

TOWARDS ECOSYSTEM VALUE: SUSTAINABLE DATA-DRIVEN BUSINESS MODELS IN THE CONTEXT OF EUROPE

May 2022

Lotte Sophie Olthof

University of Twente

MSc Business Administration

Technical University of Berlin

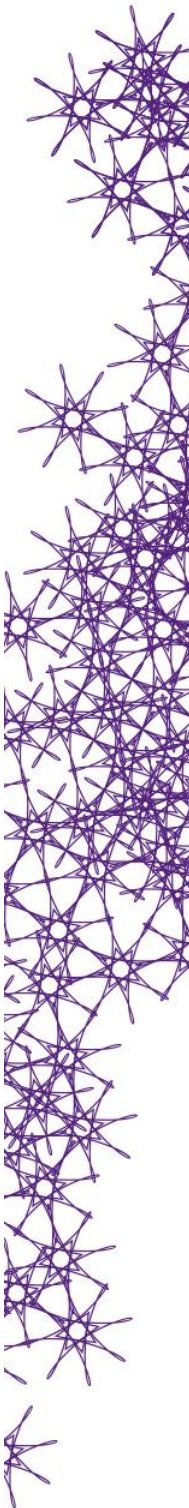
MSc Innovation Management, Entrepreneurship and Sustainability

Graduation committee:

Dr. T.G. Schweisfurth, Associate Professor of High-Tech Business, Faculty of Behavioural, Management and Social Sciences (BMS), Entrepreneurship, Technology, Management (ETM), University of Twente

Dr. M. de Visser, Assistant Professor, Faculty of Behavioural, Management and Social Sciences (BMS), Entrepreneurship, Technology, Management (ETM), University of Twente

Prof. Dr. Jan Kratzer, Chaired Professor for Entrepreneurship and Innovation Management and Founder/Academic Director of Centre for Entrepreneurship (CfE), Technical University of Berlin

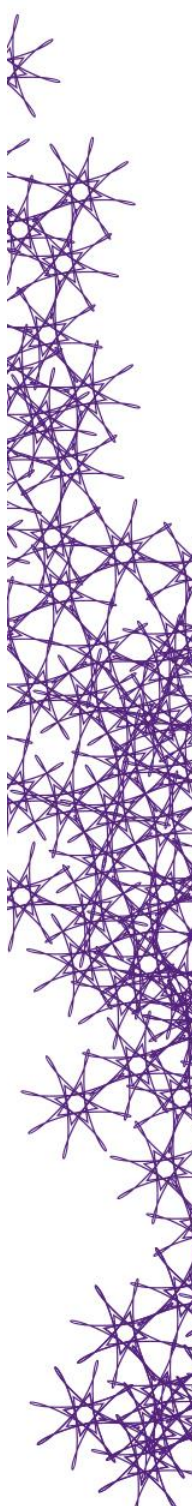


ACKNOWLEDGMENTS

With this thesis, I conclude my two-year double degree in Innovation Management, Entrepreneurship & Sustainability at the University of Twente, and the Technical University Berlin. I would like to express my gratitude to the interviewees that participated, colleagues, family, friends and teachers who helped me deal with challenges and supported me in achieving my goal of graduating.

ABSTRACT

The digital transformation challenges the sustainability of currently employed data-driven business models. Within Europe, new regulations are being developed to limit the ability of data-driven organizations to exploit personal data. Europe seeks to improve data sovereignty, for instance by enforcing mechanisms to store personal data in a decentralized manner under the user's control, or by prohibiting the processing of data unless consented by the user or data owner. With these developments, challenges occur in how to develop (socially desirable) business models. This explorative thesis aims to provide insight through semi-structured interviews with personal data operators, and experts in the field, to discover key challenges, and by extension to come up with alternative business model directions.



MANAGEMENT SUMMARY

There is an increasing trend in the use of personal data by businesses to improve or develop their services. The data owners (individuals or organizations) seem to be more than willing to share their data by using these services. Through commercial systems of data-driven organizations such as banks, energy providers, car rental companies and so on, lots of personal data is collected. In the current data ecosystems, the availability of personal data seems to be most beneficial for businesses, as they receive rich data. These data sets contain usually data such as a name, location data, age or usage data of a person. One of the biggest downfalls of the current situation is that there is an imbalance in power between organizations and individuals when it comes to pragmatic ownership over personal data. The EU has recognized this imbalance in power, and because of this is developing regulations to empower individuals over the flow of their data. The solution to empower individuals over the flow of data could be the personal data operator (also known among other concepts as trusted data intermediaries, trusted service providers or data intermediaries).

The personal data operator is characterized by placing the individual in the centre of their business model and enables the individual to take control over the flow of their data by providing a secure infrastructure for their data. However, it is still unclear what (socially desirable) business models will look like in an interoperable personal data ecosystem. This is needed because interoperable business models are an essential part of an interoperable ecosystem. This thesis aims to address this challenge by answering the following research question: *'How can personal data operators develop sustainable business models when data-driven business models are restricted in Europe?'.* The research question is tried to be answered by the use of semi-structured interviews with personal data operators, and experts in the field working at the Dutch Ministry of internal affairs, banks, and other private organizations.

Currently employed business models do not seem to be self-sustaining without the help of a larger commercial organization such as a bank or with the support of private investors that want to invest in services that guarantee privacy. There is a need for a new approach for their business models to place the individual end-user in the centre of their flow of data. The result of this thesis is an identification of key challenges and possible solution directions for sustainable business models. In conclusion, this thesis provides two alternative business model directions that need to be further researched.

The first alternative business model is based on a multi-sided platform approach. The idea behind this is that this alternative can function in the current personal data market, which is characterized by fragmentation. The second alternative needs to be further examined and is based on the service delivery network approach, which enables personal data to flow in an interoperable global network, wherein the data owner has full control over the flow of their data.

Table of Contents

1. Introduction	6
1.1. Context of data-driven organisations within Europe	6
1.2. Organizations of interest in this thesis: personal data operators	7
1.3. Ecosystem perspective	7
1.4.1. Different ways to organize personal data infrastructures	8
1.4. Trust frameworks	9
1.5. Business models of personal data operators	10
1.6. Central research question	10
2. Theoretical framework	10
2.1. The business model concept	11
2.1.1. Literature on business model patterns	11
2.1.2. Life cycle stages of a business	12
2.1.3. Consensus throughout different literature streams	12
2.2. Businesses that rely on data as a key resource: data-driven business models	13
2.2.1. Data-driven business model innovation	15
2.3. Transformation patterns of data-driven business models towards service-oriented models	15
2.4. Multi-sided platform business models	16
2.4.1 Revenue models for platform owners	17
2.5. Alternative business models operating in service delivery networks	18
3. Research method	18
3.1. Qualitative research design based on the Gioia methodology	19
3.2. Selection procedure of the sample	19
3.3. Data collection procedure through an online search, and by extension through semi-structured interviews	21
3.4. Data analysing method	23
3.5. Methodological limitations and ethical considerations	24
4. Results	24
4.1. Currently employed business models by personal data operators	24
4.2. Factors that impact the sustainability of currently employed business models	29
4.2.1. Results of part two of the semi-structured interviews with personal data operators	29
4.2.2. Analysis of key aspects of sustainable business models	31
4.3. Alternative business models for a sustainable solution	31
4.3.1. Multi-sided platform approach	32
4.3.2. Service delivery network approach	32
5. Theoretical and practical contribution	34

5.1. Theoretical contribution	34
5.2. Practical contribution	35
6. Conclusion and discussion	36
6.1. Limitations and further research	36
Reference list	38
Appendices	41
Appendix 1 – Interview questions personal data operators	41
Appendix 2 – Insight from conversations with experts in the field	43
Appendix 3 - workshop ‘Stress testing operator business models’ during the MyData in Netherlands 17-18 November 2021	44
Appendix 4 - examples of first and second-order codes	48

1. Introduction

As a result of the digital transformation, businesses are enabled to offer new or additional services based on data to generate their profits. The consequence of these technological developments is that there is a need to rethink traditional ways of creating value. There is a need to develop new business models that are suitable in the connected world of the data economy.

To provide answers on how to develop these data-driven business models, this explorative thesis contributes to management literature by developing alternative business models for data-driven organizations in the context of personal data management, which are sustainable in the European context. These organizations are data-driven in the way that they use (personal) data as a key resource.

1.1. Context of data-driven organisations within Europe

Currently, commercial systems of data-driven organizations provide highly centralized digital services. These services are often provided for free. In order to use these services, people generally need to exchange their data. This way of working means that organizations receive rich data sets on how users behave when using their apps. A detrimental property of this situation is an apparent imbalance in power between organizations and individuals when it comes to pragmatic ownership over personal data. Whereas organizations have the power to gather, trade, and access personal data, individuals seem to be subordinate. Individuals are commonly subjected to the preferences of the company gathering and processing their data. These companies typically operate on a profit motive and are commonly not transparent to the user (Cavoukian, 2013).

The EU has recognized this imbalance in power and because of this, has launched the 'European data strategy', which addresses challenges that occur in the establishment of a data-driven society. In the European context, this means the goal is to create a secure single market for data, intending to positively contribute to society as a whole. The expectation is that individuals and businesses are enabled to make more informed decisions based on the free available data. An initiative that resulted from this is the ePrivacy Directive, which is an initiative from 2009 that obligates businesses to notify data breaches. Another initiative is the General Data Protection Regulation (GDPR), which focuses in particular on data protection laws, and by extension presses fines for organizations that violate the policy, with penalties concerning potentially millions of euros (GDPR, n.d.). For example, article five describes the need for data minimization, and the requirement to process personal data securely by using appropriate technical measures. Similarly, the revised European Identification, Authentication and Trust Services (eIDAS) regulation is being developed. This regulation aims to enable citizens and businesses to have safer, faster, and more efficient data interaction (European Commission, n.d.). The main method set forth to attain this goal is to enable its citizens to share personal documents and identify themselves easily and securely by using an app on their phones. The proposed eIDAS revision is to address unmet market needs related to digital identity interoperability. These needs are currently partially addressed by single-sign-on solutions such as the OAuth/SAML integration provided by Facebook and Google. However, these solutions lack a strong level of assurance and cannot secure the privacy of the identity of the end-user. The revision of eIDAS aims to create a system that has the same convenient properties, but serves as a stronger proof of identity. A consequence of these regulations is that the ability of data-driven organizations to exploit data is limited. They seek to improve data sovereignty, for instance by enforcing

mechanisms to store personal data in a decentralized manner under the user's control, or by prohibiting the processing of data unless consented by the user or data owner.

1.2. Organizations of interest in this thesis: personal data operators

The EU has proposed the Data Governance Act in 2020, and adoption is planned for 2022. The act enforces a governance mechanism and a framework to facilitate (sector-specific) data sharing. The act provides strict guidelines for organizations that provide infrastructure or data storage for personal data. Such organizations have to assist individuals as described under Regulation (EU) 2016/679 (GDPR), for instance with:

- consent management to data processing,
- right of access to their data, and right 'to be forgotten',
- right to the rectification of inaccurate personal data, and restrict processing and the data portability right, which allows data subjects to move their data from one controller to the other,
- personal data storage space (without transmitting it to third parties).

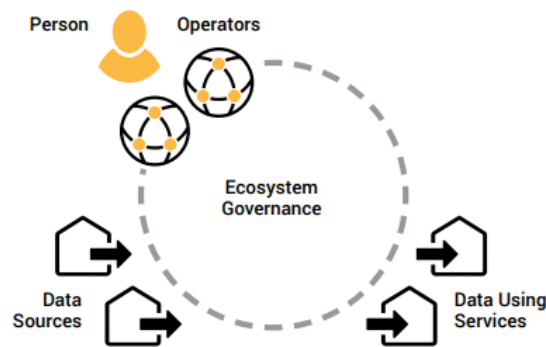
In this thesis, these organizations are named 'personal data operators' (also known among other concepts as trusted data intermediaries, trusted service providers or data intermediaries). The personal data operators provide individuals to access, use and manage personal data with a human-centric approach, which means the owner of the data needs to be empowered to take control over how their personal information is gathered and shared (Bharosa, Luitjens, Wijk & Pardo, 2018). To do this, personal data operators take either the role of an infrastructure facilitator, or provide related services, technologies, or tools concerning personal data (Zegher, Lähteenoja, Lindén & Grundström, 2020). Personal data operators function as an intermediary between data sources and data using services and provide their services within a trust framework (Poikola, Langford, Huhtamäki, Sierla & Janssen, 2019). The personal data operators provide a way for the individual to use their 'attributes'. Attributes can be described as a digital statement of an individual being someone or having something. A derived attribute can be used by an individual to verify that he or she is older than 18 years, without communicating what age they are. This allows the individual to use a service while minimizing the data they share with the service. For example: instead of sharing a birthdate to prove to be older than 18, a simple 'yes' or 'no' answer would suffice. The type of functionalities that personal data operators offer varies. Some organizations only pass on the data without storing it, whilst others function as a vault by storing personal data in a mobile app. The personal data operator concept places the individual in the centre by providing granular control over the flow of their data (MyData Global Network, 2017). This supports the goal of the EU for individuals to get granular control over their data, meaning that they can specify in great detail which attributes are shared with whom and for how long.

1.3. Ecosystem perspective

The personal data operator is part of an ecosystem consisting of different actors and roles. The different roles that can be identified are the individual, data source, data using service, and personal data operator. In more detail, the individuals use personal data operators' services to exchange their data between different organizations. The data source is responsible for the collection, and processing of personal data in their databases. For example, a governmental database in which names and addresses are processed. The data using service (f.i. a mortgage

advisor, or a landlord) can get the individual's authorization to collect and use personal data from a data source. The personal data operator takes the role of facilitator that enables secure access to use and share their data between a data source and data using service under the control of the individual. In addition to these four roles, Poikola et al. (2019) describe the need for the role of Ecosystem Governance to develop a mature personal data ecosystem. In figure 1, an overview of the roles within a personal data ecosystem is provided.

