



Authors:

Antti Poikola
Kai Kuikkaniemi
Ossi Kuittinen
Harri Honko
Aleksi Knuutila
Viivi Lähteenoja

Editor for the third edition
Viivi Lähteenoja

MyData

– an introduction to human-centric
use of personal data

3rd, revised edition

Our digital footprint is growing rapidly. Individuals find it difficult to understand what information about them is being collected by different organisations, digital services and platforms and how this information is being used. The MyData model seeks to give people access to their information and control over their data. The General Data Protection Regulation (GDPR) provides a good basis for this, but the systems under the GDPR lack easy-to-use tools for people to access and transfer information about themselves from one system to another.

Since 2014, the Ministry of Transport and Communications has been involved in supporting the MyData movement, by encouraging businesses, administration and individuals to reflect on the possibilities for and implications of a new type of personal data model. In our work on this, we have also responded to the need for international discussions and have promoted the MyData model as a means to deliver a human-centric approach that is needed for the European Commission's data strategy. Within the MyData movement, the Ministry has had a close supportive role and it is an active member of MyData Global. We believe that the view of 'the individual in control of his or her own data' should not only be considered a philosophical basis, but should also be seen as a key design principle when drafting legislation.

The Government Programme of Prime Minister Sanna Marin also acknowledges the MyData approach. The Programme outlines the need to evaluate various ways to better safeguard the rights of citizens to manage their own data in accordance with the MyData principle.

In the context of these objectives, the Ministry of Transport and Communications has supported taking the development of the MyData idea further. This updated version of the previous MyData study addresses the need for an up-to-date version of the MyData White Paper. The study sheds light on the benefits of MyData, the path to the practical implementation of it and shows the way towards an interoperable and people-oriented ecosystem for sharing personal information. The realisation of the MyData model requires data portability in practice, for example trusted personal data intermediaries called MyData operators. The updated work shares the understanding of the role of MyData in the context of EU policy-making and upcoming actions at the EU level. This work is anticipated to support especially the exercise of the sovereignty of individuals over their data envisioned in the European data strategy. The views expressed in this report are those of the writers and do not necessarily reflect the views of the Ministry of Transport and Communications.

Helsinki, 08 July 2020

Lotta Engdahl

Senior Specialist,
Ministry of Transport and Communications, Finland

Thanks

3

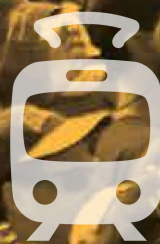
The editor for this third, revised edition would like to extend her heartfelt thanks to **Antti Poikola** and **Kai Kuikkaniemi** for the supportive discussions and to **Joss Langford** for his editorial work.

Many thanks are due also to **Maria Rautavirta** and **Lotta Engdahl** from the Ministry of Transport and Communications, Finland, for their input and support.

This work would not have been possible without the support of the MyData Global team whose dedication allowed for the editor's time to work on this report: thank you **Anna Tuomainen**, **Karolina Mackiewicz**, **Sille Sepp**, and **Teemu Ropponen**.

Table of contents

Preface	2
Thanks	3
Table of contents	4
1. Introduction: What hangs in the balance with personal data?	6
1.1 What is MyData?	7
1.2 Data protection and data utilisation	9
1.3 People and organisations	10
1.4 MyData in context	11
2. What would MyData change?	15
2.1 Actionable rights	16
2.1.1 Actionable rights for more availability of data	18
2.2 Data empowerment	21
2.3 Open ecosystems	23
2.4 The potential of MyData: Risks and benefits	25
2.4.1 People	26
2.4.2 Companies and other organisations	27
2.4.3 Society at large	28
3. MyData infrastructure	30
3.1 From platforms to networks	32
3.1.1 The API ecosystem model	33
3.1.2 Platform model	33
3.1.3 The MyData operator model	34
3.2 Human-centric personal data ecosystems	35
3.2.1 Ecosystem roles	35
3.2.2 Specialised data operators	36
3.3 MyData operator reference model	38
3.3.1 Identity management	38
3.3.2 Permission management	38
3.3.3 Personal Data Storage	40
3.3.4 Service management	40
3.3.5 Value exchange	41
3.3.6 Data model management	41
3.3.7 Personal data transfer	42
3.3.8 Governance support	42
3.3.9 Accountability and logging	42
4. Examples of MyData application areas	44
4.1 Self-assembled profiles: Portable media profile	47
4.2 Verified data: CV 2.0	48
4.3 Distributed utilisation of user data: Data from a retail loyalty card	49
4.4 Societal data collection: MyData and research	51
4.5. Data as part of interaction: Transparency of public services	52
4.6 Data about things: Is the data about my car about me?	53
References	55



The goal of the MyData movement is a fair, sustainable, and prosperous digital society in which people get value from their data and set the agenda on how it is used and for organisations, the ethical use of data is always the most attractive option.

1. Introduction: What hangs in the balance with personal data?

Recognised already in the early 2010s, personal data in the 2020s and beyond is fast becoming a defining force for the way the world works.¹ People, businesses, technologists, legislators, and societies at large are in various stages of recognising and reacting to this reality and its future potential.

Our personal digital footprints are growing rapidly. It is difficult for us to grasp what data is collected about us, what data different organisations have, and we often do not understand the ways in which this data is utilised. This raises concerns that are further reinforced by revelations about massive state data collection systems and reports of privacy-compromising legislative reforms and data breaches. Surveys clearly show a trend that people increasingly trust organisations less to use their personal data properly.²

As a result of digitalisation, the amount of stored data is constantly increasing, and at the same time its commercial and other utilisation is increasing. Much of this data is personal data. Personal data can be used to develop, among other things, proactive healthcare and applications for managing one's own life and learning about oneself. Personal data allows companies and other organisations to tailor their services to better meet people's needs. At the societal level, personal data can be used as a basis for decision-making or, for example, for a more precise targeting of public services. However, the benefits from the increasing use of the rapidly growing amounts of personal data generated are currently heavily concentrated in the hands of a small number of organisations. As a result, much of the potential use and benefits from personal data goes unrealised.

Two fault lines emerge from the current situation. The first is the tension between data protection and data utilisation, and second is the existing imbalance between the interests of and benefits for people on the one hand and organisations on the other. MyData is a model through which both sets of considerations can be balanced.

