketplace provider could cause several impacts on its platform. If the platform has a centralized structure, the data will not be stored in the platform anymore, but they will remain with the data provider. Moreover, if a data marketplace focuses only on data exchange offerings, it would be able to offer a new type of product in its platform (e.g. insights). Finally, the adoption of the MPC in a data marketplace could cause additional overhead in the functioning of the platform.

The main results are summarised in the overview table below.

MPC features	Affordances of MPC for data marketplace operators	Control objective of data marketplaces	Factors	Impacts
Information-theoretic security/computational security	Preserving the data	Ensure data and privacy security	Perception of technology	New product offered
Agreement protocols among the participants before starting the process and identification mechanisms if someone deviates from it	Enabling data ownership	Guarantee that the data owner has full control over its data	Need for technology	Decentralized architecture
Execute the computation correctly	Preserving the result of the computation	Ensure the correct execution of the computation	Resources availability	Additional overhead

Table 5: Affordances of MPC for control over data in data marketplace ecosystems

3.3 How MPC transforms the business model of data marketplaces

In a third qualitative study, we examined how privacy-preserving technologies change the business model of data marketplaces, focusing on multiparty computation (MPC) as one specific technology.

3.2.1. Approach

As our primary data collection method, we conducted semi-structured interviews with MPC experts and practitioners. Experts were sourced by looking into relevant publications and white papers, as well as by utilizing the personal networks of the WP2 researchers. Our sampling approach resulted in fifteen interviewees with varying diverse backgrounds from academia, research institutions, and industry. The majority of them are at the senior level of their respective affiliation and have more than seven years of experience working on MPC.

In the interview, first, we explain the definition of the data marketplace and example use-case. After that, we also explain the meaning of MPC concept and an illustrative use-case. We validate this presentation with an expert before presenting it to the interviewees. The interviews questions were semi-structured and took around one hour on average, including the presentation. We asked questions related to business models: the value proposition of MPC in a data marketplace, how MPC changes the

architecture of a data marketplace, and the new revenue sources enabled by MPC for a data marketplace. Transcripts were made and anonymized. Transcripts were coded and analyzed through a qualitative data analysis software tool. In analyzing the transcript, we follow the three steps of coding: open coding, axial coding, and selective coding (Bryant & Charmaz, 2007).

3.2.2. Results

We provide here an overview of results, an extensive version is currently under review for the European Conference on Information Systems.

Regarding the value proposition of MPC for data marketplace stakeholders, we find three main types:

- **Privacy:** protecting the input data. Through MPC, data buyers can learn only about the output of a computation, and will not be able to learn from input. This prevents competitors to get advantage (e.g. reverse engineering).
- **Control.** MPC enables to control what kind of query other parties can run on the data. This needs to be agreed before running a query.
- **Get rid of trust.** Through MPC, there is, in theory, no need to trust a third party or the data buyer.

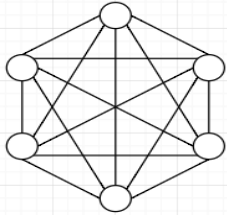
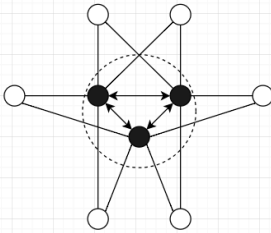
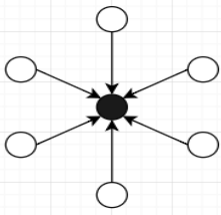
We find two new main ways of capturing the value that is enabled by MPC:

- **Pay-what-you-need:** instead of paying for an entire dataset, data buyers only need to pay for the insights/aggregation from multiple data sellers.
- **Protect the buyer:** MPC can also protect data buyers in revealing what they actually want to data sellers. So, it also prevents knowledge spillover from data buyers to data sellers.

Regarding the value architecture, we find three main scenarios from the interviews.