Figure 1 Personal data ecosystem roles and ecosystem Governance



(Poikola et al., 2019)

Jong, Duits & Rikken (2020) have recognized different streams of revenue that are generated within the personal data ecosystem, some examples are:

- Data using service pays directly for the personal data (often a subscription with a combination of fixed costs and variable costs depending on usage).
- Pay-per-use model (pay per request of needed information. (Implemented by Qii)
- Pay-per-user model (pay per unique user, implemented by Itsme)
- Data source pays the data using service (implemented by IRMA)

1.4.1. Different ways to organize personal data infrastructures

Within the ecosystem, the personal data operators are not the goal in themselves but a means to achieve a secure and safe infrastructure for the exchange of data within the ecosystem. This ecosystem can be organized in different manners. Langford, Poikola, Janssen, Lähteenoja & Rikken (2020) have described four likely scenarios of how the personal data market can be organized, namely:

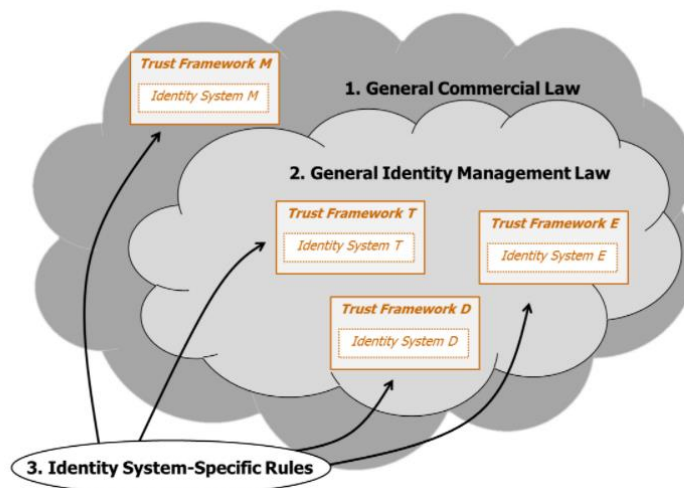
1. Fragmented, with many small-sized organizations that function without the establishment of interoperability between their services.
2. Monopolistic data platforms, where few larger organizations provide connectivity within the ecosystem, and without the incentive to enable interoperability.
3. Fully decentralised, which operates in a peer-to-peer approach with standardized infrastructure such as blockchain technology. The end-user directly manages the flow of their data towards the data using service, or by using a personal cloud-based application that can be hosted on their device or by a third party.
4. Competition-based interoperable operator network, wherein many organizations are competing with each other on an interoperable network, and together provide global-level connectivity for the end-user. This scenario is often compared to the contemporary network of banks, and telecom operators.

The first two scenarios are not desired to achieve an interoperable network for the exchange of personal data but do describe the current scenario of the personal data ecosystem, and help to design a market towards more desirable scenarios. Whereas the fully decentralized scenario is already more suited, this is not yet possible with the current technological solutions. To develop a global-level network for the exchange of personal data, the competition-based interoperable operator network would be most ideal and realistic to achieve this goal in the coming years. In such an ecosystem, the value is created from network effects and reduced costs through collaboration, standardization, and risk sharing (Langford et al., 2020). The third and fourth scenarios can exist in co-existence so this should not be excluded to create an interoperable network.

1.4. Trust frameworks

In order to achieve a competition-based interoperable operator network, there is a need to develop governance on an ecosystem level, as described in section 1.3. This is needed to establish a safe and dependable infrastructure for the exchange of personal data by a set of rules, specifications and policies, and this concept is described as a 'Trust Framework'. Such a framework is to create functionality and trustworthiness among the participants in an ecosystem. In the current personal data landscape, trust and functionality are aimed to be achieved by Service Level Agreements or other agreements set between parties. However, these types of agreements between parties are proven difficult to scale, challenging from a liability perspective, and lack the transparency needed to establish trust between the participants (Makaay, Smedinghoff, Thibeau, 2017). Figure 2 provides an overview of how Trust Frameworks are depicted in the overall legal framework. Level 3 represents the Trust Framework, in which the personal data operators have the freedom to concrete and agree on system rules as long as they are contractually agreed on. The Trust Framework is useful as this complements already existing legal frameworks by addressing more specific issues in the context of identity systems and transactions. On the other hand, levels 1 and 2 are outside of the control of the personal data operators of an identity system and are determined by public law. Hence, the Trust Frameworks provide additional governance for specific challenges.

Figure 2 Trust Framework in the overall legal framework



(Makaay et al., 2017)

1.5. Business models of personal data operators

The current personal data ecosystem can be described as immature. That is why we still see many data point solutions. Instead of mature ecosystem-wide standards for process methods and technical standards such as can be seen in the aforementioned network of telecom operators. It is still unclear what (socially desirable) business models will look like in an interoperable personal data ecosystem. This is needed because interoperable business models are an essential part of an interoperable ecosystem. The goal of this thesis is to map how currently employed business models by personal data operators and to provide insight into alternative business models that can function in an interoperable personal data ecosystem. Literature on business models for personal data operators is limited to explorative research. There is a need to extend the literature by providing more insights into how business models of personal data operators function to find out what their business models could look like in an interoperable ecosystem.

1.6. Central research question

This thesis aims to achieve the research goals as set out in the introduction by answering the following research question: ***'How can personal data operators develop sustainable business models when data-driven business models are restricted in Europe?'***. To answer the central research question, the following sub-questions need to be answered:

- What business models are currently employed by personal data operators?
- Why are currently employed business models of personal data operators not sustainable?
- What alternative sustainable business model for personal data operators can be designed?

This thesis is structured as follows: in section 2 relevant literature on business models, and in particular for data-driven business models is described; in section 3 the research method is explained; in section 4 the results will be described on currently employed business models, and alternative business models will be discussed; in section 5 the practical and theoretical contribution will be further explained; in section 6 the answer to the central research question will be provided.

2. Theoretical framework

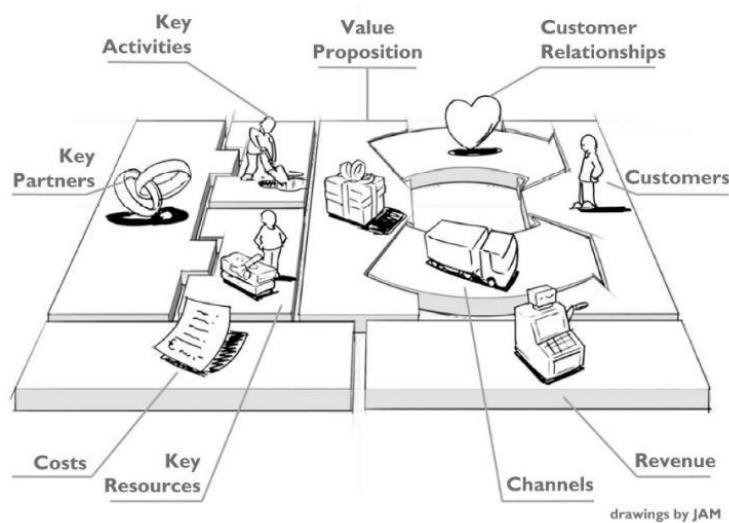
The goal of this chapter is to elaborate on the concept of data-driven business models by addressing theoretical concepts. The focus of this chapter is to elaborate on business models by looking at them beyond the traditional organizational borders, by including the network around and the value chain wherein the data-driven business models operate in, and finally analyse business models that place the end-user in the middle.

In section 2.1 the concept, patterns, and life cycle stages, and describes different literature streams on business models. The goal of this section is to provide background on the literature on business models. Subsequently, the following sections elaborate specifically on literature concerning data-driven models. Section 2.2 elaborates on models that rely on data as a key resource i.e. data-driven businesses. Section 2.3 provides insights into the transformation patterns of data-driven business model to provide a starting point for describing alternative data-driven business models. To provide insights on alternative data-driven business models in the context of the personal data operator, section 2.4 describes literature on multi-sided platform business models, and section 2.5 provides literature on service delivery network business models.

2.1. The business model concept

Different definitions of business models have been proposed throughout the literature. At a general level, the most well-known concept definition is described by Osterwalder & Pigneur (2010) as the blueprint for how they want to implement their overall strategy through systems, processes, and structure. The development of a business model can be used as a tool to implement the overall business strategy. In order to describe a business model, it is helpful to describe the different components upon which a business model is built. The process of describing a business model helps to improve the alignment of the business model's components to improve the overall value, and the assessment of a business model is an annually used tool by management to evaluate the health of the organization. Osterwalder & Pigneur (2010) have described nine components of a business model in the Business Model Canvas, which can be seen in figure 3.

Figure 3 Business Model Canvas



(Osterwalder & Pigneur, 2010)

The Business Model Canvas is used to assess a business model through the nine building blocks, which together reveal the logic behind the decisions an organization makes to be profitable. These building blocks cover four main areas, which are: customer, offering, infrastructure, and financial viability. These are the four key parameters used for describing and improving a business model. An important insight of the research of Osterwalder & Pigneur (2010) is that using a simplified model to develop a business model, such as the Business Model Canvas, helps to use a strategic and holistic approach to design a business model. The Business Model Canvas is often used as a template to assess the most vital components to developing a sustainable business model. Another important aspect is that this approach takes into account an analysis of the external environment of a business, and in particular the need for a business to enter key partnerships with their suppliers and competitors. Hence, the Business Model Canvas is the basis for mapping and developing alternative business models in this thesis.

2.1.1. Literature on business model patterns

In the research of Osterwalder and Pigneur (2010) five business model patterns have been identified, which help to understand important concepts in literature on business models, and contribute to a common understanding of business models. These business model patterns include Long Tail (selling less for a higher price), Multi-Sided Platforms (brings together two or more groups

of customers), FREE (one customer segment is free-of-charge), and Open Business models (systematically collaborating with partners outside of a firm). Business model patterns have been widely studied.

2.1.2. Life cycle stages of a business

The development from an innovative idea to a mature business can be described by the use of life cycle stages. This is helpful to analyse as it provides insight into the phase of an organization and by extension the challenges and opportunities. Lester, Parnell and Carraher (2003) came up with the following stages of organizational growth:

1. Existence (entrepreneurial phase, focuses on viability, identifying customers to support existence, decision-making and ownership by the entrepreneur(s), the environment is unanalysable. Within this stage, starting phase, creating/ finding their environment.)
2. Survival (seek for growth, develop some structure or formalization, develop distinctive competencies, organizational goals become more clear, the primary goal to generate enough revenue to continue and become financially independent and to stay competitive, important in this phase if the operations can be scaled, otherwise the organizations stays in this phase)
3. Success (often referred to as the 'red tape' phase, which means that formalization, mature, hierarchical reporting relationships and bureaucracy are the norm. Seek ways to protect gained success instead of finding new changes, and the environment is analysable)
4. Renewal (by collaboration and teamwork, innovations and creativity are needed, decision-making is decentralized, an organization is still bureaucratic and large, customer needs play a central role.)
5. Decline (characterized by power and politics, the organizational members are more concerned with personal goals, rather than what is best for the organization. Not able to meet external goals, experience lack of power/ loss in market share, and control and decision-making is centralized.)

These phases are used to describe the life cycle stages of the personal data operators in section 4.1 because it allows to describe the growth of different sizes of organizations, takes into account decline, which is useful to describe the phases of distinctive organizations.

2.1.3. Consensus throughout different literature streams

Zott, Amit and Massa (2011) have contributed to the literature on business models, by combining the insights from different research domains. This was needed because literature was mostly developing in silos and within different research domains, which are literature on e-businesses and the use of information technology, strategic issues such as value creation, and technology and innovation management. Their review revealed the lack of clarity in concept description used by different authors and concluded that these different topics can be described under the same umbrella as literature on business models. This multiple subject-matter lens of combining these three domains has led to more consensus in the research field about business models because Zott et al. (2011) have identified emerging themes that apply to the literature on business models in general. The following paragraph elaborates on these themes.

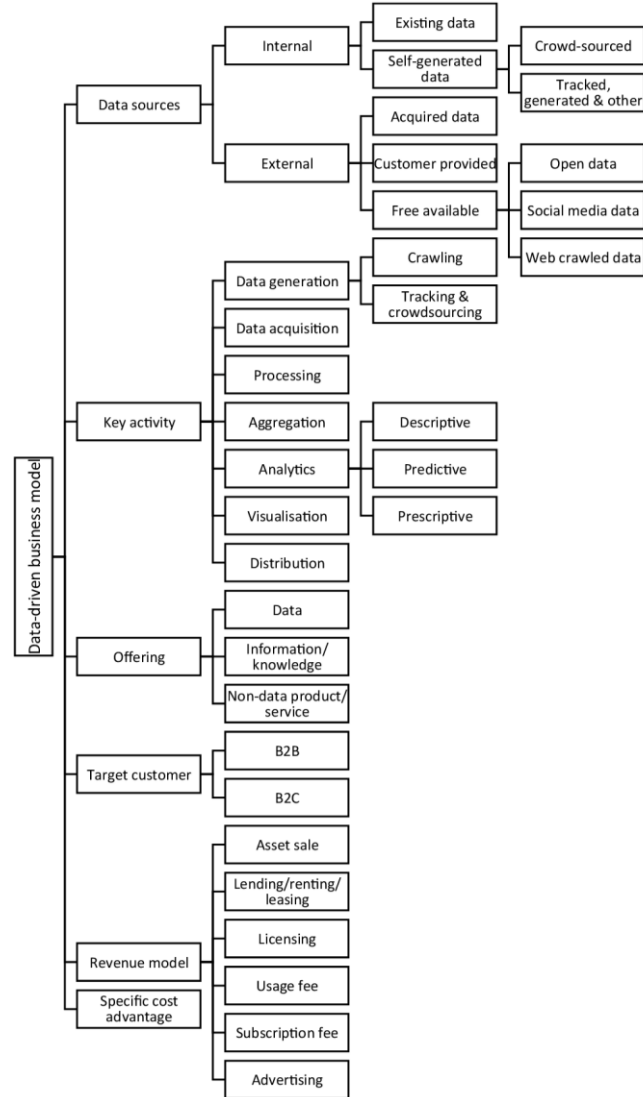
First, there is a consensus on the need of researching a business model as a unit of analysis distinctive from the traditionally used unit of analysis, which are the firm's output, the firm itself, the

industry it operates in, and the network around it. This helps to bridge the more traditional unit of analysis so the business model of the focal organization can be seen beyond the traditional organizational boundaries. Second, to describe how businesses function, there is a need to use a holistic and systematic perspective. This means that to provide a business model typology, there is a need to not only assess what products and services the business provides, but also how this is marketed to customers. Third, to conceptualize the business model of a specific firm, the activities of both the firm and its partners need to be taken into account. Throughout literature, this is described as an 'activity system perspective'. Fourth, the business model is used to explain both the value creation and capture.