In MyData thinking, the usability of personal data is approached by placing the person at the center of the use of data about themselves. In this way, data protection and data usability do not contradict each other, but on the contrary, support each other. Strong data protection and transparency in the use of personal data increases trust between people and organisations and thus opens up opportunities for the development of innovative services based on personal data. The goal is that an ethically sustainable human-centric way of managing personal data would always be the most practical and also economically profitable way for organisations to operate in the future while providing people the power to set the agenda for how data about them is used and to benefit from its use.

1 World Economic Forum 2013.

2 Sirkkunen & Haara 2017, Salesforce 2019.

1.1 What is MyData?

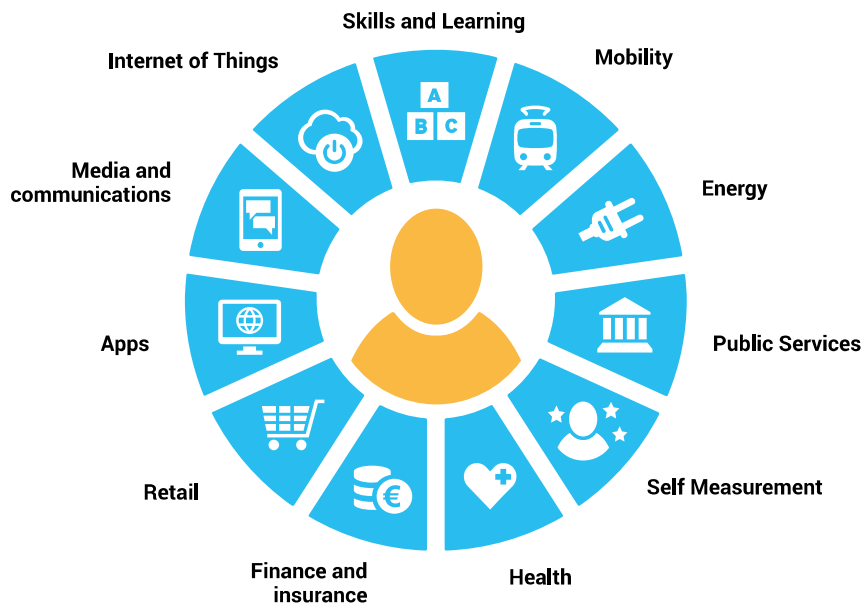


Figure 1.1 Increasingly, growing amounts of personal data are created in different areas of human life. The MyData principles are particularly conducive to cross-industry data transfer. Areas that are important to MyData and that produce a lot of personal data include mobility, the health and well-being sector, and banking and insurance.

MyData refers to, on the one hand, a human-centric model for the management and utilisation of personal data, which seeks to endow people with self-determination regarding data about themselves and, on the other hand, to a growing movement working towards the realisation of such a model in digital societies globally. The premise of MyData is that people themselves can use, manage, and give permissions regarding data collected about them, such as shopping, mobility, financial, or health data. Human-centric data management creates interoperability and reconciles digital human rights and high data protection standards with the promotion of data availability and business opportunities.

MyData is a model for personal data use. Legal and technical definitions of personal data vary between jurisdictions and disciplines. From a MyData perspective, the distinctive element of personal data is the “personal” element. The set of personal data about a person is a digital representation of their personhood and has analogous rights, legal and beyond, attached to it. In addition, MyData thinking recognises the fact that no person is an island. In other words, our personal data is rarely data relating to our individual person alone, but rather is generated through interactions. My date of birth is data about the person who gave birth to me, my click is a share of your post. Individuals inevitably exist in groups, communities, and societies with which they are inextricably linked.

MyData is becoming established as an international concept. The thinking and the community of actors identified and identifying with the moniker MyData has expanded significantly since the publication of the first version of this study.³ A subsequent series of international meetings and conferences resulted in the establishment of an international nonprofit organisation, MyData Global,⁴ headquartered in Finland, in 2018. The purpose of MyData Global is to empower individuals by improving their right to self-determination regarding their personal data.⁵

³ Poikola, Kuikkaniemi, & Kuittinen 2014.

⁴ <https://mydata.org/>.

⁵ <https://mydata.org/bylaws/>.

Primarily, and in addition to the movement and the organisation stewarding it, MyData refers to a phenomenon and a change in mindset that seeks to move the management and processing of personal data from the current model to a human-centric one. In 2017, the global MyData community published a Declaration of Principles,⁶ in which it outlined the major shifts it strives for as well as the principles based on which those shifts can be achieved. This declaration is the first version of an evolving document, which is revised as events unfold.

MyData declaration

The shifts needed for realising the human-centric personal data model:

1. From formal to actionable rights

In many countries, individuals have enjoyed legal data protection for decades, yet their rights have remained mostly formal: little known, hard to enforce, and often obscured by corporate practices. We want true transparency and truly informed consent to become the new normal for when people and organisations interact. We intend access and redress, portability, and the right to be forgotten, to become “one-click rights”: rights that are as simple and efficient to use as today’s and tomorrow’s best online services.

2. From data protection to data empowerment

Data protection regulation and corporate ethics codes are designed to protect people from abuse and misuse of their personal data by organisations. While these will remain necessary, we intend to change common practices towards a situation where individuals are both protected and empowered to use the data that organisations hold about them. Examples of such uses include simplifying administrative paperwork, processing data from multiple sources to improve one’s self-knowledge, personalised AI assistants, decision-making, and data sharing under the individual’s own terms.

3. From closed to open ecosystems

Today’s data economy creates network effects favoring a few platforms able to collect and process the largest masses of personal data. These platforms are locking up markets, not just for their competitors, but also for most businesses who risk losing direct access to their customers. By letting individuals control what happens to their data, we intend to create a truly free flow of data – freely decided by individuals, free from global choke points – and to create balance, fairness, diversity and competition in the digital economy.

MyData principles:

1. **Human-centric control of personal data**
2. **Individual as the point of integration**
3. **Individual empowerment**
4. **Portability: access and reuse**
5. **Transparency and accountability**
6. **Interoperability**

6 <https://mydata.org/declaration/>.