- **Peer-to-peer:** data markets only act as broker/matchmaking two sides of the market, MPC agent deployed in each data providers, direct data exchange between buyers and sellers with MPC
- **Intermediaries:** data markets as a broker and hosting computational infrastructure (multiple MPC servers)
- **Aggregator:** data markets aggregate data from various companies, then sell the data to data buyers using MPC protocol

We found that there are mainly three types of MPC deployment scenarios in a data marketplace. Each type could result in a different degree of trust requirements as well as privacy and security guarantees for data providers. We also look into value architecture's implication of MPC towards value finance in terms of complexity and resource provision for both data marketplace operators and data providers.

Aspect	MPC deployment scenario in a data marketplace		
	Peer-to-peer	Intermediaries with multiple & independent computing server	Intermediaries with a single computing server
Illustration			

Data marketplace type	Data broker & Data aggregator	Data broker	Data broker
Computation type	Synchronous	Synchronous & Asynchronous	Asynchronous
The trust required by data providers	No trust required towards the intermediaries, as no intermediaries involved Only need to trust that the MPC software runs correctly	Trust towards intermediaries are distributed to multiple entities that provide computing servers Data providers need to trust intermediaries that they do not collude and reveal the input data	Data providers only need to trust one computing server that it will not see and reveal the data
Complexity & resource provision	More effort is needed to set up the MPC software and infrastructure on the data providers' side	Data marketplace operator needs to establish a partnership with multiple entities to provide computing servers jointly	Data marketplace operator only have to provide one computing server that they can provide themselves
Privacy & security guarantee	More robust privacy & security guarantee as the computation is performed on the data providers' side	Privacy & security is guaranteed as long as there is at least one honest computing server All computing servers may collide and combine the secret-shared data	If there is a breach or leak, then the input data can be easily exposed

Table 6: Cross-analysis for three deployment scenario of MPC in a data marketplace

We found that all three of the MPC deployment scenarios are suitable for the data brokering model. The focus of this model is to provide a matchmaking service between data providers and data consumers and not facilitate data exchange between both parties. Hence, data marketplace operators could opt for peer-to-peer architecture and provide technical expertise to install MPC protocol on the client-side. In this way, the data exchange is performed directly between data providers and data consumers without involving data marketplace operators. However, data marketplace operators could also choose to offer computational infrastructure as a service to ease the burden for data providers and consumers. In this regard, the intermediary architecture (either single or multiple servers) would be more suitable. While this architecture requires data marketplace operators to be involved in the computation, they would not be able to see the data as it remains encrypted throughout the process, and only data consumers can access the computation results.

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

In terms of the computation type, we also found that synchronous computation is most compatible with the peer-to-peer architecture. MPC protocol generally requires all parties to be online and present at the same time. Peer-to-peer architecture would make this possible, as MPC protocol will be installed in all parties and allowing them to be connected and present during the computation. The synchronous computation can be organized independently without the need to have a trusted third party in the middle. Nevertheless, we also see the potential of using multiple computing servers as intermediaries to facilitate synchronous computation. In this setting, all participating parties do not need to be present simultaneously, but only the multiple servers in the middle. For the asynchronous computation, intermediaries' presence is essential to coordinate the computation process between all parties to participate at different points in time. For this reason, the intermediary architecture (either single or multiple computing server) is the most suitable approach for the synchronous computation.

Looking at how different architecture could result in a different value proposition, we found that data providers do not need to trust data marketplace operators in a peer-to-peer architecture as they are not

involved in the computation. Since the computation runs on each party (i.e., data providers and data consumers), this architecture offers a robust privacy and security guarantee. However, more effort is needed to set up the infrastructure and MPC software for each participating party, increasing the cost and complexity.

Meanwhile, a single server architecture is relatively straightforward because data providers and consumers do not need to prepare additional infrastructure and software on their side. Instead, data marketplace operators only have to provide a single computing server on their side as an MPC engine that they can provide themselves. This approach would reduce adoption costs for data providers and data consumers in a data marketplace and possibly attract them to join the platform. Concerning the trust requirement, data providers need to trust this single computing server to perform the computation correctly and not see the original data. The challenge would lie in the privacy and security guarantee: if there is a breach or leak on this single server, the input data can be easily exposed.