2.2. Businesses that rely on data as a key resource: data-driven business models

Within this thesis, the focus lies on business models that support organizations that capture value by providing data-driven services. Hence, this section, and the following sections elaborate specifically on the characteristics of data-driven business models. Hartmann, Zaki, Feldmann, and Neely (2016) use the term data-driven business models for organizations that rely on data as a key resource to be operational. To provide more insights into data-driven business models, Hartmann et al. (2016) have developed a taxonomy (framework) that provides a set of data-driven business model attributes for the individual business model aspects and helps as a starting point to describe how business models of data-driven organizations can be mapped. This taxonomy has been added in figure 4 and is useful as a starting point to analyse data-driven business models.

Figure 4 Data-driven business model taxonomy



(Hartmann, et al., 2016)

Susha, Flipsen, Agahari and Reuver (2020) have recognized that data-driven business models can be described by using two discriminating variables, which are: the degree to which the data is available or restricted, and the value these organizations add to the data, to help describe the personal data operators. Noteworthy is the 'Data controls model' described by Susha et al. (2020) because this description overlaps with the organizations that are of interest within this thesis. This model describes that personal data operators that fit this model, add value from their data expertise and technical excellence of their product, and seem to both depend on their standardization and interoperability efforts, and buy-in from their key stakeholders. For example, organizations that fit this description are AMdEX and Civity, as they provide a data structure to enable the individual (referred to as data owner) to control the data they share and with what party. The findings of their research are:

- The difference in the field they target (cross-sector/ sector-specific)
- Issue of balancing between competition, innovation and openness (trust)
- Both create value by offering a trusted data infrastructure
- The business model of AMdEX is still developing/ different for CIP as a platform provider
- Issues of standardization and interoperability

2.2.1. Data-driven business model innovation

To gain more insight into important themes that impact how data-driven business models can be developed, Ciasullo, Montera & Romeo (2021) have carried out a systematic literature review on conceptual and empirical peer-reviewed articles and conference proceedings concerning data-driven business models. Their review shows that even though the growing interest of academia in data-driven business models, this literature stream can still be described as emergent as it contains mostly research focused on conceptualising the main characteristics of data-driven models and understanding them as a phenomenon. Instead of providing theoretical and practical guidance on the process of the development and the transformation from an existing into a data-driven business model. Additionally, literature on data-driven business models can be described as fragmented in distinctive disciplines. There is a need to overcome these separate disciplinary silos. Also, most literature investigates data-driven business models that focus on the perspective of a traditional business, without taking into account the dependencies of other organizations i.e. focal firm perspective. These models need to be examined through a holistic view. Another important view is addressing ethical issues when developing data-driven business models, as current literature on data-driven business models because unethical use of big data is occurring.

Additionally, their research showed that all of the studies classify the development of a data-driven business model into technical (data standards, formats), organisational (objectives, strategies, structure) and financial (resources needed to process, manage and maintain data) dimensions. An overview of the benefits and barriers has been added in table 1, aggregated from the article of Ciasullo et al. (2021).

Table 1: Overview of benefits and barriers collected from Ciasullo et al. (2021)		
Dimensions	Benefits	Barriers
Technical	Availability of data with the right standards and formats, high-performance data analyses	Data security, data license, data privacy, data quality
Organisational	Increased competitive advantage, change in the value proposition	Changes in human resources' skill sets and technical infrastructure need for collaborations with partners having specialised knowledge
Financial	New market segments, improved decision-making process, new ways of cooperation, better use of resources, reduction of costs, higher productivity Increased revenues	High investments in physical and infrastructural resources, extended payback

On a more practical level, more research is needed since many businesses recognize the value of data, and are willing to invest in exploiting data. However, many of these businesses seem to be unable to turn data into value, and by extension an increase in profits. Ciasullo et al. (2021) describe this as the "limbo stage". The use of business model innovation is acknowledged as a way to develop stronger businesses and provides a competitive advantage (Bashir & Verma, 2017).

2.3. Transformation patterns of data-driven business models towards service-oriented models

Enabled by the developments in the field of information technology, opportunities for alternative data-driven business models arise. The alternative business models presented in this thesis are based on the service-oriented business models to enable an interoperable global network. Tax, McCutcheon, and Wilkinson (2013) have defined this concept as: "two or more organizations that, in the eyes of the customer, are responsible for the provision of a connected, overall service"(p. 455). Zolnowski, Christiansen and Gudat (2016) used a multiple case study method to come up with

transformation patterns from which two enable a shift from a product to a service-oriented business model, which are:

- (1) cooperative value innovation (customers and partners are actively integrated as a co-creator and co-producer, strong impact of data)
- (2) customer-centric value innovation (generation of new services of development of a new market)

In order to develop alternative service-oriented business models that function in an interoperable network, it is useful to take into account both of these transformation patterns as it is used for telecommunication services (Zolnowski et al., 2016). In addition, the service business model is distinctive from the product business model, as it focuses on the services as a process between the interaction of different actors, who all add value to the service in their way. This means that the end-product is a service that can only be provided to the end-user with the cooperation and resource integration of the actors together. Hence, the service business model approach is in particular useful to come up with alternative data-driven business models as this allows to look at the creation process in a chain of actors that co-create to add value for the end-user. In order to analyse data-driven business models, Zolnowski et al. (2016) used an adoption of the Business Model Canvas mentioned in 2.1, namely the Service Business Model Canvas as can be seen in figure 5. The Service Business Model Canvas focuses on a service-oriented approach by taking into account co-creation and the chain of resource integration. This model provides insights into important aspects that need to be taken into account to develop alternative business models.

Figure 5 Service Business Model Canvas

Customer perspective	Customer (Customers in the business model)						
	(Costs borne by customers)	(Resources provided by customers)	(Activities carried out by customers)	(Value proposition for customers)	(Contribution of customers to maintain the relationship)	(Channels provided by customers)	(Revenues captured by customers)
Company perspective	Cost Structure	Key Resources	Key Activities	Value Proposition	Relationship	Channels	Revenue Streams
	(Costs borne by the focal company)	(Resources provided by the focal company)	(Activities carried out by the focal company)	(Value propositions of the focal company)	(Contribution of the focal company to maintain the relationship)	(Channels provided by the focal company)	(Revenues captured by the focal company)
Partner perspective	(Costs borne by partners)	(Resources provided by partners)	(Activities carried out by partners)	(Value propositions for partners)	(Contribution of partners to maintain the relationship)	(Channels provided by partners)	(Revenues captured by partners)
	Key Partner (Partners in the business model)						

Zolnowski et al. (2016)

2.4. Multi-sided platform business models

Within the market for the exchange of personal data, a multi-sided platform approach enables new opportunities. Value is created by providing a secure infrastructure, which simplifies the process for organizations that use such a platform. An example of such a multi-sided model is a platform ecosystem, which is characterized by the convergence of different organizations in one place, to provide a service together. Sorri, Seppänen & Valkokari (2019) have identified eight key characteristics useful to develop an ecosystem platform business model:

- Value, monetizing, producers, users, filtering, governance, resilience, network effect

These characteristics can be defined by answering guiding questions by the platform owner, which can be seen in table 2. The process of answering these questions helps to explore business model

innovation for an organization that aims to operate as a platform ecosystem by involving the other ecosystem participants (Sorri et al., 2019).

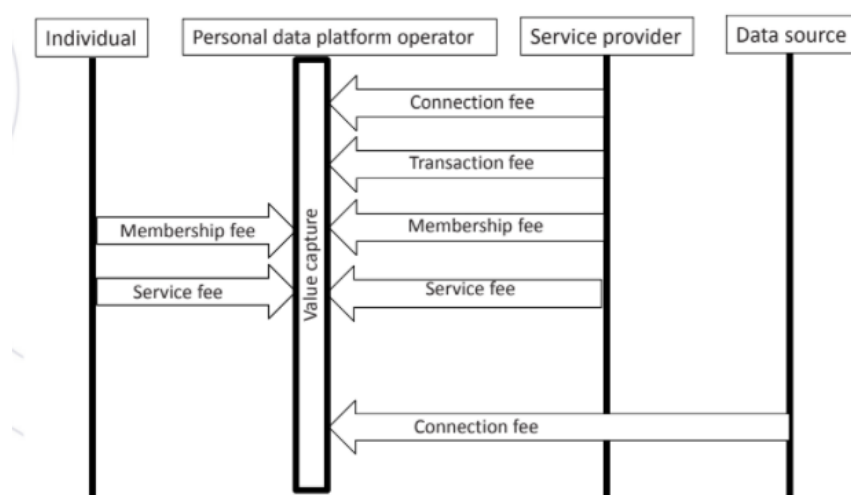
Table 2 Guiding questions to develop a platform in an ecosystem Sorri et al., 2019

Characteristics	Questions
Value producers	Who are the value producers and what motivates them to create the value? Through which channels do they produce the value?
Value users	Who are the value users, and what motivates them to consume the value? Through which channels do they consume the value?
Value	What are the different values that are created? How does the platform attract participants? How is the chicken-and-egg problem solved? Which friction does the platform reduce?
Filters	What data are acquired to match producer and user? Which filters does the platform need to serve the relevant content to consumers and connect them to the relevant value producer?
Network effects	Which types of network effects are achieved?
Value capture	What currency does the user provide to the producer in exchange for value? How does the platform capture some portion of this currency?
Governance	What are the tools for lowering the barriers to entering the platform? Which creation/curation/ customization/ consumption tools does the platform provide?
Resilience	To what extent are the boundary resources defined?

2.4.1 Revenue models for platform owners

In order to develop a better understanding of possible alternative data-driven business models, the following section elaborates on the revenue streams used by platform operators in the context of personal data. Kemppainen, Koivumäki, Pikkarainen and Poikola (2018) examined emerging revenue streams used by platform operators that facilitate the exchange of personal data from a person towards a data using service. Hence, these platforms can be described as multi-sided, and often this means that there is a group of users that pays in combination with a group of users that is subsidized, similar to the business structure of a large commercial organization such as LinkedIn. In figure 6, the revenue models for a platform approach are presented. This model is used in section 4.3 to develop an alternative sustainable business model.

Figure 6 Revenue models and the key stakeholders of a personal data platform operator



(Kemppainen et al., 2018)

Kemppainen et al. (2018) emphasize the need for two important requirements for business models to align with the EU strategy to enable interoperability and to increase the empowerment over control of the individual, which are no-advertising and free-for-users models. This means that

current platform models used by large commercial organizations cannot fulfil, and alternative business model solutions are needed. In their research, one of the organizations used a revenue-sharing model. For example, the personal data platform operator charges a data using service with a transaction fee, which is a percentage of the value of the transaction. These alternative models would be beneficial if the market stays fragmented. The personal data operator becomes more self-sustaining. An example of this approach is that of the organization Visions, a France personal data operator that aims to build a platform business model.

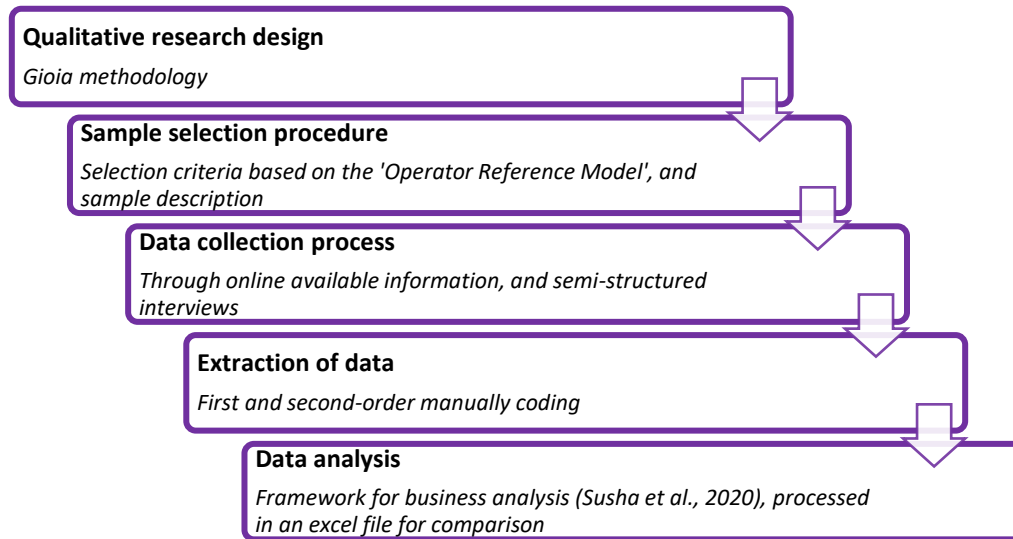
2.5. Alternative business models operating in service delivery networks

In order to develop alternative service-oriented business models mentioned in the prior section, there is a need to develop a better understanding of challenges and opportunities in the development process. To do so, the following paragraph elaborates on the findings of Pikkarainen, Koivumäki and Iivari (2020), who have explored the possibilities for business models that operate in a service delivery network such as health insurance organizations and other actors in the network (data using services, personal data operators). First, on an operational level, the transformation of a business model is a challenging process, and for now impossible due to concerns about responsibility taken by each organization from the viewpoint of the individual end-user, and a lack of trust among the actors involved in the network. Pikkarainen et al. (2020) emphasize the need for experimentation through case studies to build trust and find robust solutions for the exchange of personal data. Second, to increase trust in the overall services, there is a need to communicate with stakeholders about what measures have been taken to handle protection issues. Third, the service delivery network and customer-journey starting viewpoint helps to understand challenges and opportunities to develop sustainable alternative data-driven business models and creates new opportunities to develop new types of services together with the other roles in the personal data ecosystem. Fourth, Pikkarainen et al. (2020) emphasize the need for further development of data regulation and legislation regarding the collaboration of public and private sectors.