1.2 Data protection and data utilisation

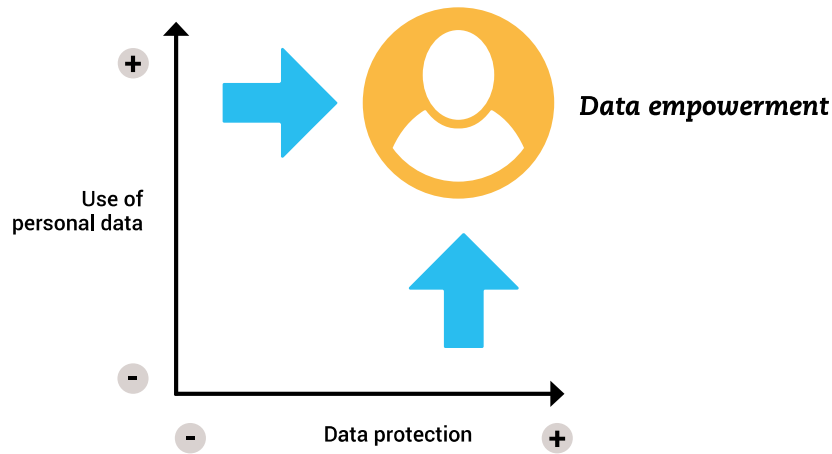


Figure 1.2 MyData allows for the effortless use of personal data so that the benefits are maximised and the weakening of privacy protection is minimised.

The regulatory challenge in terms of personal data is to balance measures for data protection and the availability of personal data for utilisation. Any collection and use of personal data may compromise privacy. For this reason, one of the starting points of the strong European data protection regulation is the principle of data minimisation by organisations, according to which only personal data that is necessary for a purpose defined can be stored: “the less personal data is collected and shared by organisations, the lower the risk of privacy compromise”. In this instance, the regulator has chosen to err on the side of data protection, potentially at the cost of reducing the value from data that accumulates for people themselves and in a way that contrasts with the megatrend of increasing the amount and use of personal data.

On the other hand, U.S. law and practice allow organisations fairly free reign to utilise personal data as long as the user has checked the box that they have read and agree to the terms of use. In this instance, the regulator has chosen to emphasise the wide use and re-use of personal data, but this is accomplished at the expense of privacy protection and solely on the terms set by private companies. According to the U.S. Federal Trade Commission (FTC), companies that collect and sell consumer data possess accurate data repositories which cover almost every U.S. consumer and which have been collected without the knowledge of those consumers.⁷

MyData thinking strikes a balance between these two regulatory approaches by bringing the perspective of data utilisation on an equal footing with privacy protection. This is achieved by providing people with opportunities to utilise the data collected about them by companies as well as the ways and means to control how data is collected, processed, utilised and further shared.

7 FTC 2014.

For fair, sustainable, and prosperous digital society through a human-centric approach to personal data



Figure 1.2 MyData aims at a fair, sustainable, and prosperous digital society where people get value from the data about themselves and set the agenda on how it is used, and for organisations, the ethical use of data is always the most attractive option.

Another, very closely linked, aspect in which MyData is a balancing force is when it comes to the relative power of people and organisations. The current economies and societies of data are by and large designed and dominated by organisations and their needs and benefits. The burden of reading terms and conditions to be sure one's data is not used in ways which one does not agree with or which compromise one's privacy in an unacceptable way on the one hand, or putting in the effort to exercise their legal rights to gain better access and use of the data about oneself on the other, is left for the individual person. This is the result of organisation-centric design.

The MyData approach is a way to address this imbalance of power by placing people in the centre of the data about themselves and thus shift the current paradigm, and the organisations operating within it, to better serve people and societies. Society's activities are increasingly based on the collection and utilisation of data. The starting point for MyData thinking is human-centricity, in which the functioning of society is built around people. It is a counterbalance to the existing trend that focuses solely on the operating capabilities of organisations. The crucial difference is whether the mechanisms for collecting and utilising data are designed primarily from the perspective of people or of organisations. When a mindset is adopted that prioritises functioning and flourishing capabilities of people, a society can emerge that serves both people and organisations in a balanced and fair way.

1.4 MyData in context

Personal data is a subset of the data currently circulating in what is called the data economy. Technically, there is nothing particular about data that is personal compared with non-personal data. In terms of business models, the law, ethics, and policy, however, the difference is drastic, and these differences also have technological implications.

Designation of data as personal or not is not static or irreversible: consider processes of data anonymisation and enrichment. Definitions of personal data typically include the notion of not only directly identifying data about a natural person (such as a social security number), but also such data as can be, in some combination, used to identify such a person. As circumstances change, e.g. the controller of a data set or the technologies available to use on a data set, previously non-personal data may become such that it can in fact identify a natural person both in theory and within reason also in practice.

An ongoing debate exists over whether the distinction between personal and non-personal data is enduringly useful. Some⁸ argue against the distinction because de-intentification is not truly possible, while others⁹ maintain that it is. This paper does not intend to shift this particular debate but recognises that in fact more and more of the data in the data economy is and will be of the kind that can reasonably be used to identify a natural person. For this reason, the MyData model for personal data use and management is one that ought to be incorporated into any regulation or governance framework intended for the broader data economy.

The human-centric vision of MyData has strong roots in the European value tradition as well as an important place in EU policy to come. The European commission published in February 2020 their data strategy, at the core of which is the understanding that in "a society where individuals will generate ever increasing amounts of data, the way in which the data are collected and used must place the interests of the individual first, in accordance with European values, fundamental rights and rules".¹⁰ Further, the "Commission's vision stems from European values and fundamental rights and the conviction that the human being is and should remain at the centre".¹¹

MyData is a complex idea with implications for businesses, legislators, technologists, and societies at large. It is also an idea that is continuously being developed by a global community of experts and thought leaders. This paper provides an overview of the landscape of MyData thinking so far.

8 Ohm 2010, Narayanan and Felten 2014.

9 Cavoukian and Castro 2014.

10 European Commission 2020, p. 1.

11 Ibid. p. 4.

Can data be owned?

MyData is often associated with the idea of data ownership. In everyday language, it is natural to say that “people should own their data”. However, the concept of data ownership is not unproblematic for a number of reasons and instead we in paper write about non-property rights to personal data and prefer the framing “data about a person” over “a person’s data”.