A multiple server architecture serves as an alternative that positioned itself between the previously mentioned architectures. In this setting, each computing server is offered by an independent entity that is not related to each other. Together, they act as intermediaries that perform MPC computation. Hence, instead of only trusting one intermediary, data providers need to distribute their trust towards those different computing parties. In particular, data providers need to trust that those computing parties do not collude and reveal the input data, which might be a drawback of this architecture. Nevertheless, privacy and security are guaranteed as long as at least one computing server does not behave maliciously. Like single server architecture, data providers and data consumers can simply use the data marketplace without deploying additional infrastructure and software. However, data marketplace operator needs to establish a partnership with multiple entities to provide computing servers jointly, which might increase complexity on their side.

4 Business models for the use cases

Section 3 gives a broad and generic understanding of how Safe-DEED technologies enable new business models. In this section, we apply these insights for the specific use case setting of WP6. We omit WP7 here because that use case was extensively treated in deliverable D2.4. We apply the tools developed in T2.3 to visualize the business model scenarios.

4.1 Data sharing between departments.

Sometimes, data is not being allowed to share between departments. For instance, data collected about customer behaviour that is necessary for the operations of a telecom firm cannot, without the consent of customers, be transferred to a marketing department. Besides, employees may not trust others from other departments, or the management of a firm may not trust its employees handling raw data. All these instances demonstrate a need for privacy-preserving technologies and deanonymization checks when sharing data between departments