3. Research method

In order to answer the research question as described in chapter one, a qualitative research design will be used. The following sections aim to describe the research method used in this thesis, to enable reproduction. In section 3.1. the qualitative research design of Gioia is explained. Section 3.2. describes the selection procedure of the sample. Section 3.3. describes the data collection procedure. In section 3.4. the data analysing method is explained. Lastly, section 3.5. elaborates on the methodological limitations of the research design, and an explanation is provided on how ethical aspects of the data collection have been taken into account. An overview of the research design is schematically shown in figure 7 on the following page.

Figure 7 Overview of research method



3.1. Qualitative research design based on the Gioia methodology

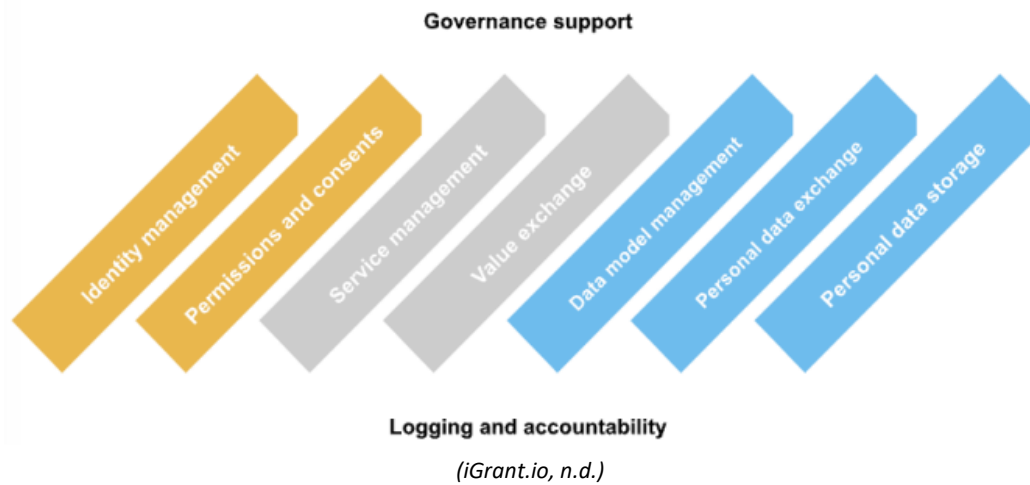
The methodological approach that is used in this thesis to explore alternative sustainable data-driven business models is qualitative. A qualitative approach is suited to answer 'why' and 'how' research questions, as used for the central and sub-research questions. For this thesis, the research method is based on the Gioia methodology. This method is suited to gather in-depth information and to critically compare different perspectives of interviewees in a small sample size. This research method can be described as inductive, which means that the data is first collected and after that compared to theory (Gehman et al., 2018). Within the Gioia methodology, a suited method of collection of data is through conversations (McLeod, 2019). For this thesis data was collected via semi-structured interviews. Therefore, the data that is obtained is non-numerical and focused on the subjective story as explained by each interviewee. An alternative qualitative research design that was considered at the beginning of this research is the Eisenhardt method, which is mainly used for case studies. The Gioia methodology has been chosen over the Eisenhardt method because this approach allows one to look beyond the borders of specific case studies, to gather enough information to develop alternative business models based on the views of the interviewees.

3.2. Selection procedure of the sample

The unit of analysis that has been studied in this thesis is the personal data operator in its context, as described in chapter one. Throughout literature, the personal data operator has been researched before, and mostly by the use of an explorative research design with small sample sizes (2-6 personal data operators) (Susha et al, 2020; Kemppainen et al., 2018) This approach allows to find key aspects, rather than finding generalize results to gain more insights in how these organizations function. This is in line with the research design used in this thesis because it aims to answer an explorative research question in the context of a personal data operator. The sample criteria in this thesis are similar to the MyData Operator reference model and is described below. This model describes nine functional core elements of the desired type of personal data operator as can be seen in figure 8. The MyData Operator reference model is developed by MyData Global, this organization is focused on practical characteristics of governance and technology for infrastructures regarding personal data sharing, use and management. MyData Global is an international non-profit

organization that works together with personal data operators to empower individuals by giving them the right to self-determination concerning personal data.

Figure 6 MyData Operator reference model



The core elements can be clustered into three groups, which are: 1. mediation of data transactions and permissions, 2. services and value exchange, and 3. data management and utilization. The purpose of this model is to create a common understanding and language to facilitate a human-centric personal data infrastructure that enables interoperability (iGrant.io, n.d.). This reference model is based on the study of a variety of functions of current personal data operators. The element 'value exchange' is particularly important in the context of this thesis because it describes how value can be generated by the operator. The value that personal data operators provide can be hard to quantify due to their position as continent-wide digital infrastructure. They create value for data using services by, for example, providing verified data, enhancing efficiency in administrative processes, and being outsourced to comply with the General Data Protection Regulation (GDPR). In addition, the personal data operators provide value to data sources by enabling them to comply with requests for data in line with GDPR. Personal data operators provide a user-friendly automated mechanism for fulfilling their obligation to comply with a request from an individual to receive their data. The individual receives value through the greater ease of access and selectiveness of sharing, giving them more options when it comes to handling their data, and this for better services without needing to manually collect and share their data with data using services. The operator reference model is used as a tool to assess personal data operators on their human-centric infrastructure approach for personal data management (iGrant.io, n.d.). MyData Global provides a list of personal data operators that meet the requirements set by the Operator reference model yearly. The majority of the sample size in this thesis is selected from this list. To provide insights into the sample of this thesis, the personal data operators, a short description of these businesses is provided in table 3 below. The sample includes a variety of different data-driven organizations that are active in the European market. This variation of the sample is needed to enable broad analysis.

Table 3: Description of the personal data operators' sample

Name	Business description	Interviewee	Duration (Min)
Ockto	A single-use platform that enables a person to collect their data from different data sources, and let the data flow towards a data using service. This is for specific services such as applying for a mortgage, renting a house, finance services, and so forth.	Product Director	69
Doccle	A reusable service platform, the individual end-user can make an account to distribute digital documents.	Executive	69
IRMA	Personal data operator and wallet provider. Self-sovereign identity platform. Work together with public governmental organizations.	Product owner	57
Own Your Data	Personal data operator and wallet provider. Focuses mainly only on healthcare-related services. Also, promote awareness of the need to manage personal information via the newsletter "Weekly Digest".	Product manager	45
Data Yogi	Personal data operator and wallet provider to provide the individual end-user with control over the flow of their data. Aims to co-operate with others such as suppliers.	Data Analytics Officer	37
Digi.me	Personal data operator and wallet provider. The individual end-user can choose the data storage place, which can be their OneDrive.	Business developer	78
Visions	Data ecosystem orchestrator to use by the individual end-user and businesses, an intermediary that provides services to share data. Their platform provides data sharing concerning skill and education. Visions is convinced of the need to separate the facilitator and wallet roles. Therefore, the organization cooperates with 'Cozy Cloud' as an independent wallet provider.	Co-founder	49
Schluss	Wallet provider is directly offered for the individual end-users, which they describe as the one-stop-shop for data management on demand of the person. Besides that, they are building a community together with businesses and governments. Schluss focuses on creating a community of actors that value sovereignty, privacy, and working 'open source'.	Co-founder	71
DataKeeper	Personal data operator and wallet provider for specific services such as car and house renting and field labs.	Sales, acquisition & strategy	74

3.3. Data collection procedure through an online search, and by extension through semi-structured interviews

The personal data operators have been studied by analysing their websites and other online available publications. The sample used in this thesis contains different types of organizations, and this desk research helped to become familiar with their specific services and products. To further analyse the personal data operators, semi-structured interviews were conducted. An overview of the interview questions is added in appendix 1. The semi-structured interviews have been divided into two parts as can be seen in table 4. The first part is based on the components of the Business Model Canvas from Osterwalder as described in chapter two.

Table 4: semi-structured interview questions preview based on the Business Model Canvas

Part one		
BM components	BM canvas blocks	Interview question
Customers (who)	Segments	Who are your customers? (What does your typical customer look like?)
	Relationships	How would you describe the relationship with your customers?
	Channels	What distribution channels are used to contact customers?
Offerings (what)	Value proposition	What bundles of services are you offering?
Infrastructure (how)	Key activities	What are the key activities to operate successfully? (core-business)
	Key partners	Who are your key partners? (describes the network of partners that make the business model work)
	Key resources	What key resources are needed to make the business model work?
Financial viability (why)	Costs	What are the biggest cost drivers inherent in your business model?
	Sources of revenue	From which organisations do you receive revenue streams?
Part two		
Discussion themes	Revision of eIDAS, API economy, monopolisation, inclusivity challenges, input from the interviewee	

The second part concerns discussions with the interviewees on certain topics to allow their input on themes that are expected to have an impact on the sustainability of the business model. To come up with these discussion themes, two steps have been undertaken. These two steps have been added to the appendix and function as support to answer the second sub-question. The first step was to gain insights through conversations held with experts from the field, working at the Dutch Ministry of internal affairs, banks, and other private organizations, in appendix 2. In the second step, these insights have been discussed in the workshop “Stress testing operator business models” during the MyData in Netherlands 17-18 November 2021¹, with a group of personal data operators, the results have been added in Appendix 3. The aforementioned steps have been used to come up with developments that are expected to have an impact on the sustainability of the business models, as input for part two of the semi-structured interviews held with the personal data operators.

The interviews have been both held with English, as well as Dutch-speaking interviewees. Therefore, the interview questions have been translated from English to Dutch. The interview questions were both closed- and open-ended. This approach provides flexibility for the collection of in-depth knowledge about specific contexts. The consequence of this was that the output of the interviewees varied (some were more focussed on the business model (Ockto, Irma, Digi.me, Doccle, Visions), and others were more focussed on the interviewee’s perspective of the market or from a research background (Own Your Data, Data Yogi, Schluss, Datakeeper). The interviews took place via a communication platform due to the COVID-19 pandemic and convenience because this allowed for an international sample. Before the interviews took place, the interviewees received the interview

¹ https://mydata.org/netherlands-2021-full-programme/?utm_campaign=InnoValor&utm_content=187531070&utm_medium=social&utm_source=twitter&hss_channel=tw-2260366520

questions to enable them to prepare for the interview. The interviews were audio-recorded and were worked out by two different approaches, some of the interviews have been transcribed, and others have been summarized per interview topic. The reason for this was that, that some of the interviews were more structured, following the topics of the interview question. Whereas the order of discussed topics of others was less structured. Although data collection through semi-structured interviews is not unobstructive, this approach enables one to find key challenges and come up with possible alternative business models.

3.4. Data analysing method

To analyse the semi-structured interviews, the audio records have been processed in a word processor program. The coding process contained two rounds, which helped to analyse the data by starting from one specific interview, towards checking each interview with theory and trying to come up with general conclusions. Following the Gioia method, this thesis used two rounds of coding. The first round contained open coding, to reduce the qualitative data to allow a more manageable focus (Hahn, 2008). This means that large quantities of data have been labelled without directly thinking about answering the research question. The benefit of this is to start with a broad analysis so informant-centric terms can be labelled in the form of comments in each of the transcripts. In his part of the process, the codes are used for entire paragraphs or sentences. Also, during this process, striking statements of interviewees have been highlighted in blue. After that, the goal of the second round of coding was to further filter the data that has been selected during the first round of coding. During this process, themes and categories have been identified that are expected to help to answer the research question. Examples of the two rounds of coding have been added in appendix 4. Besides that, to enable a comparison of the different interviews, the results have been placed in an excel file, with the above-mentioned dimensions. This helped to compare the different organizations systematically. The result of the general conclusions of the analysis of the currently employed business model is in section 4.1. provided. This two-step approach for coding helps to systematically analyse the transcripts (Gehman et al., 2018). If this thesis is redone, it will be expected to have similar outcomes broadly speaking because the results of the different interviewees were fairly similar, and matched with outcomes of other research. To enable a structured presentation of the results, the results have been mapped by using the four main areas of the Framework for Business Model Analysis, as adapted by Susha et al. (2020). The framework of operationalized business model elements can be seen in table 5.

Table 5 Framework for business analysis Susha et al., 2020	
BM Elements	Operationalization
Value Proposition	Business logic for creating value by offering products and services for targeted segments (data providers on one side and data users on the other)
Value Architecture	Architecture for the technological and organizational infrastructure used in the provisioning of products and services
Value Network	Collaboration and coordination with other organizations
Value Finance	Pricing and revenue breakdown associated with sustaining and improving the creation of value

This helps to categorize the business model of the personal data operator in a standardized manner, and has been chosen since it is specifically focused on analysing business models of organizations that rely on data as a key resource for business, which is in the direction of what is researched in this study.

3.5. Methodological limitations and ethical considerations

Within this thesis, the research method is explorative to discover key challenges to come up with alternative business models. In order to do so, the sample has been randomly selected and involves distinctive business types to gather a broad view by including different perspectives. The consequence of this is that the external validity is limited, which means for this thesis that the results solely provide insights into possible solution directions for alternative business models. Besides that, the expectation is that the ethical considerations for this research are expected to be minimal since it is focused on the exploration of business models. This is verified by requesting ethical permission from the research board of the University of Twente.