MyData aims to give people the right, opportunities, and means to utilise data about themselves and set the agenda for how it’s used and by whom. However, this does not mean that organisations that collect or otherwise use personal data do not or could not have rights to the same data.

Ownership is easy to understand in the case of personal property or real estate. The owner may exercise their ownership to the exclusion of others, i.e. property is an excludable good. The owner of the chair can usually decide who is allowed to sit in the chair or what color the chair is painted. Owning data is not quite so straightforward. Many people may know the same things, i.e. can be characterised as a non-rivalrous good. The fact that one person knows something and utilises it does not in itself prevent others from simultaneously knowing and utilising the same information or data. Similarly, when data is copied, the use of one copy does not prevent the use of other copies.

As a general rule, information or data are not subject to exclusive rights, no one owns data. Other rights to data are usually negative rights, i.e., they oblige inaction on the part of others. Through the exercise of such rights, the availability of and control over data can be limited so that, practically, only few parties have the opportunity to utilise it. Some data may be subject to more limited rights due to, for example, to copyright, trade secrets, or privacy protection. This practice can be characterised as creating artificial scarcity, or limiting the use of a good beyond what is strictly necessary.

It can also be the case that many parties may have a legitimate interest in any number of data sets. For example, in a customer relationship, a store has a good reason to be able to use the customer data it collects, even if customers have rights to the same data, such as the ability to retrieve the data for themselves or delete the data at the end of the customer relationship.¹²

12 Pitkänen 2014.

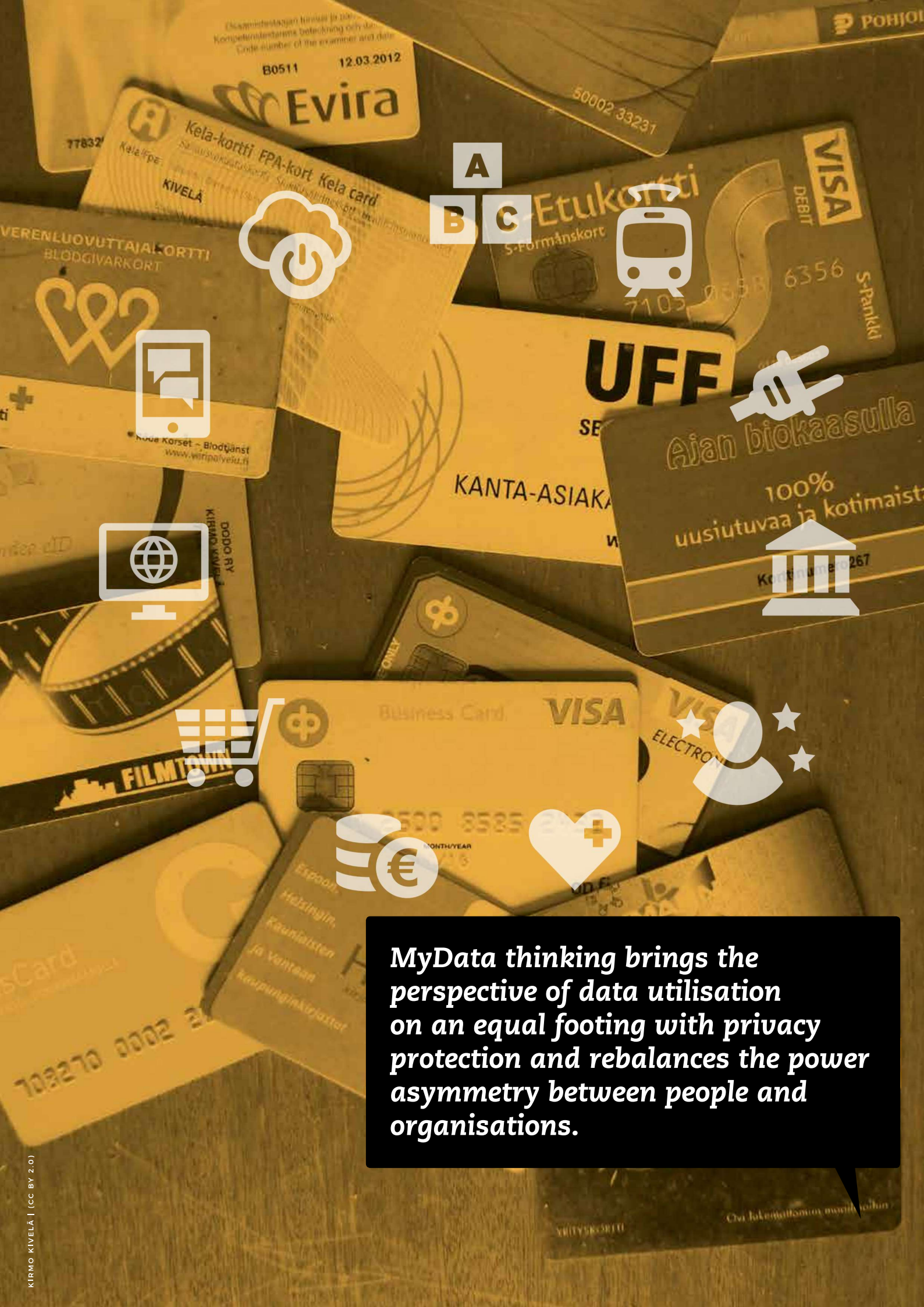
Public Administration and MyData

The public sector involves a wealth of legally mandated personal data processing that is not based on the consent of the people themselves. However, the legal basis for processing does not preclude the implementation of the minimum requirements of the MyData principles.

Rights and possibilities to control the data collected about one vary on a case by case basis. The minimum requirement for implementing the MyData principles is that people get access to data about themselves and can use it also elsewhere and for other purposes. Full control that would allow, for example, editing or deleting that data from the original source is not a prerequisite for MyData. The key to implementing the MyData model is that personal data is technologically easily usable by people and by those services and actors for whom the person wants to allow the use of that data.

The processing of personal data is necessary in the performance of authorities' duties: e.g., tax authorities need personal data for taxation and people may not remove themselves from tax registries. Public administration actors can comply with MyData principles by making personal data processing transparent and by providing people, in a format that is machine-readable and allows for easy re-use, all the data they have access to in a traditional online service interface.