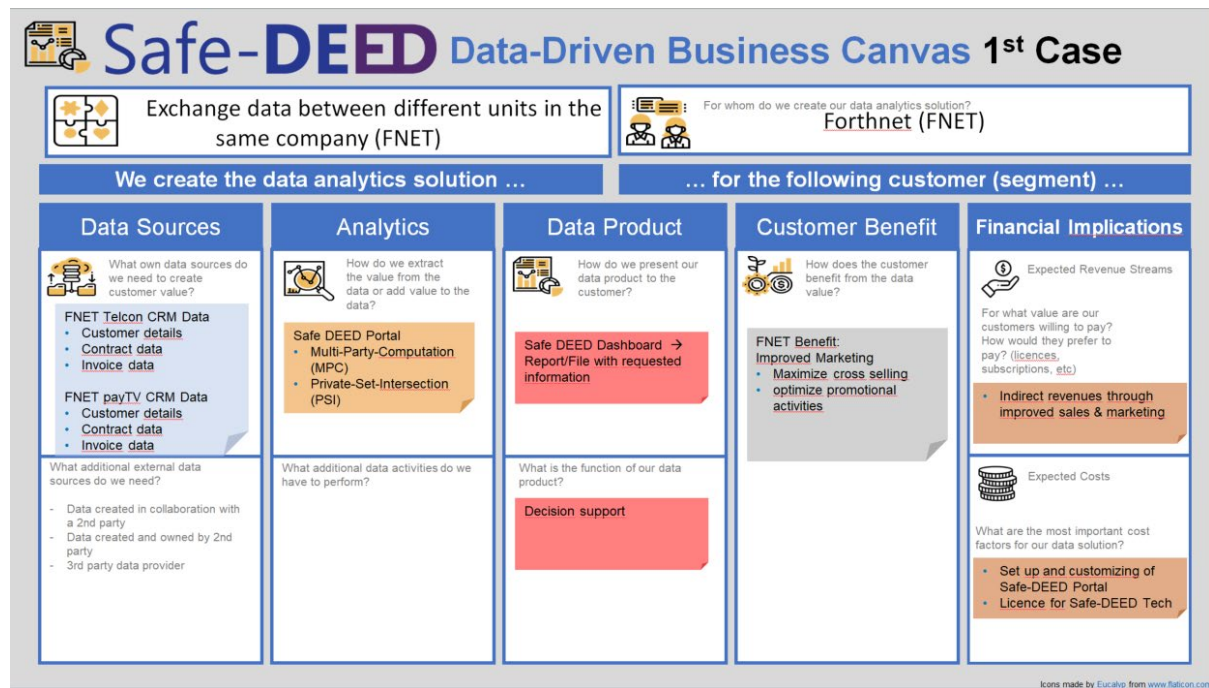


Figure 4: Data sharing between departments

4.2 Data sharing between two firms that have a joint interest.

Here, for example, firms could share customer relationship management data to improve their joint marketing programs. The example of a bank and a telecom firm is mentioned, which have an overlapping geographical reach, and would like to find out where opportunities exist to target each others' clients.

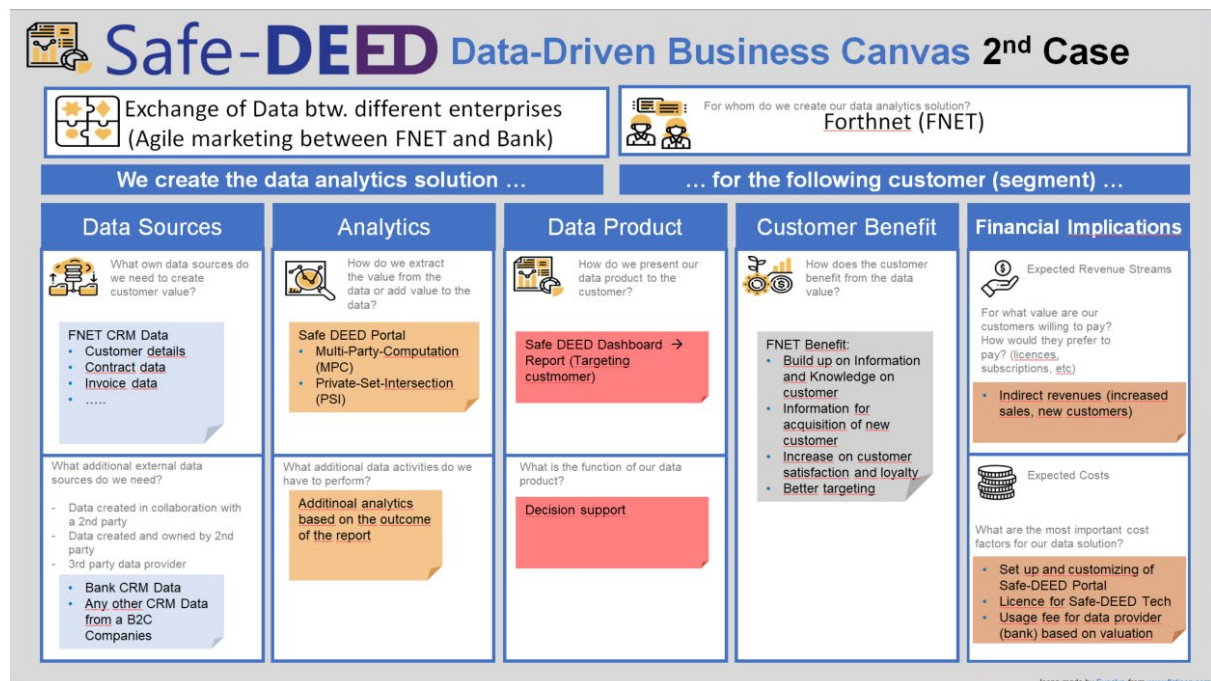


Figure 5: Data sharing between two different firms

4.3 Selling data to firms in other industries.

This scenario becomes the closest to a data marketplace. The telecom firm would make its data available to firms in other industries