4. Results

Within this chapter, the results will be presented on currently employed business models, and alternative business models will be discussed based on semi-structured interviews with the personal data operators. The following sections provide answers to each sub-question as mentioned in section 1.6. In order to do so, section 4.1 is devoted to describing currently employed business models of personal data operators based on part one of interviews held with personal data operators. Subsequently, based on part two of the interviews, section 4.2. elaborates on the discussion themes based on part two of the semi-structured interviews, and provides an overview of key challenges that impact the sustainability of business models. Lastly, section 4.3. describes alternative data-driven business models that enable an interoperable global network, which are based on a platform and service-oriented approach. The findings of this thesis correspond largely to the findings of Sussha et al. (2020) and Pikkarainen et al. (2020) and aim to be novel by providing insights into the shortcomings of the sustainability of these business models, and based on these shortcomings, provide alternative data-driven business models for personal data operators.

4.1. Currently employed business models by personal data operators

The first sub-question concerns: 'What business models are currently employed by personal data operators?'. The answer to this question is that, in general, the personal data operators employ a typical traditional business model, which means in this thesis that these organizations employ a product-oriented business model. In table 6, an overview has been added of the different business model patterns that have been identified.

Table 6: Overview business model patterns personal data operators					
	Flat-rate	Pay-per-use	Freemium	Shop-in-shop	Subscription
Ockto	√	√		√	
Doccle		√			
IRMA	√	√			
Own Your Data					
Data Yogi			√		
Digi.me	√				
Visions			√		
Schluss					√
DataKeeper				√	√

The personal data operators have some significant differences when looking at their business model. The following section elaborates on the differences and similarities that stood out during the data analysis, and are discussed per business model component to provide a more in-depth analysis.

The role that the personal data operators take in the ecosystem is that of an infrastructure facilitator or wallet provider, or these two roles combined. Throughout the interviews, some of the personal data operators are convinced that their role should be neutral, which means that they should not provide both a vault and the infrastructure used to let the data flow from the data source to the data using service. For example, Digi.me describes the need for a 'neutral data facilitator', which means that the personal data operator should not provide the infrastructure and wallet simultaneously.

"The personal data operator must be neutral, so only the facilitator for the person. The person needs to have the freedom to retrieve information from different data sources and to transfer it under his direction, and to have it rotated under different data using services." Digi.me

When comparing current personal data operators, many seem to be specialized in a certain sector of data using services. Some of them mentioned focusing on a field such as finance, house rental, health care, human resources, e-commerce, or education. Some of the organisations offer customised packages of their offerings to a verifier. For example, Datakeeper offers customised data packages on request of a verifier, which they describe as the 'cafeteria model'. Others mentioned offering a generic service, cross-sector. The products and services that personal data operators offer vary. Some examples are:

- Digital document distribution
- Consent management
- Digitally signing
- Validate and place in the scheme (businesses)
- Service Level Agreement (businesses)
- Additional services: Customization of the interface (businesses)
- Personal data storage (wallet)
- Services conform to PSD2

There is a difference between the personal data operators when it comes to their legal structure. Some of these organizations have a limited liability company legal structure. They depend on large commercial organizations as their main revenue stream. Whereas others are a foundation or not-for-profit legal structure. The consequence of this is that these organizations have different financial goals and incentives.

"... because we have a non-profit motive, people are more willing to put in their time and energy."
Schluss

Another difference is the way personal data operators store the data differs. Some of them use a decentralized manner, others centralized. For example, the data is only temporarily saved by Ockto

when an individual requests data from the source and sends it to the data user service. The personal data is stored in an encrypted database. The data is erased within 90 days or anytime an individual requests this if the 'Ockto bron data service' is used. The data is secured with a code that is only known to the individual. On the other hand, an organization such as Own Your Data enables the individual end-user to download their information, and have the choice to store their data in a decentralized manner on a device of choice. Own Your Data uses a combination of blockchain and a central database to save costs. If the individual end-user wants to use the free service, it is centrally stored. Otherwise, this will cost extra. The option of centralized or decentralized data storage can both offer good options for security. Which one is better from a security perspective is a long-standing debate. In short, a central approach can be secured more easily (you have much more control over where something is stored and what measures you can take), but when things go wrong, all data is endangered. On the other hand, a decentralized approach is more difficult to secure properly, but the impact of an individual leak is much more limited. In this research, the organizations that most emphasize privacy by design and data sovereignty indicated that they had challenges improving their user experience, once they had a working proof of concept. A common hazard for uptake seems to be that developers with the most technical expertise tend to prefer the most expert-friendly interfaces, which might sacrifice a clear layout for a greater level of control. Table 7 provides an overview of the maturity and business domains for each personal data operator.

Table 7: Description of the personal data operators' sample		
Name	maturity	Business domain
Ockto	Phase three success (aim to protect gained, success, the environment is analysable)	Financial-sector
Doccle	Phase four renewal (gained success aim to find new markets/ in other countries)	Cross-sector
IRMA	Stage two survival (the next step is to go to the market)	Cross-sector
Own Your Data	Stage one: existence (publicly available, thousands of users in total)	Health and mobility sector
Data Yogi	Stage one: existence (test phase, not ready for the market)	Cross-sector
Digi.me	Stage two survival (goals are clear, developing structure, seeking for growth)	Cross-sector
Visions	Phase two survival (change of business model, develop different organizational structure)	Education
Schluss	Stage one: existence (beta version available)	Cross-sector
DataKeeper	Stage one: existence (immature, a project in a large organization)	Cross-sector

Value proposition

The value that is created by the personal data operator is currently mostly acknowledged by data using services. The reason for this is that these organizations expect that investments will lead to an optimization of internal processes and by extension ease customer processes. This is mainly because the personal data operator structures the data in a standardized manner, and the data can contain verified attributes from the original data source. On the other hand, it is less clear if and what the data sources experience as value. The personal data operators expect that these could save costs by replacing a desk employee if the process becomes optimized. This is due to the interference of the personal data operator, which is expected to lead to a simplification of the flow of data. For example, to automate the process to meet a subject access request (regulated by GDPR), which is the right of an individual to request that an organization discloses what personal data they have, and

how this is used in their processes. The individual uses these services in particular for ease of use. This means that the individual is only willing to use the service if this fits in their use-case journey.

"We don't think it makes sense at the moment, because the person isn't going to pay for a wallet. A person is willing to pay because they want to rent a car, right? A person is always in a use-case journey, so we are going to target the data using service who lets the person join, so we don't want to bother people there anymore." Datakeeper

Besides that, for society, personal data operators are a way to secure that personal data flows. These organizations enable an ethical way to share and control data. In addition, this information can be used on an aggregated level for analysis to enable better-informed decision-making.

Value architecture

The current infrastructure used by the personal data operator needs to be further developed. The current market is fragmented solutions, and there is a need to come up with Trust Frameworks to come up with agreements on technical standards. The data using services seem to be willing to invest in the implementation of the services of the personal data operator. This is needed because the implementation requires changes in the internal processes of the data using service. On the other side, the data sources are important for a personal data operator to provide its services. How the data flow from a data source toward the individual end-user can be improved. In many cases, the personal data operator has to gather the data for the person by using 'web scraping', which involves scraping data from a human-readable webpage (which you gain access to by having the user login first, ID), instead of automated data exchange via, for example, an API.

"It's still inconvenient, web scraping. Prefer to retrieve data via APIs so that there is 24/7 stability, but that is not there yet. It's still really early in that regard." Ockto

This is unstable because if the data source makes changes in the interface/ layout of the webpage, this automated data collecting does not work anymore. The relationship between the personal data operator and the data source needs to be professionalized by the establishment of the use of standard technical communication means such as via standardized APIs offered by the data source to enable a better data flow. In this case, the personal data operator creates a token to gather data from the data source. The data is solely saved in a way the individual prefers. For example on their smartphone in a data wallet of their choice. The consequence of this is that the personal data operator does not have the personal data themselves, the data can directly flow from the data source to the person. The personal data operators mentioned the need to develop a more stable network connection with the data source. Many of the personal data operators mentioned the need for a network wherein the data source offers an API to share the data with the personal data operator.

“If it has to be opened just like when it comes to the phone lines. KPN has a network, and years later other parties have been allowed to use the network. For example, a private network with an external registry with access to data sources. This could work, the personal data operator has its network with access to data sources.” Ockto

The current market is not developed enough to support this model on a large and interoperable scale. The growth of this market requires investments in better infrastructure.

Value Network

The personal data operator is dependent on other organizations to make their business model work. Most of the personal data operators describe the data using service as a key partner for commercializing their services. Mostly to start the use case because these organizations are willing to invest. The personal data operators cooperate in these cases with the data using service. Mostly because they value the data that is provided by the personal data operator. The relationship between the personal data operator and the data source seems to be less obvious and should be further investigated. Also, there is a need for personal data operators to co-create to develop a mature market.

“There is a trend of uniting forces, different organizations working together around standards, general governance, human-centric approach. We are going to collaborate as any growing industry will. If there are centralized standards, governance standards, that is the root to unite in a more organized manner.” Visions

Value finance

The financial viability among personal data operators is limited due to the immaturity of the market. The personal data operators seem to be aware of this and try to develop along with the market. The risk for personal data operators, for now, is to invest too early. Similarly, Yoon, Copeland and McNally (2018) emphasized the lack of organizational resources for personal data operators to either maintain their current services or develop new services. In particular, smaller-sized personal data operators seem not to have enough resources to maintain for years without having clear revenue streams. Therefore, the position of these organizations can be described as first movers or initial entrants. The consequence of this is that the majority of the personal data operators are not financially independent. During the interviews, the personal data operators explained that their financial resources mostly depend on a larger organization. For example, large banks that invest or are (partly) owners of a personal data operator. This means in the context of the personal data operators, that the data using services pay, and the individual end-user can benefit for free and without obligations. The pricing model that is used for data using services is mostly based on a:

- connection fee (depending on the number of connected data sources, and type of licenses),
- pay-per-transaction or pay-per-use.

For now, this type of business model seems to be able to cover the average costs of the individual end-user, and the data using service seem to be willing and able to pay. The personal data operator services are still developing, it is not logical to let this be subsidized by another actor within the personal data ecosystem. In this developing market, it is important to have a role within the

ecosystem that is paying for the establishment of the infrastructure to enable personal data operators to develop into mature organizations. The personal data operators do not seem to agree yet on the role of the data source when it comes to their pricing model. Some of the personal data operators mentioned that they are willing to pay the data source in return for an API or another standardized manner of data sharing. On the contrary, other personal data operators mentioned the possibility to have an issuer as a customer. An example of this is the approach used by IRMA. The data source pays IRMA a yearly fee. The reason for this is according to IRMA that there is a need to validate and add them to a scheme. Data sources want to pay for this because it reduces costs and simplification of customer processes.

“Data sources that want to allocate attributes within IRMA, must be validated and included in the scheme. IRMA charges an annual fee for this.” IRMA

Besides that, IRMA asks broker organisations for a flat fee. On the contrary, the data using service is enabled to download IRMA technology for free. The costs are for most of the personal data operators build up out of payments to software developers, hosting services, technical operational costs, and employee costs. The largest costs for personal data operators are for the technical expertise and for hosting.

4.2. Factors that impact the sustainability of currently employed business models

The following section elaborates on the second sub-question: ‘Why are currently employed business models of personal data operators not sustainable?’. The main reasons for this are the restrictions set by Europe concerning the exploitation of (personal) data, the risk of data using services or big tech players investing in similar services. On the other hand, the personal data operators seem to lack internal resources to keep improving and developing their services and the lack of trust among these organizations.

To provide a more in-depth analysis, the outcomes of part two of the semi-structured interviews, held with personal data operators, are explained in section 4.2.1, and to provide an overview, section 4.2.1. summarizes the key challenges that are expected to have the most impact on the sustainability of the business models.

4.2.1. Results of part two of the semi-structured interviews with personal data operators

The following section elaborates on challenges and opportunities derived from part two of the semi-structured interviews with personal data operators. The outcome in the following paragraph is described per theme, namely the consequence of the revision of the eIDAS, data using services that internally provide personal data operators, potential risk of a big tech player that takes over the market.

Consequences of the revision of the eIDAS regulation

The revised eIDAS regulation has direct consequences on the sustainability of business models employed by personal data operators. The consequences of this that are expected to have the largest impact are:

- not allowed to collect usage data/ use or exploit data for other purposes than agreed to,
- service needs to be free for the individual end-user,

- high-level of security assurance (for the enrolment of new users),
- public (and designated private) data using services need to accept the wallet.

The strategy of the EU governance is based on and dominated by a Self-Sovereign Identity approach. An example of this approach is moving away from the use of centralized wallets. According to IRMA, is a challenge to assess which level of reliability can be used for which attribute, and the fact that technology is accelerating with the consequence that law and regulations are still lagging. The majority of the personal data operators describe the revision of the eIDAS as a development in line with their business goals. For example, Ockto already fits the requirements (they are working on becoming a standard with IDIN because they are already used by the Dutch government). Whereas for others the revision of the eIDAS will require some attention when it comes to their business model. For example, an organization like Schluss needs to make changes in its pricing model as they aim to let the individual end-user pay to use the wallet. Own Your Data proposed the need for data points and governance, a systematic description, and a need for standardization in governance. Besides that, the status of the personal data ecosystem was often described by the interviewees as immature and experimental with similar proof of concepts or small-sized case studies. There is a need for attention of both the EU and national governments need to make decisions according to the personal data operators. The lack of this decision-making hampers companies from making big investments in the current market.

The data using service internally provides personal data operator related services

There is a trend among data using services (in particular banks and health-related businesses) to invest in personal data operator-related services or to offer it themselves. Ockto explains that these organizations seem to experience the value a personal data operator can add, and therefore is willing to invest. A negative consequence of this for the customer journey is that this had led to lots of separate mobile applications, which means that the individual end-users use every service with a separate solution, which is a data point solution. During the interviews, the personal data operators stated that the data using service should not take a personal data operator role, in particular, due to the restrictions under the revised eIDAS. There is a need for a personal data operator to increase trust by individual end-user, and the expectation is this is increased if the personal data operator is an independent entity from data using services and data sources. The personal data operators emphasize exceptions, which are data using services in the sector such as energy, banking, and the health industry. For other sectors, the expectation is that the personal data operator is needed because of their technical expertise.