Public sector actors can even set an example in the development of personal data interfaces. Personal data management that complies with MyData principles could facilitate the interoperability between public and private services when the person themselves is allowed to transfer data from one service to another. The public sector has several unique data pools not available elsewhere. For example, the data contained in basic registers is essential basic data, official and reliable in nature, and therefore particularly valuable.



MyData thinking brings the perspective of data utilisation on an equal footing with privacy protection and rebalances the power asymmetry between people and organisations.

2. What would MyData change?

15

2. What would MyData change?

MyData strives to give people the opportunity to use their personal data for their own purposes and to share it securely on their own terms. In Europe, the GDPR provides a good basis for this in principle, but it must be recognised that previous data protection legislation has also offered people such rights. The problem so far has been that they have often remained largely nominal and formal rights because they are unknown and difficult to use by people on the one hand, and difficult for the regulator effectively to monitor and enforce. The MyData declaration of 2017 states as a goal for MyData that “access to personal data, correction and portability, and the right to be forgotten evolve into one-click rights that are as simple and effective to use as the best online services of today and tomorrow”. This is referred to as the shift from formal to actionable rights.

The second shift described in the MyData declaration is the one from data protection only to people’s empowerment with data. This describes a new kind of mentality and mindset, whereby people are not seen merely as passive beings to be protected from harm but also as fully-fledged citizens and agents with the will and capacity to better their lives through the use of personal data gathered about them.

The third shift that MyData seeks to effect is the one from closed to open ecosystems. The rapid growth of the data economy is based on network effects. Their most effective utilisers are those platform actors, who have the ability to collect and process the largest amounts of personal data. These platforms close markets not only from their competitors but also from many other companies that are now in danger of losing direct contact with their own customers. By placing the people at the centre of the data about them, MyData will enable truly free flow of data and to create balance, fairness, diversity, and competition in the digital economy.

2.1 Actionable rights

The right to control data about oneself is often considered as a human right in the digital age. MyData principles are a way to ensure people both have these rights and also have the practical tools to exercise them. These tools help people manage the data about themselves, their privacy, and their lives on- and, increasingly, offline though exercising those rights.

Exercising control over the data collected about oneself starts with being made aware of the nature and extent to which personal data about them is being collected. A person ought also to have access to that data and have it corrected or deleted if needed. They ought to be able to check and see which actors are using what data. In addition, and crucially for being able to benefit from this data, people must have the rights to obtain the data about them and to be able to share it for whichever use they wish as well as to stop any such sharing. Putting the person in the driver's seat and in control of the data about them obviously does not mean absolute or unrestricted rights to do everything one might please. Just like being in control of the bicycle one's driving doesn't mean there are no traffic rules and good manners to follow, so also people's rights regarding control of the data about themselves are at times restricted by law or other means to protect themselves and others and to ensure a pleasant time on the road for all.

Legal rights in many jurisdictions support the implementation of MyData principles. Yet simply following the requirements of the law alone is not enough to realise the MyData model and to create a human-centric personal data ecosystem. Legislation, regulation and technological changes can all contribute to the realisation of MyData. Regulation can have an accelerating or decelerating effect, but legislation alone will not bring about change. The challenge is the internationalisation of the data economy and the application of regulation to services operating globally.

MyData aims to give people more opportunities to control their own data than often is the minimum legal requirement, thereby making personal data a reusable resource with privacy protection kept in mind. In Europe by law, data controllers may only collect, store, process, and use personal data for pre-defined purposes. But the persons themselves whose data is involved are not subject to such restrictions. A person themselves can benefit from this data by using it flexibly for purposes they themselves define. In practice, this can be done by permissioning the re-use of personal data and sharing data between services according to a person's own needs and wishes.

The minimum implementation of MyData is that people can download the data about themselves in a machine-readable format for themselves. However, for innovative applications, it would be better to have continuous access to up-to-date data through standardised programming interfaces (APIs). In this way, updating data would not require visits to the data provider's website, but services could be automated. For example, purchase data is most useful if you receive an electronic e-receipt automatically as soon as you pay for your purchase, like a paper receipt today.

E-receipt for businesses and consumers

For many, the so-called “receipt circus” is familiar. Paper receipts fill up your pockets and wallet, some are personal small purchases, but there are also important documents among them that should be kept as proof of warranty or attached to a business expense claim for the employer.

In Finland alone, 1.3 billion card payments are made annually, of which ten percent with corporate cards. For consumer card payments some banks are already sending an electronic receipt in near real time to a mobile application provided by the bank. The broader idea of digital so-called e-receipts would mean that the buyer would receive, regardless of the payment method, a data receipt, which would also include purchase line item data, to their desired location without the need to request one separately during the purchase transaction. Businesses need receipt information for their accounting systems and consumers could have their own data wallet in which eReceipts are automatically saved.

The difference between a leather wallet bursting with paper receipts and a data wallet is the fact that in the latter, useful programs can be installed that process and visualise information. In the data wallet could function, for example, a real-time financial monitoring program, a warranty receipt archive, and other personal services related to purchasing behavior. It is up to the user to choose which programs to install in their data wallet, but the merchant can also recommend programs that specifically take account of the receipt data they send.

Financial management automation is one of the key areas for productivity development of businesses. Succeeding requires both standardised structured data and open interfaces. The savings potential of an electronic e-receipt compatible with accounting systems has been estimated at almost EUR 800 million per year. The Finnish Taltio project has defined an ecosystem for the transmission of e-receipts based on the e-invoicing standard. Follow-up projects with Real-Time Economy aim to start receipt transmission extensively, and in cooperation with Estonia among others, to add machine-readable data to business e-invoices¹³ which can, for example, deepen the analysis services of accounting firms.¹⁴

Although the automation of corporate financial management is now a driver in the development of e-receipts, the same mechanisms should be used also on the consumer side, so there would be no need to build a separate approach to transmission of consumer receipts at a later stage. For consumers, the e-receipt would act as a line-by-line warranty certificate as well as tool for recall, service reminder, etc notifications.

¹³ E-invoices in European Union are supported with Directive 2014/55/EU of the European Parliament and of the Council of 16 April 2014 on electronic invoicing in public procurement by empowering companies to request invoices in electronic format.