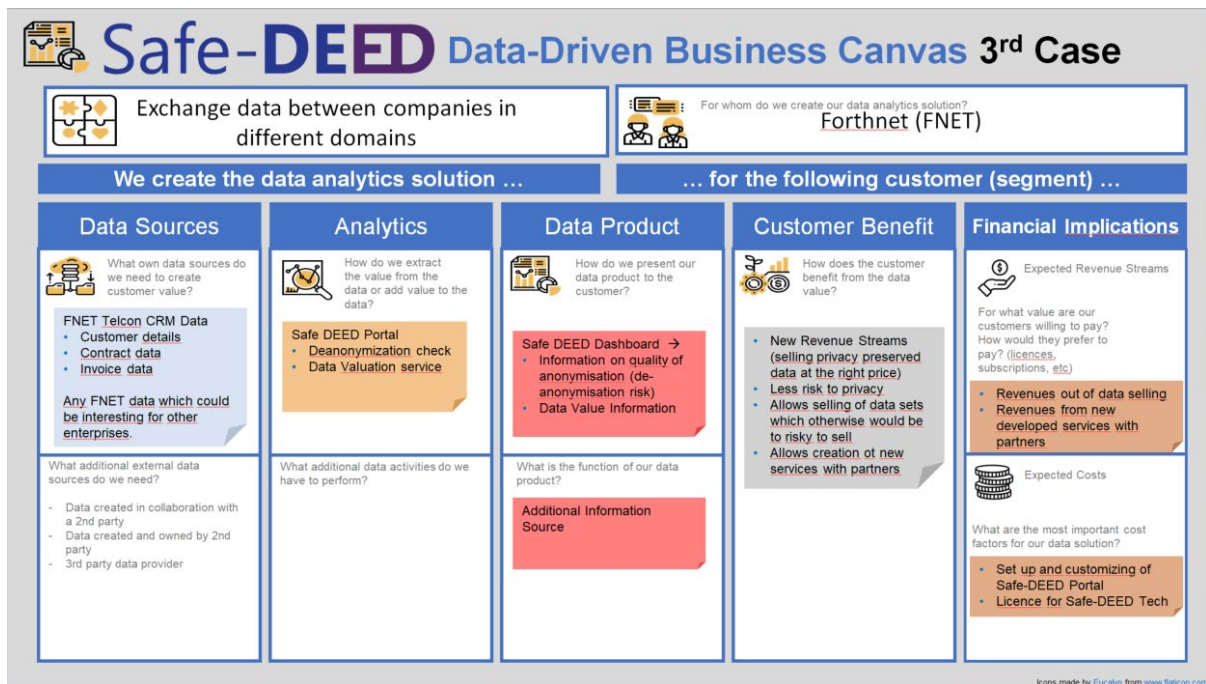


Figure 6: Selling data to firms in other industries

5 Conclusions

In this deliverable, we examined how Safe-DEED technologies enable new business models within the context of data marketplaces. This understanding is essential to lay a basis for assessing the economic impact the Safe-DEED brings to the data economy.

Since Safe-DEED technologies can be used in a wide variety of use contexts, and since there is hardly any related work on privacy-preservation and business models, we took a largely exploratory and qualitative approach. In Section 2, we provided a preliminary analysis of the business model implications of Safe-DEED technologies, based on desk research. Next, Section 3 presented three qualitative studies. We found the significant barriers and incentives for data sharing through a workshop. We examined the affordances of privacy-preserving technologies (as developed in WP5) through qualitative interviews with data marketplace operators. In the third study of Section 3, we examined the business model implications of privacy-preserving technologies in data marketplaces. We found new value propositions (i.e. privacy, control, no need for trust), new value capturing models (i.e. pay what you need, protect the buyer) and three new value delivery architectures (peer-to-peer, intermediary, aggregator). We examined the business model for each of these value delivery architectures. In Section 4, we contextualized our findings for the specific use case of WP6 and used the business model tools from T2.3 to visualize three business model scenarios that Safe-DEED enables.

Given the new and largely uncharted area of privacy-preservation and business models in a data marketplace context, our main ambition is to lay a basis for further work. Our exploratory studies (Section 3) are complementary in terms of their focus and perspectives and provide a basis for a more systematic and hypothesis-testing approach, which will be done in T2.4. The contextualized business models of Section 4 provide a basis for illustrating and testing the business model tools from T2.3.