Risk of big tech players taking over the market

A realistic threat mentioned by personal data operators is that big tech organizations will take over the market. This is realistic as these organizations can invest more because of economies of scale. The consequence of this is that the process is not interoperable, and becomes more complex for the individual end-user due to fragmentation as data using services each offer their solutions, which are not interoperable. This approach is described as the 'killer app strategy'. The consequence of this is that this stimulates the individual end-user to only use the services of one business, which leads to vendor lock-in. The challenge for personal data operators is to create trust and cooperate, this is needed to be able to compete with these big tech organizations such as Microsoft.

4.2.2. Analysis of key aspects of sustainable business models

This thesis aims to describe sustainable alternative business models that can function in a personal data market that is a combination of fully decentralised and competition-based interoperable operator network because this enables interoperability that is scalable on a global level. In order to achieve this, there is a need to take into account the following key aspects when designing sustainable business models:

- Trust among actors in the personal data ecosystem, need to develop Trust Frameworks
 - Technical solutions are not the main challenge
 - Need for agreements, technical standards, the collaboration between public/ private organizations
 - Validation (which level of reliability can be used for which attribute)
 - Decide what the price for data is
 - This needs to be further developed by organizations
- Expected consequences of eIDAS, moving away from centralized solutions, are:
 - not allowed to collect usage data/ use or exploit data for other purposes than agreed to
 - service needs to be free for the individual end-user
 - high-level of security assurance (for the enrolment of new users)
 - public (and designated private) data using services need to accept the wallet
- Both EU and national governments need to make decisions (lack of decision making hampers investments in current markets)
- Data using services invest internally in personal data operator solutions
 - Negative consequence for customer journey, data point solutions, fragmentation, for every service an app, not scalable
 - Lack of trust experienced by data owner, need for an independent entity from data using service, need for a personal data operator
- Risk of a big tech player that takes over the market
 - Economies of scale, personal data operators lack internal resources to compete (Scarcity of resources, for small-sized organizations it is challenging to invest to develop and improve their processes)
- Separation is needed for the role of infrastructure facilitator and wallet provider (important to have a data intermediary who stores the data, and another data intermediary who processes the data)

To overcome these challenges and enable interoperability, need for collective and open business models (Chesbrough, 2006) for personal data operators, to come up with sustainable solutions. The following selection elaborates on two alternatives have been mentioned that will be used in 4.3. to describe the alternative business models, based on a multi-sided platform and service-delivery approach.

4.3. Alternative business models for a sustainable solution

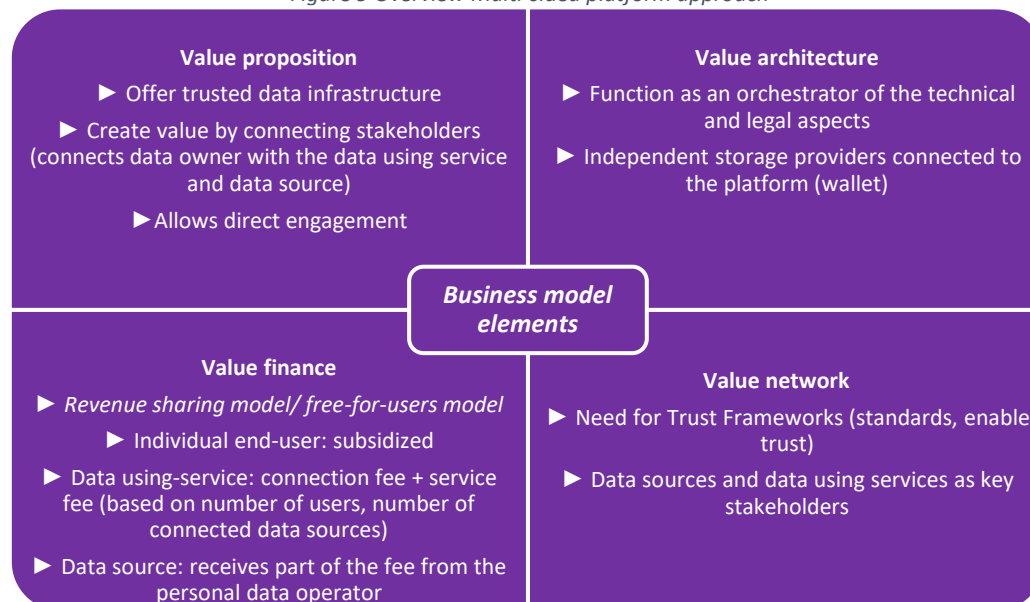
Within this section, an attempt has been made to come up with alternative data-driven business models, to answer the third sub-research question: 'What alternative sustainable business model for personal data operators can be designed?'. To answer this question, section 4.4.1. provides an alternative business model based on a platform approach that can function in the current market.

This current market is characterized by fragmented solutions, provided by small-sized organizations and market leaders that depend on data using services as financing party. In section 4.4.2. the aim is to describe an alternative data-driven business model based on the service delivery network approach that can lead to an interoperable network, and by extension can lead towards the ideal market described in section 1.4.1., a combination of the fully decentralised and competition-based interoperable operator network similar to the network used by telecom operators. For both of the alternatives, the focus lies on the transformation from a product to a service-oriented business model to enable the development of an interoperable global network. As described by Zolnowski et al. (2016), this transformation is possible through developments in the field of information technology and data availability.

4.3.1. Multi-sided platform approach

Within this section, the multi-sided platform business model is described to overcome current challenges for personal data operators in the current fragmented market such as scalability. To overcome this, the personal data operator can function as a multi-sided platform in which the data sources and data using services are connected at the request of the data owner. In particular, due to the developments in information technology, this type of linking service can be offered. The revenue models can be based on a multi-sided perspective where one group subsidizes the other group. For example, the data using service could be on the paying side and the individual end-user can use the services for free. For now, a risk is that a big tech organization is leading in developing such a platform, without a human-centric approach, so the data owner does not become empowered over the flow of their data between organizations. Below in figure 9, an overview is added with the business elements that are important when designing multi-sided platforms.

Figure 9 Overview Multi-sided platform approach



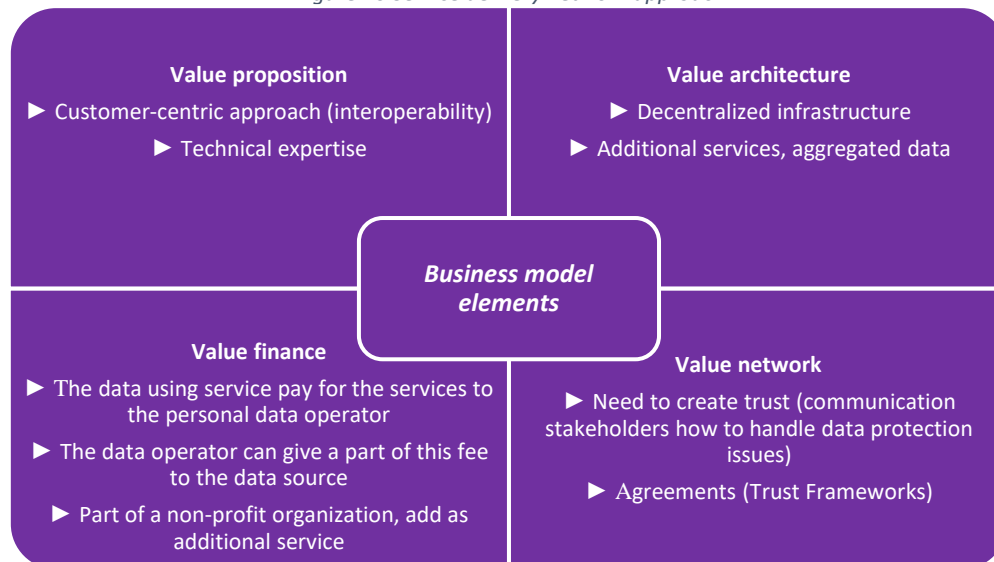
4.3.2. Service delivery network approach

The service delivery network approach helps to explore alternative business models that could function in a future personal market in which the network is interoperable and functions on a global level. The following section is based on the research of Pikkarainen et al. (2020). In their research, they emphasize the need for a service delivery network approach, which means that the individual

end-user can choose between different businesses to purchase a service, which prevents data lock-ins. In order to achieve this, the personal data operator needs to work together with the other actors in the ecosystem. The personal data operator creates value by providing a secured and transparent infrastructure that empowers the data provider (business/ individual) with granular control over their data. The value created by the data source, data using service, and the personal data operator together enables synergy. This means that the value for the individual end-user increases if these organizations cooperate. For the personal data operator to distribute their services, there is a need to enable a decentralized infrastructure. The idea behind this approach is that the person is placed in the centre, and thus the freedom to choose what organizations can use their data in the person's service journey. The personal data operator becomes responsible for securely connecting these organizations to let the data flow at the request of the individual end-user.

There is a need for the personal data operator to invest in collaborations with other organizations, to be able to provide their services. In a networked approach, the revenues can be broken down for different actors. Based on the results of the interviews, it could be possible to use a value chain perspective. The idea is to let the data using service pay for the services to the personal data operator. The data operator can give a part of this fee to the data source. This could be a way to create a self-sustaining network for personal data management. The personal data operator could use the pricing model that is currently used by some of them, which is a connection fee for organisations that want to join the network. Below in figure 10, an overview is added with the business elements that are important when developing service delivery network models.

Figure 10 Service delivery network approach



5. Theoretical and practical contribution

The following sections elaborate on how the outcomes of this thesis propose to contribute to business scholars and the professional field.

5.1. Theoretical contribution

This thesis aims to contribute to the academic community by providing insights into how sustainable business models can be developed for organizations that use data as a key resource i.e. data-driven businesses. Even though the impact of data on the sustainability of currently employed business models is widely acknowledged by academia, there is a need to examine specifically how data-driven business models can be developed. This is needed as literature on this topic is characterised as emerging since the majority of the literature is conceptual and focuses on understanding the phenomenon of data-driven business models, rather than providing empirical observations of these types of business models (Ciasullo et al., 2021). There is a need to extend the literature in the form of empirical research. This thesis hopes to contribute to the academic community by providing insights into how data-driven business models can be developed, taking into account the restrictions set by Europe regarding the use of (personal) data. This thesis uses an explorative research design in the form of semi-structured interviews to map currently employed business models and to gain insights into developments that impact the sustainability of data-driven business models. Therefore, this thesis aims to contribute to the literature by providing an overview of key aspects that need to be taken into account when designing sustainable data-driven business models in the context of a personal data operator to provide insights into possible sustainable alternative data-driven business models in the European context.

Findings of current business model theories cannot directly be translated to data-driven businesses that function in a personal data ecosystem (Kemppainen et al., 2018). Due to developments in the field of information technology, and new regulations developed by the EU regarding the exchange of personal data. The studies that take into account the specific context of the personal data operator are limited and have an explorative character. Therefore, this thesis aims to contribute to the academic community by providing insights into how data-driven organizations can employ sustainable business models for organizations that use personal data as a key resource. The alternative business models (multi-sided platform, service delivery networks) suggested by this thesis have not been tested and therefore need to be taken as suggestions rather than solutions that can be directly used as guidelines to develop data-driven business models in the context of a personal data operator. In addition, literature on business model innovation is characterized by its experimental and iterative character to develop a (new) business model. This is in line with the results of this thesis, as it shows that the currently employed business models are being developed iterative, and are developed with an experimental approach to come up with sustainable business models. Hence, this thesis aims to contribute to the literature by providing insights into key challenges for data-driven organizations in the context of the personal data operator in Europe, which can be used as a starting point to further research on how these business models can be developed.

An important aspect of researching business models of data-driven organizations, such as personal data operators, is the use of an ecosystem perspective. Prior literature has focussed on the perspective of the focal organization when researching data-driven business models (Ciasullo et al., 2021). This thesis aims to contribute to business scholars by taking into account an ecosystem

perspective by including the dependencies of stakeholders such as the individual end-user, the data using service, the data source, and policymakers. The alternative business models presented in this thesis both use an ecosystem-based approach, which allows cocreating the services offered by data-driven business models. This thesis aims to contribute by providing alternative business models that allow a holistic approach when designing business models.

Lastly, another important aspect of this study is to address ethical issues in the development of an alternative data-driven business model, by enabling the data owner with empowerment over the flow of their data within the personal data ecosystem. Since the market is developing, it is difficult to state what factors need to be taken into account. The implication of this thesis is to further research how a human-centric approach can be guaranteed in the design process of a business model.

5.2. Practical contribution

Even though the use of data as a key resource is nothing new for businesses, technological developments provide new ways of doing business, and this requires new types of business models to enable the added value of these developments.

This thesis aims to practically contribute by providing insight into how sustainable data-driven business models can be developed, in particular for organizations that handle personal data. This is needed because personal data operators are currently uncertain about their business models, and cannot operate independently with a human-centric approach. This is shown by a large number of initiatives and the limited number of successful case studies, let alone mature businesses. This challenge is becoming increasingly prominent due to trends such as increasing interest from big tech players and regulations regarding handling personal data. Even though this thesis looks specifically at the business models of personal data operators, other organizations may also benefit from these alternative business models due to the digital developments. In addition, the outcomes of this thesis show that the personal data operators have not yet found a best practice when it comes to the development of their business model. The consequence of this for practitioners is that there is a need to continuously monitor technological development and developments in the ecosystem, improve their data-driven services, and come up with alternative business model directions. This is a continuous process. Also, the personal data operators are faced with lots of uncertainties when it comes to legislation on a European and national level. Hence, this thesis aims to contribute to practice by providing insights into key challenges that need to be taken into account when developing a data-driven business model in the context of personal data operators. Hopefully, the analysis of key challenges provides insight for practitioners to develop sustainable data-driven business models.

For practitioners, it is important to come up with sustainable solutions that enable scalability in order to shift towards a mature personal data ecosystem. The importance of this lies in the shift from a product to a service-oriented business model to create an interoperable global network. The first alternative based on a multi-sided platform approach can help to overcome scalability issues in the current fragmented market. The second alternative based on the service delivery network approach needs to be further researched as this will allow creating an ideal market, a combination of the fully decentralised and competition-based interoperable operator network similar to the network used by telecom operators.