¹⁴ Taltio and Real-Time Economy Projects: <https://www.lvm.fi/en/-/interoperable-and-better-digital-services-in-financial-data-957006>

2.1.1 Actionable rights for more availability of data

As described above, European data protection law is based on the protection of privacy and the accuracy of the data contained in personal data registers, not so much on the realisation of the potential benefits of personal data. The EU Charter of Fundamental Rights states that “everyone has the right of access to data which has been collected concerning them and the right to have that data corrected”.

According to the GDPR, personal data may in principle only be “collected for specified, explicit and legitimate purposes and not further processed” (Art. 5). Examples of permitted secondary uses specified by law are “archiving purposes in the public interest, scientific or historical research purposes or statistical purposes” (Art. 5). From a privacy perspective, it is justified that organisations should not be able to use personal data for new purposes out of the blue and especially without the knowledge of the people concerned. On the other hand, sensible uses of data may only emerge after it was originally collected. Establishing a legitimate basis for processing after the fact of collection is often difficult in practice, even if the new use is considered acceptable by the people whose data is involved. In accordance with the spirit and purpose of the law, the purpose of the use of personal data should always be specified as precisely as possible. This feature of the regulation has often been criticised as limiting the availability of personal data.

On the other hand, a significant new change in terms of the availability and usability of personal data is the rights of the data subject to request copies of data being processed about themselves under the right of access and to obtain them in a commonly used electronic format (Art. 15). In addition to the right of access, the GDPR defines a new right, the right to data portability from one controller to another (Art. 20). This right seeks to ensure that data collection does not become a factor restricting competition between services, but that people are free to choose between competing service providers and also transfer their data with them if they decide to switch services.

The European Strategy for Data recognises the fact that the formal rights granted to individuals by the GDPR need to be made more practical. In the strategy, the commission highlights the need for the empowerment of people to exercise their rights, writing that people “suffer from the absence of technical tools and standards that make the exercise of their rights simple and not overly burdensome” before going on to acknowledge that “the potential of Article 20 of the GDPR to enable novel data flows and foster competition is recognised [but], as a result of its design to enable switching of service providers rather than enabling data reuse in digital ecosystems the right has practical limitations.”¹⁵

Through the MyData approach, cited also on p. 10 of the data strategy as promising “significant benefits”, the availability of personal data is enabled through the concept of the person as a point of integration regarding the data about themselves. Framing the person as in charge of the personal data about themselves allows for infinite use and re-use, on the terms set by the person, of personal data. A person is allowed freely to process, use, and share the data about themselves without the need to provide a pre-defined purpose for such processing, use, or sharing. However, it should be noted that, under the GDPR, receiving consent from the data subject for the collection and processing of personal data does not override the requirement for the data processor to comply with other principles of data processing. This means that even if a person consents to the collection and use of personal data about themselves, i.e. there exists a legal basis for processing, the processor must also be able to demonstrate that such processing is in fact justified and compliant with the other principles enumerated in Article 5 such as data minimisation and purpose limitation.

¹⁵ European Commission 2020, p. 10.

Mixed data sets and the GDPR

The Regulation on the free flow of non-personal data by the European Commission, applicable from May 2019, is the regulation governing the movement of all data which falls outside the GDPR's definition of personal data. In its later "Guidance on the Regulation on a framework for the free flow of non-personal data in the European Union",¹⁶ the commission considered the notion of mixed personal and non-personal data sets. They conclude that "[m]ixed datasets represent the majority of datasets used in the data economy and are common because of technological developments such as the Internet of Things (i.e. digitally connecting objects), artificial intelligence and technologies enabling big data analytics." Further, they stipulate that mixed data sets where the personal and non-personal data are "inextricably linked", the whole set is to be subjected to the GDPR governing personal data.

As one of the goals of the MyData approach is the increased re-use of personal data by way of, for example, combining data from different sources, it is likely that the number of mixed data sets created further increases with the adoption of MyData practices and policies. This will not, however, prove problematic, as strong data protection is already baked into the MyData principles.

¹⁶ European Commission 2019.

Data portability in practice

Article 20 of the GDPR introduces a new right for people to download their own data either to themselves or transfer it directly between services.

Data portability covers all personal data concerning a person which they have provided to a controller. The data must be available in a structured, commonly used and machine-readable format. This right does not trigger the right to be forgotten, data transfer does not automatically initiate data deletion from the original source.

This is one of the points in the regulation for which companies are the least prepared. Some companies already offer the possibility for their customers to download their own data from their websites, but for most companies, this is completely new and especially the ability to transfer data directly between two organisations at the customer's request requires development efforts. The Working Party of European Data Protection Supervisors has provided clear guidance on the interpretation of the right to data portability,¹⁷ but there is still a long way to go before this can be implemented in practice in organisations.

In the French Dataaccess¹⁸ project, a telecom operator has convened an open working group to develop definitions, guidelines, openly licensed design elements and open source implementations, aimed at providing a common, easy-to-use model for practical implementation of personal data portability in organisations. The working group includes the French Data Protection Supervisor (CNIL) as well as energy companies, banking and insurance sector actors and cities.

Dataaccess has been in part modeled on American Green and Blue Button practices that promote consumers' access to energy consumption (green button) and health (blue button) data to themselves.¹⁹ Instead of industry-specific solutions, Dataaccess strives to reduce the challenges of data portability in all organisations and thereby to enable users of different services to have a consistent and easy user experience of data portability. The goal is that, with a practical implementation model, the requirement of the law turns from a necessary evil into an opportunity to develop trust-based customer relationships and foster innovation.

Dataaccess has published definitions of three complementary use cases: 1) downloading data onto one's own computer 2) transfer data between applications and 3) transferring data to a personal data wallet (personal cloud). In addition, The Dataaccess publication outlines a typical data portability implementation project and divides it into easily manageable parts.

¹⁷ European Commission 2017.

¹⁸ http://mesinfos.fing.org/wp-content/uploads/2018/03/PrezDataaccess_EN_V1.21.pdf.

¹⁹ <http://energy.gov/data/green-button> and <http://www.healthit.gov/bluebutton>.