Finally, our understanding of business models will be used as input for the economic impact modelling, which is the final deliverable of T2.4.

References

- Archer, D. W., Bogdanov, D., Lindell, Y., Kamm, L., Nielsen, K., Pagter, J. I., ... & Wright, R. N. (2018). From keys to databases—real-world applications of secure multiparty computation. *The Computer Journal*, 61(12), 1749-1771.
- Bestavros, A., Lapets, A., & Varia, M. (2017). User-centric distributed solutions for privacy-preserving analytics. *Communications of the ACM*, 60(2), 37-39.
- Bogdanov, D., Jõemets, M., Siim, S., & Vaht, M. (2015, January). How the estonian tax and customs board evaluated a tax fraud detection system based on secure multiparty computation. In *International conference on financial cryptography and data security* (pp. 227-234). Springer, Berlin, Heidelberg.
- Bogetoft, P., Christensen, D. L., Damgård, I., Geisler, M., Jakobsen, T., Krøigaard, M., . . . Pagter, J. (2009). *Secure multiparty computation goes live*. Paper presented at the International Conference on Financial Cryptography and Data Security.
- Bouwman, H., de Vos, H., & Haaker, T. (Eds.). (2008). *Mobile service innovation and business models*. Springer Science & Business Media.
- Chesbrough, H. (2010). Business model innovation: opportunities and barriers. *Long range planning*, 43(2-3), 354-363.
- Choi, J. I., & Butler, K. R. (2019). Secure Multiparty Computation and Trusted Hardware: Examining Adoption Challenges and Opportunities. *Security Communication Networks*, 2019.
- Fricke, S. A., & Maksimov, Y. V. (2017). *Pricing of data products in data marketplaces*. Paper presented at the International Conference of Software Business.
- Fruhworth, M., Rachinger, M., & Prlja, E. (2020, January). Discovering Business Models of Data Marketplaces. In *Proceedings of the 53rd Hawaii International Conference on System Sciences*.
- Hemenway, B., Lu, S., Ostrovsky, R., & Welser IV, W. (2016, August). High-precision secure computation of satellite collision probabilities. In *International Conference on Security and Cryptography for Networks* (pp. 169-187). Springer, Cham.
- Koutroumpis, P., Leiponen, A., & Thomas, L. D. (2020). Markets for data. *Industrial and Corporate Change*, 29(3), 645-660.
- Richter, H., & Slowinski, P. R. (2019). The Data Sharing Economy: On the Emergence of New Intermediaries. *IIC-International Review of Intellectual Property Competition Law*, 50(1), 4-29.
- Roman, D., & Vu, K. (2018). *Enabling Data Markets Using Smart Contracts and Multi-party Computation*. Paper presented at the International Conference on Business Information Systems.
- Schomm, F., Stahl, F., & Vossen, G. (2013). Marketplaces for data: an initial survey. *ACM SIGMOD Record*, 42(1), 15–26. doi:10.1145/2481528.2481532
- Schreieck, M., Hein, A., Wiesche, M., & Krcmar, H. (2018). The challenge of governing digital platform ecosystems. In *Digital marketplaces unleashed* (pp. 527-538): Springer.
- Spiekermann, M. (2019). Data Marketplaces: Trends and Monetization of Data Goods. *Intereconomics*, 54(4), 208-216.
- Stahl, F., Schomm, F., Vossen, G., & Vomfell, L. (2016). A classification framework for data marketplaces. *Vietnam Journal of Computer Science*, 3(3), 137-143.
- van den Broek, T., & van Veenstra, A. F. (2018). Governance of big data collaborations: How to balance regulatory compliance and disruptive innovation. *Technological Forecasting & Social Change*, 129, 330-338.

Zhao, C., Zhao, S., Zhao, M., Chen, Z., Gao, C.-Z., Li, H., & Tan, Y.-a. (2019). Secure Multiparty Computation: Theory, practice and applications. *Information Sciences*, 476, 357-372.