6. Conclusion and discussion

When it comes to market development, the personal data market still is underdeveloped and needs to become mature. The goal is set out by the EU. However, how this goal is reached seems to remain unclear. It was often mentioned that there is a need for more clarity on what is a suitable solution because for now, personal data operators are hesitant to make investments. A good start could be for governments to stimulate this market by opening up their data sets to enable a more stable connection with the personal data operators. The currently employed business models are not self-sustaining without the help of a larger commercial organization such as a bank or with the support of private investors that want to invest in services that guarantee privacy. There is a need for a new approach for their business models to place the individual end-user in the centre of their flow of data. This requires an EU that is willing to invest in these types of solutions, and policies and guidelines that force personal data operators to take a neutral role. The consequence of this is that the business models employed by business models need to fit into the personal data ecosystem to overcome interoperability challenges. As described in the results is this for now the greatest challenge to overcome. This means that personal data operators need to cooperate to establish the infrastructure and standards to enable a secure personal data ecosystem. This needs to be anchored in the business model. Therefore, the service delivery network approach to business models could help to fundamentally tackle the current challenges that personal data operators face.

The main research question of this study is: *'How can personal data operators make profitable business models when the exploitation of personal data is undesirable within Europe?'*. Even though this thesis does not provide deterministic answers to what type of business models are sustainable, the results provide insights into possible directions. The first alternative can be employed in the current market, which is the multi-sided platform business model to overcome the current, and a business model that functions in a service delivery network approach that can function in a further developed and ideal market.

6.1. Limitations and further research

The limitation of this thesis is the lack of relevant literature on data-driven business models that take into account the current EU strategy regarding personal data, and place the data owner in control over the flow of their data. The majority of literature on data-driven business models places the organization in the middle, instead of using a human-centric approach. Further research is needed on how a data-driven business model can be developed using a human-centric approach. Another limitation is the small sample size used in this thesis. Future research should involve larger sample sizes to come up with generalizable results to provide more insights into what aspects impact the sustainability of data-driven business models. Besides that, the Business Model Canvas was used to analyse the currently employed business model. However, in the completion stage of this thesis, the Service Business Model Canvas was discovered. The expectation is that this model would have been more suited to analyse the business models of personal data operators as it focuses on the external network of an organization to make the business model work. Lastly, more research is needed to test the potential of the alternative business models presented in this thesis. For the first alternative, a multi-sided platform, there is a need to test the feasibility with a large sample size. There are some examples of organizations that aim to develop such a model. If more research test the feasibility, results can be validated. For the second alternative, service delivery networks, there is a need to further research important characteristics, as this type of business model is new for data-driven

business models in the context of personal data. Therefore, further research can map important factors to implement such a model.

Reference list

- Bashir, M. & Verma, R. (2017). Why business model innovation is the new competitive advantage. *IUP Journal of Business Strategy*, 2017, 14(1)7
- Bharosa, N., Luitjens, S., Van Wijk, R., & Pardo, T. (2018, May). Panel: Removing the barriers for personal data management. In Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age (pp. 1-3). doi: 10.1145/3209281.3209327
- Cavoukian, A. (2013). Personal Data Ecosystem (PDE)—A Privacy by Design Approach to an Individual's Pursuit of Radical Control. *Digital enlightenment yearbook 2013: The value of personal data*, 89-101. doi:10.3233/978-1-61499-295-0-89
- Chesbrough, H., Ahern, S., Finn, M., & Guerraz, S. (2006). Business models for technology in the developing world: The role of non-governmental organizations. *California management review*, 48(3), 48-61. doi: 10.2307/41166349
- European Commission (n.d.). Discover eIDAS. Retrieved from <https://digital-strategy.ec.europa.eu/en/policies/discover-eidas>
- Gehman, J., Glaser, V. L., Eisenhardt, K. M., Gioia, D., Langley, A., & Corley, K. G. (2018). Finding theory–method fit: A comparison of three qualitative approaches to theory building. *Journal of Management Inquiry*, 27(3), 284-300
- Hahn, C. (2008). *Doing qualitative research using your computer: A practical guide*. Sage. Retrieved from https://books.google.nl/books?hl=nl&lr=&id=ier-g1UBEe4C&oi=fnd&pg=PP7&dq=Doing+Qualitative+Research+Using+Your+Computer&ots=1gOFntyQ Sq&sig=dWYOHgMW2mzoTesT53Yh1IPLwc&redir_esc=y#v=onepage&q=Doing%20Qualitative%20Research%20Using%20Your%20Computer&f=false
- Hartmann, P. M., Zaki, M., Feldmann, N., & Neely, A. (2016). Capturing value from big data—a taxonomy of data-driven business models used by start-up firms. *International Journal of Operations & Production Management*
- iGrant.io (n.d.). MyData Operator reference model. Retrieved from <https://docs.igrant.io/docs/mdo-ref-model/>
- Jong, J., Duits, I., Rikken, M. (2020). PDM landschap 2020. Regie op gegevens in Nederland. Retrieved from <https://www.rijksoverheid.nl/documenten/rapporten/2021/01/11/pdm-landschap-2020-regie-op-gegevens-in-nederland#:~:text=Onderzoek%20naar%20de%20mogelijkheden%20die,PDM%20staat%20voor%20persoonlijk%20datamanagement.>
- Kemppainen, L., Koivumäki, T., Pikkariainen, M., & Poikola, A. (2018). Emerging revenue models for personal data platform operators: when individuals are in control of their data. *Journal of Business Models*, 6(3), 79-105

- Langford, J., Poikola, A., Janssen, W., Lähteenoja, V., & Rikken, M. (2020). Understanding MyData Operators. MyData Global. <https://mydata.org/wpcontent/uploads/sites/5/2020/04/Understanding-Mydata-Operators-pages.pdf>
- Lester, D. L., Parnell, J. A., & Carraher, S. (2003). Organizational life cycle: A five-stage empirical scale. The international journal of organizational analysis. Chicago. retrieved from https://libres.uncg.edu/ir/uncp/f/Organizational%20Life%20Cycle_A%20Five%20Stage%20Empirical%20Scale.pdf
- Makaay, E., Smedinghoff, T., & Thibeau, D. (2017). OpenID Exchange: Trust Frameworks for Identity Systems. Retrieved from <https://openididentityexchange.org/networks/87/item.html?id=175>
- Vincenza Ciasullo, M., Montera, R., & Romeo, E. (2021). What about Data-Driven Business Models? Mapping the Literature and Scoping Future Avenues. International Journal of Business and Management 16(8)1. doi: 10.5539/ijbm.v16n8p1
- McLeod, S. (2019). What's the difference between qualitative and quantitative research? Retrieved 2021, November 4th, from <https://www.simplypsychology.org/qualitative-quantitative.html#:~:text=Qualitative%20research%20is%20the%20process,meaning%20to%20their%20social%20reality>
- MyData Global Network (2017) 'Declaration of MyData Principles'. MyData Global Network (before the MyData Global association was established). Retrieved from <https://mydata.org/declaration>
- Osterwalder, A., & Pigneur, Y. (2010). Business model generation: a handbook for visionaries, game changers, and challengers (Vol. 1). John Wiley & Sons
- Pikkarainen, M., Koivumäki, T., & Iivari, M. (2020). Seizing the business opportunities of MyData service delivery network: Transforming the Business Models for Health Insurance Companies. Journal of Business Models, 8(2), 39-56. doi: 10.5278/ojs.jbm.v8i2.3322
- Poikola, Langford, Huhtamäki, Sierla and Janssen (2019). Discussion Paper What is the MyData Operator? Retrieved from <https://mydata.org/wp-content/uploads/sites/5/2019/09/Discussion-paper-MyData-operator-final.pdf>
- Sorri, K., Seppänen, M., Still, K., & Valkokari, K. (2019). Business model innovation with platform canvas. Journal of Business Models, 7(2), 1-13
- Susha, I., Flipsen, M., Agahari, W., & Reuver, M. D. (2020, August). Towards generic business models of intermediaries in data collaboratives: From gatekeeping to data control. In International Conference on Electronic Government (pp. 304-315). Springer, Cham.
- Tax, S. S., McCutcheon, D., & Wilkinson, I. F. (2013). The service delivery network (SDN) a customer-centric perspective of the customer journey. Journal of Service Research, 16(4), 454-470. doi: 10.1177/1094670513481108
- Zegher, M. D., Lähteenoja, M. A., M., Lindén, F., & Grundstrom, C. (2020) MyData: Applying human-centric

principles to health data. Medical Writing. 29(2)

Zolnowski, A., Christiansen, T., & Gudat, J. (2016). Business Model Transformation Patterns of Data-Driven Innovations. In ECIS (Vol. 2016, p. 146)

Zott, C., Amit, R., & Massa, L. (2011). The business model: recent developments and future research. *Journal of Management*, 37(4), 1019-1042. doi: 10.1177/0149206311406265

Appendices

Appendix 1 – Interview questions personal data operators

The results of this interview will be used as input for my master thesis for the study of Business Administration at the University of Twente. The goal of this research is to examine what operators' business models look like and how developments can impact this. This interview is divided into three parts. The first part will start with general questions. The questions of the second part are based on the components of the business model canvas from Osterwalder and Pigneur. This will help to systematically analyze the various business models by having standardized questions. The third part describes developments based on conversations held with experts from the work field, working at the Dutch internal affairs, banks and other private organizations. Several developments have been mapped that are expected to have an impact on the operators' business model in the coming 5 years. The questions go into whether you expect these developments to impact your organizations' business model. The duration of the interview will be 60 minutes, recorded (audio only) and transcribed. The transcripts will be shared afterwards to enable you to review the transcript. The results of this research will become publicly available.

Thank you for participating in this research!

General questions

- In what year is your organization established?
- How many employees does your company have?
- In which industry do you now supply your service?
- What is the legal structure of your organization?
- What is the total number of unique users?
 - And the number of recurring users?

Part one - Business Model Canvas

The following questions will be asked in the context of your role within the personal data ecosystem

The roles that are often identified throughout the literature are the person, data source, data using service and operator/ wallet.

- *The person is an individual that uses services to exchange their data between different organizations*
- *The data source is responsible for the collection and process of personal data in their databases. (Governmental database wherein names and addresses of individuals are processed)*
- *Data using service can get authorization to collect and use personal data from a data source (mortgage advisor, landlord).*
- *The operator takes the role of a facilitator who enables secure access for individuals to use and manage their data and the outflow of data between a data source and data using service.*

Specific questions regarding the business model

Who – customers

- Who are your customers? (What does your typical customer look like?)
- What role(s) fulfil(s) your customers within the personal data ecosystem?
- How would you describe the relationship with your customers?

What - proposition

- What bundles of services are you offering?
- Does your organization offer different services to different groups of customers?
- Which problem are you helping to solve with your service?
 - Which customer needs are you satisfying?

and value

- What values are you creating for other roles?
 - What value do you receive from other roles?

How - activities

- What are the key activities to operate successfully? (core-business)

Key resources

- What key resources are needed?

channels

- What distribution channels are used to contact customers?
 - How do you reach the roles that pay you?
- Is the personal data stored in a central database?
 - If so, does your organization have read access to the data?
- How much of your codebase is open source, and how much is closed source?
- What programming language is the back-end logic mostly written in? (Java, Go, Python, ...)

and partners

- Who are your key partners? (describes the network of partners that make the business model work)
- Which key resources are you acquiring from partners?
- Which key activities do your partners perform? (For example, outsourcing It, Sales, Marketing, ...)
- Who are your key suppliers?

Why - costs

- What are the biggest cost drivers inherent in your business model?
- Which key resources are most expensive?
- Which key activities are most expensive?

and revenues

- From which roles do you receive revenue streams?
 - Besides the roles of the ecosystem, are there other organizations that provide value streams? (e.g. from subsidies from the government/ sponsoring from public organizations)
- For what value are the roles willing to pay?
- How much does each revenue stream contribute to overall revenues?
- What type of fees does your organization ask or pay to/ for each role?
 - (For instance: Onboarding fee, Account fee, Roaming fee, Sales commission, Transaction fee, Connection fee)
- How would you describe your business model design?
- Follow-up question; Could another design work for your organization, why?

Part two - Discussion questions

- Do you want to share your vision on how the following developments could impact the business model?
 - *Revision of the eIDAS puts limits on wallets by the European law*
 - *European digital identity wallets > Citizens can identify themselves worldwide and can identify and share their attributes.*
 - *Requirements: free of charge, wallets need eIDAS high safety, public and private service providers need to accept wallets, provider wallets not allowed to gather usage data*
 - *Big tech player takes over the market*
 - *Big tech companies already investing and trying. Monopolistic data platforms*
 - *Data using services will invest in this to optimize their internal services.*
 - *Data using services offer operators' services to their customers*
 - *The personal data operator is no longer needed for data using services*
 - *Development X?! – What are your thoughts on other developments with impact?*

Appendix 2 – Insight from conversations with experts in the field

In the starting phase of this thesis, several experts have been asked about their views on developments that impact the sustainability of business models of personal data operators. The result of this is the three developments described below. These developments have been chosen because they are expected to impact the operators' business model in the coming 3 – 5 years.

D1 - Revision of the eIDAS puts limits on wallets by the European law

The first development is the consequence the revision of the eIDAS could have for personal data operators within Europe. This is a framework for secure electronic identification and verification across Europe. The consequence of the eIDAS is that data owners can safely share their data/ digital documents between two EU countries is safe, legally compliant, and regulated. Also, the European digital identity wallets > Citizens can identify him or themselves worldwide and can identify and share their attributes. Requirements i.e.: free of charge, wallets need eIDAS high safety, public and private service providers need to accept wallets, provider wallets are not allowed to gather usage data.

D2 – Data using services offer operators' services to their customers

The data using services start to invest in the infrastructure that is needed to offer operators' tasks to its customers. Will there still be enough work for the operator? In which market? Data using services will invest in this to optimize their internal services.