2.2 Data empowerment

The ambition of the MyData model is to provide ways in which people, while being protected, can be empowered by and with the data that exists about them. In a world where we are continuously assaulted by news of data breaches, bombarded by requests for our data for purposes we don't fully understand, and startled by revelations of the adverse societal and global effects of what personal data (mis)use is capable of, it is easy to understand that we often feel powerless or become apathetic regarding our own power regarding the data about us. MyData is a way to empower people to use their data to make a difference for themselves and for their communities and societies.

Loyalty card data for more sustainable consumer behaviour

Data about how people behave can empower them to make more educated choices that are aligned with their values. For example, the Finnish retail chain S-Group has developed a carbon footprint calculator based on customers' purchase data.²⁰

The service combines a consumer's purchase data, the product group data involving data on agricultural production, processing, packaging and transportation to the shop, and reference data from the Natural Resources Institute Finland²¹ allows the service to visualise the carbon footprint of the consumer's shopping habits over time. This data can then be used by the consumer themselves to evaluate their consumption habits, be more value-driven in their behaviour, and to monitor the effects their shopping choices have on their overall carbon footprint as visualised in the service.

Empowering services like the carbon footprint calculator will become even more useful when they begin to be able to incorporate data from various different sources according to MyData principles.

20 <https://s-ryhma.fi/en/news/leading-finnish-retailer-s-group-launches-a-calcul/7bJ1UjgsE6C47YBQxdr6s6>.

21 www.luke.fi/en.

Another aspect of data empowerment is the development of ways for organisations to innovate by being able to combine and re-use data sets from various and also cross-sectoral sources, again with the knowledge and permission of the people involved, to provide services and products that serve people better than ever.

A major hurdle for organisations in this respect is the prevailing lack of trust among consumers towards how organisations currently handle personal data. A 2019 study²² found that "73% of customers say companies' trustworthiness matters more than it did a year ago, and 54% say it's harder than ever for a company to earn their trust." Further, "63% of customers say most companies fail to use their data transparently, and 54% believe most companies don't use their data in a way that benefits them."

Adopting and clearly communicating the MyData approach to personal data use is a way for organisations to signal that they act in a trustworthy manner regarding the data about their customers.

Trusted European cloud services

The European data strategy includes the development and adoption of a European cloud federation. This project is connected with the creation of a horizontal data governance framework for European data spaces, which aims to contribute to the availability and use of data in the EU. MyData remains agnostic as to the specific technologies used to process and store personal data, including the cloud (at the edge or not) but welcomes the emergence of good and trustworthy options both for people and organisations in terms of which providers they use. When realised, the MyData vision includes options for all actors in digital societies to use technologies and service providers whom they trust. To ensure the desired trustworthiness of future European cloud services, such as are being developed in Germany in the Gaia-X project, the MyData approach ought to be incorporated from the first design phases onwards.

2.3 Open ecosystems

A significant consequence of the implementation of MyData principles is the splitting of personal data value chains and the centralisation of data management around the person. This opens up opportunities for new entrants and breaks the traditional boundaries between sectors and industries.

One of the goals included in MyData thinking is to open up the business environments related to personal data to development, competition, and cooperation so that monopolistic scenarios²³ about the future of the information society do not materialise. In a functioning constellation of services, the aim is to avoid that any single organisation, or a small set of organisations, has a monopoly on personal data. Open personal data ecosystems mean avoiding a situation in which all or most personal data is located in one or a few services, or that personal data based services can only be implemented with one type of technology.

The value chain of personal data processing consists of the collection, transmission (management) and utilisation of personal data. Traditionally, the entire value chain exists within a single organisation. For example, when data about a person's account transactions is generated in a bank's systems, the bank processes and transmits the data to produce account statements and online banking views to its customers. With MyData, a shift occurs from value chains exclusively implemented and controlled by individual organisations to open and decentralised value networks, in which actors specialising in different stages of the chain can emerge.

PSD2 & Open banking

An example of a successful splitting of a personal data value chain can be found in the banking sector in the EU and the UK. After the introduction in the EU of the revised Payment Services Directive (PSD2), the UK Competition and Markets Authority (CMA) issued a ruling requiring major banks to open up account transaction data to licensed third parties in 2016. These third parties are able to offer services such as personalised financial management based on the data available from multiple banks in which a person might have accounts or enable business customers to receive funding decisions faster than ever before. Similar open banking initiatives have since been launched in Australia and Nigeria.

In a traditional data processing value chain, the company or service where the data is generated and collected is the gatekeeper. For example, purchases made with a retail chain's loyalty card generate data, the use of which is decided by the loyalty system administrator. But if the service has implemented MyData principles by making that data available to the card user(s) though, e.g., a personal data interface, then the user of the service becomes the gatekeeper and they can decide to share that data with others. This allows for the person-controlled use, re-use, and combination of data from different data sources.

Splitting the personal data value chain is a significant change. It is also detrimental to some of the business models that are typical today. For example, a business model based on attracting as many users as possible by offering a free service and then generating revenue from selling on the data collected from those users does not function in an environment where people can quickly and easily share that same data with other processors as they wish. On the other hand, splitting the value chain creates a framework for new types of business models to emerge and businesses should see it as an opportunity to enter into more transparent and more deeply trust-based customer relationships. Companies could also develop their business models based on better customer insights and leverage customers in the production and design of their services.

An important role in the decentralised value network is played by specialised providers of personal data transmission and management services. The MyData model calls these MyData operators, which are described in more detail in chapter 2 of this paper. In the MyData model, the provision of personal data infrastructure is decentralised in the same way as banking or telecom services. These decentralised basic services should work together in the same way that payment by debit card works independently of the card issuing bank and mobile calls are connected from one telecom operator's network to another. For the customer, this would mean that it is easy to change services so that the person's own data goes with them. The possibility of switching increases competition and trust in services and thus accelerates the development of services.

Personal data spaces and the EU Data Act 2021

The European data strategy makes repeated reference to, but does not exactly define, "personal data spaces", and mentions potential regulation for emerging actors to provide such spaces to be explored in the context of the upcoming Data Act of 2021. It is of utmost importance that these spaces, their role within the horizontal data spaces as well as the sector-specific spaces to be created, are designed first and foremost from the point of view of the people and their societies about whom that data is collected and used. The MyData infrastructure, described in chapter 2 of this paper and further in the "Understanding MyData Operators" white paper,²⁴ provides a developed and researched basis for implementing such personal data spaces according to the human-centric principles shared by the MyData model and the European Commission.