- invest in their infrastructure
- optimization of internal processes
- what does this mean for standards?
- competition (neutral) operators
- limited choice customer?
- examples; Datakeeper Rabobank/ Ockto ABN AMRO

D3 – Big tech player takes over the market

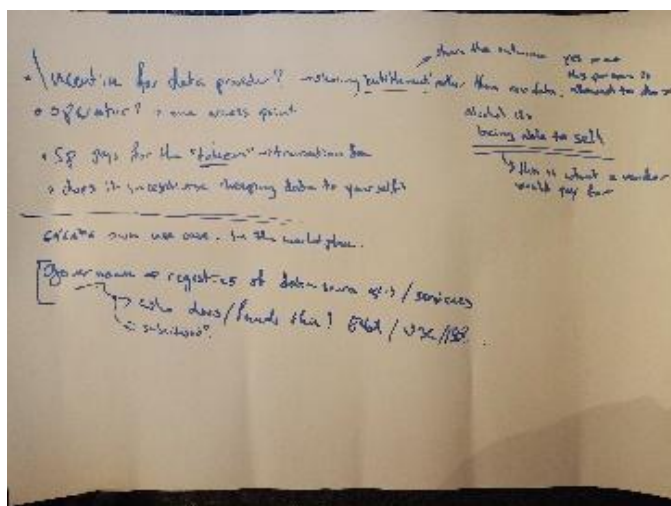
Big tech companies already investing and trying. Monopolistic data platforms (few platforms provide data sharing and connectivity within their ecosystem. There is no incentive for interoperability between the platforms). For example, Microsoft/ Samsung

- there is little to no competition
- the incentive for interoperability is missing
- risk of vendor lock-in
- charging a higher or lower price compared to competitive market

Appendix 3 - workshop 'Stress testing operator business models' during the MyData in Netherlands 17-18 November 2021²

Overall conclusions

- The incentive for data providers? > entity check > share the outcome > yes or no this person is allowed to do X
 - Alcohol is 18+ being able to sell > this is what an organization would be willing to pay for
- Operator? > one ace point
- Service provide pays for the 'token' > transaction fee
- Does it incentivise keeping data to yourself?
- Create your own use case in the marketplace
- Governance > registration of data sources APIs/ services
 - Who does/ fund this? EU? W3c/ ISO
 - Subsidies?
- One of the operators looks at the entire value stream within the ecosystem. Need to look at an ecosystem business model canvas
- Need to distinct sectors, across-sectoral
- (Led MyData4Pandemics). To see business aspects: www.mydata.org/pandemics
- Provides data storage platform. The individual decides whether third parties are allowed to receive personal data. Make also use of APIs
- The individual can use the service for free. Organizations always need to pay a fee for the membership



² https://mydata.org/netherlands-2021-full-programme/?utm_campaign=InnoValor&utm_content=187531070&utm_medium=social&utm_source=twitter&hss_channel=tw-2260366520

	Entity...	development 1	development 2	development 3
who customers	Individuals / privacy aware	Big Tech moving in	standards &	Digital ID
what proposition and value	Companies: Reestablish Trust	<ul style="list-style-type: none"> • ease of use • Price • small deployment • long term security • Price effective 		
how activities and partners	Service Providers: Creating service services	<ul style="list-style-type: none"> • fancy easy UX / Potential big market • costly to market outside Big Tech 		
why costs and revenues	Data Vault			
	GeNode: white label cloud with collaboration			
	SDK: develop 3rd party service on infra	<ul style="list-style-type: none"> • big development budget / money at hand 		
	Providers			
	Growing infra structure with 'generic' services	<ul style="list-style-type: none"> • fancy services • changes customer expectation 		
	NPO with strict governance			
	Companies: reselling, creating services			
	Fee for infrastructure, maintain, stake-of-the-infrastructure			
	White label fees (volume, transaction, users)			
	Add-on services fees (free → premium)			

Based on Business model stress testing: A practical approach to test the robustness of a business model from Timber Haaker, Harry Bouwman, Wil Janssen, Mark de Reuver from 2017

Business model stresstest		name		
		development 1	development 2	development 3
who customers	LEARNERS	Skills Data Space (fancy) PP		
what proposition and value	EMPLOYERS Enter Hard Hatters / train TRAINING COPS. APPS + APPS + APPS MORE CHOICES, 2nd. MORE RECOG, STATS MORE AGENCY FIND MATCHING PROFILES FIND TRAINED PROFILES Scale up + more off lower stamp. relationship + APPS	More traction multimedia + demand + demand		
how activities and partners	Infrastructure + GOV AZ services, apps	Standardization, Cloud in Europe + + + Specific #2. Models (exp. transparency account, treated #2) => TRUST		
why costs and revenues	infra + GOV + HL + change → more to do → gain time efficiency → reputation → gain visibility → time 10x → less			

Based on Business model stress testing: A practical approach to test the robustness of a business model from Timber Haaker, Harry Bouwman, Wil Janssen, Mark de Reuver from 2017

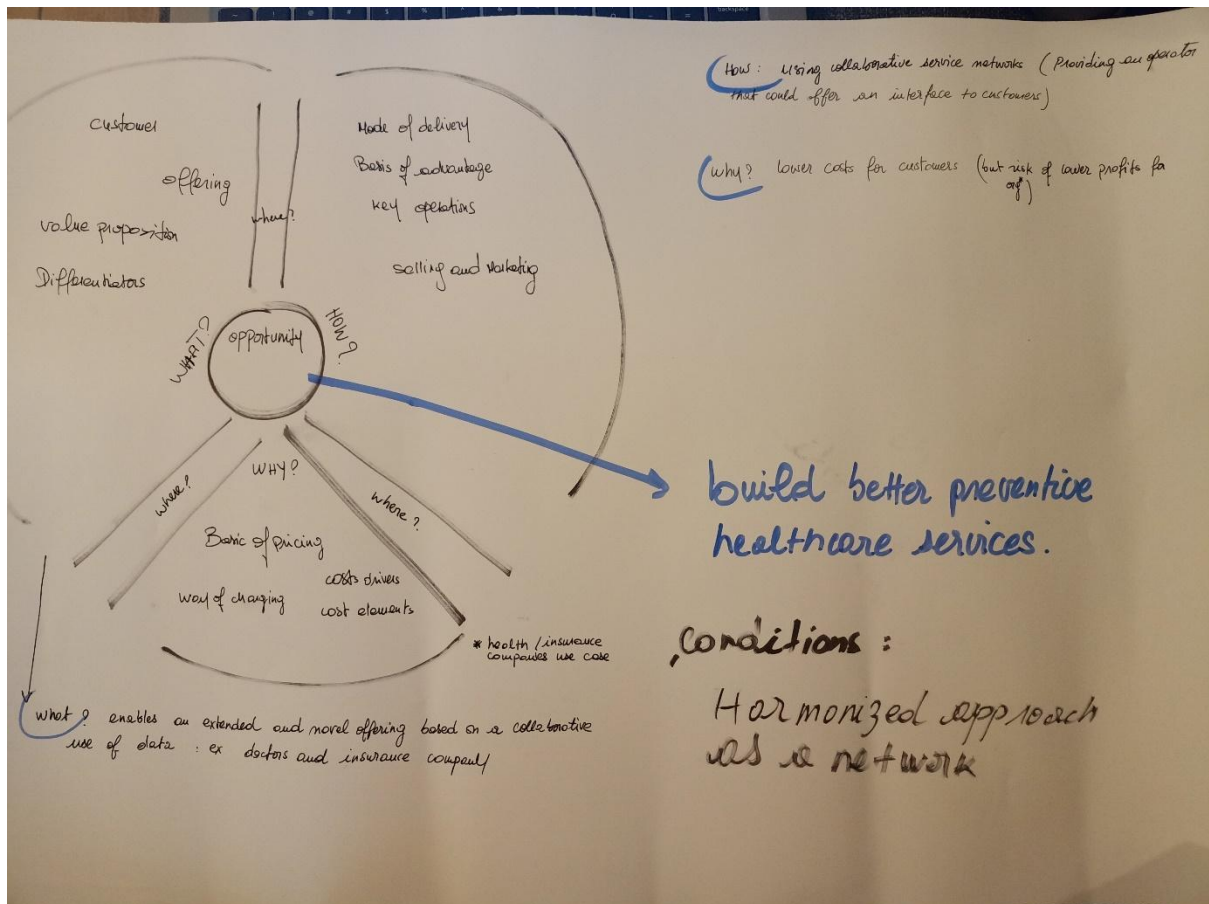
Problem Top 3 problems <ol style="list-style-type: none"> 1. Lack of useful – and tailored – information to citizens to help them understand how they should behave 2. Limitations of movement of citizens with impact on <ul style="list-style-type: none"> • Economy; decreased activities • Health: hospital overloaded, delay in less critical procedures, mental impact • Education: limited access to schools 3. Increase Health risk (morbidity, death) in case of infections 	Solution Top 3 features <ol style="list-style-type: none"> 1. Intelligent assistant (BOT) with tailored and adaptive information 2. Adaptive Mobility score to be provided on demand 3. Access to high quality, up to data personal and population data to support decision making 	Unique Value Proposition <ol style="list-style-type: none"> 1. Empower citizens and make them responsible partners, part of the solution (rather than complaining & part of the problem) 2. Same solution works across all sectors (no need to develop different solutions in silo) 3. Maximizing use of personal data while protecting human and data privacy right: citizens are always in control of their data 	Unfair Advantage <ul style="list-style-type: none"> • Citizens in control of quality – not authorities or other business • Same solution across sectors 	Customer Segments <i>(Note: target customers in one country at a time, cross fertilize across countries)</i> <ul style="list-style-type: none"> • Authorities: partnership, ensure they provide & maintain the rules/ knowledge needed (local and international experts) • Citizens: understand how to respect requested measure while maximizing freedom to move • Businesses (hospitality, travel, employers,...): ensure their business re-open ASAP
Cost Structure <ol style="list-style-type: none"> 1. Knowledge rules – in kind from government/ per country 2. Technology- develop : 5 MEUR ± 20%; host & maintain: 10 MEUR/y ± 20% 3. Communication & marketing: 2 MEUR/y ± 20% 4. Management & overhead: 1.5 MEUR/y ± 20% 5. Return to citizen: based on revenue 		Revenue Streams Revenue Model (DURING PANDEMIC, WHEN DEPLOYED): Freemium model for citizen with monthly subscription fee, yearly subscription model for organizations Revenue (for Belgium with 10 millions inhabitants): Worst case scenario = ± 75 MEUR/Y - Best case scenario = ± 350 MEUR/y		

PRODUCT

MARKET

		development 1	development 2	development 3
		Anti-FB	Pollution, violence, crimes, mobility	Big Brother
who customers	cities individuals ecosystems.	Transparency, Risk of dictatorship	Data for managing cities sustainability Data forgery	afraid of too much control and surveillance -Transparency - education
what proposition and value	GDPR services/abilities faster mobility, more usable services optimization of resources	market place, able to create use case and monetize it Transaction fee promotes restriction of volume of data	The value might not be apparent, only when the whole chain events is simple	
how activities and partners	proof of concepts inter city partnership citizen engagement positive awards	test the business models who pays for moving citizens between cities and countries		
why costs and revenues	Freemium cost of delay/± not doing it, missed opportunity	visualization - comparison to reality		

Based on Business model stress testing: A practical approach to test the robustness of a business model from Timber Haaker, Harry Bouwman, Wil Janssen, Mark de Reuver from 2017



business model stresstest

name

		development 1	development 2	development 3
		eIDAS Revision		
who customers	Public domain Enterprises with group companies Personal Operator	The local municipality can advertise the Need to separate the operator		
what proposition and value	open-source platform sharing data by individual ecosystem	Community engineer can help Personal data will be more secure		
how activities and partners	De-identify data and upload to a secure platform Implement new features Deploy PoCs	No one would Who is going to More companies		
why costs and revenues	Cost of supporting the platform Cost of data management Revenue from consulting and deployment fees	Who is going to Who is going to More companies		

date

Based on Business model stress testing: A practical approach to test the robustness of a business model from Timber Haaker, Harry Bouwman, Wil Janssen, Mark de Reuver from 2017



Appendix 4 - examples of first and second-order codes

Statements of respondents	First-order codes	Second-order codes	Aggregated dimensions
<ul style="list-style-type: none"> - market is still in an early development stage. so they don't have a market readiness level - takes at least five years away to build a significant markets - the wheel is reinvented very often - we know that at the same time you have to give an incentive to all parties, otherwise the product will not be adopted on a large scale - this could make it successful if that entire ecosystem is already digital. but yeah, that's just not the case 	market still in the early stages of development	market is underdeveloped	market development
<ul style="list-style-type: none"> - we want to minimize data, a really important one. when it comes to privacy - the person has a vault on their phone with data that can be her-used - with a service that shares data digitally, there is often less waiting. can it sometimes be near real-time? 	Value added for other roles in the ecosystem	added value	value proposition
<ul style="list-style-type: none"> - it's a chain where all 3 actors are essential and so at the moment the success of such a solution - it's about the emergence of a data ecosystem, that's the phase we're in, collaboration with data sources and data using services is needed - need to work on proof of concepts, as a proof that the service works 	requirements to establish a mature ecosystem	ecosystem perspective	personal data ecosystem
<ul style="list-style-type: none"> - we don't think it makes sense at the moment, namely because the person is not going to pay for a wallet - a person is always in a use-case journey, so you're going to target the person who lets you join, so we don't want to bother people - paradox: people express that privacy is important, but don't behave that way - people are starting to realize more and more that if the service is free, they become the product 	individuals' willingness to pay	willingness to pay	revenue streams
<ul style="list-style-type: none"> - no direct, marketing or communication with end customers, so we now do that via the organizations that provide the services directly to the customers - there is a need for support from data using services to enable adoption - if the market is ripe, then use data providers and data buyers to roll out the wallet via their customer groups 	data using services as distributors	distribution channels	adoption