24 Langford et al. 2020.

2.4 The potential of MyData: Risks and benefits

Thus far we've described the vision of what MyData would change if realised. In this section, we discuss the potential benefits and possible risks involved with implementing the MyData model from the points of view of individual people, organisations working with personal data, and societies collectively. The motivation behind MyData is to create and speed up change towards increasingly fair, sustainable, and prosperous digital societies. Benefits for people and organisations, and through them ultimately for societies at large, are the driving force behind this change, without which nothing will happen. This is why, from the beginning, it is important to find applications of the MyData model that concretely serve people in their daily lives and improve the functioning of businesses and other organisations. Positive collective and societal effects, on the other hand, can be seen as guiding goals. Solutions for the infrastructure and regulation of personal data must be such that they simultaneously generate more personal data use and at the same time steer that use in a direction that allows for the realisation of collective benefits well into the future.

As with any vision for change, MyData also has potential risks. One such risk is that it never becomes reality. Turning the MyData vision into reality is currently the work of a more or less loosely connected network and community of actors. The continuously growing number of companies of different sizes, public administration organisations, research institutions, non-governmental organisations and other organisations working towards the realisation of the MyData vision each come with their own motivations. The appealing concept and practical implementations emerging create expectations for the future and bring more actors, funding, and overall activity to the domain. There are therefore a number of foreseeable factors which can cause the current MyData momentum to wither or to be overrun by some more powerful force.

However, as developments move us closer to the MyData vision becoming a reality and the field matures, some convergence is inevitable and ideas will become more fixed. Certain technology standards will become widespread, certain players gain a significant market position, a "killer app" sets the standard for a working business model, or an internationally accepted governance model and its institutions emerges. Such convergences stabilise the field and the operating environment and allow for the growing adoption of the MyData model. On the other hand, after something becomes fixed, it can no longer be easily changed. Below, we consider also some factors that may either slow down the development and adoption of the MyData model or cause it to be fixed in an unwanted position and thus lead to harmful outcomes.



Figure 2.1 The international MyData slogan “make it happen, make it right” illustrates the two sides of the development of MyData. On the one hand, care must be taken to ensure that the development cycle towards real and functional MyData stays on track and does not wither. On the other hand, it must be understood that development can, as a result of different actors and activities, change direction and, at worst, become fixed in some undesirable position.

2.4.1 People

Most people are not interested in data itself, and they can't be bothered with the effort to manage their own data. Nor are the strengthening of rights or better data protection and privacy the primary motives for the introduction of new types of services or policies. There is a need for fully productised and easy-to-use services that use data to better people's daily lives.

Thanks to data portability, service providers can be switched in an agile manner and the burden of managing one's own data is reduced. As the availability of data becomes easier, companies can specialise in the diverse offering of products to even small customer groups.

The opportunity to share one's own data in a controlled and effortless manner with different organisations affects the relationships between people and organisations, be they customer relationships or other interactive relationships. For example, if employees share data they collect about their own well-being with their employer, companies may have better ways to improve the work environment.

Risks

- **Frankenstein's MyData:** The transmission, aggregation and analysis of personal information is smooth and takes place ostensibly under human control. However, people do not have genuine choices, as broad consents to the use of personal information are sought everywhere. A normal everyday life does not go well unless you give up your own data – like it or not.
- **Unbearable responsibility for the individual:** MyData brings long-awaited digital rights and power to people. At the same time, the responsibility shouldered by the person grows; they must take care of their own privacy and where their data is shared. The market is filled with scam services and data phishers: the former fish for personal data with false promises and the latter give false, misleading or even harmful interpretations and “advice” based on people's data. People's abilities and opportunities to manage their own data matters are unevenly distributed: some do well and others don't. In a society that emphasises individualism, people are left alone with their data.

Benefits:

- **New types of services:** Opportunities for the production of completely new types of services open up when data from different sources can be utilised and combined more easily than at present. People's daily lives would be bettered, for example, by more accurate recommendation systems, targeted health advice and applications to monitor their own finances and to understand consumption choices.
- **Freedom of choice and interchangeability of services:** Easy data portability protects customers from being locked into a single service. People can change service providers and at the same time move their own data to a new location. The situation is comparable to the situation where it became easier to change telecom operators when a customer was allowed to keep their old telephone number.
- **Equity of information:** The position of people in relation to organisations is strengthened when they are offered better means to understand and manage the data that organizations collect about them. Increased transparency makes it easier for NGOs and public authorities to deal with invasions of privacy.

2.4.2 Companies and other organisations

At the level of business models, MyData is a response to a change in which we are currently moving rapidly from silos to value networks through, among other things, data sharing, APIs, and blockchain technologies that enable distributed trust. Human-centric personal data management and new types of decentralised digital identity solutions are essential enablers of networked digital economy business models.

All organisations that handle personal data can promote MyData by, e.g., opening up personal data interfaces. At first glance, the motivation to do this may be unclear, as customer data is a significant competitive asset for many companies. To a certain extent, legislation requires data portability for all, and in this new regulatory environment, proactive participation in data exchange can provide a frontrunner position. Entities collecting data pools benefit if others provide services that complement their offerings based on the same data. For example, if the data collected via a loyalty card can be widely used in different services, then the card will become more valuable to customers and will have a positive effect on card use and customer satisfaction. Actors committed to open practices can also gain visibility and brand image improvement if they stand out as ethical and fair actors.

Risks:

- The continued development of MyData globally requires that a large number of independent actors see MyData's promise for the future as sufficiently interesting and worth pursuing, as well as actually possible to realise. The pursuit of the impossible is not exciting.
- Adopting the MyData approach and implementing its principles offers no first mover advantage.
- Organisations prefer to see and place themselves in the centre rather than as parts of ecosystems where the person is in control of the data about themselves.
- Personal data managed by organisations is often treated with a protectionist attitude and with the view that it's of no benefit outside its original use or organisation collecting it.
- Organisations are afraid of legal sanctions and default to "protecting the people from their own data".

Benefits:

- **Network-like business models:** Human-centric personal data management enables the production of seamless service combinations in the networks of several organisations without a centralised platform entity. The ability of people to decide for themselves who can use their data also provides small actors with equal access to data that is now only accessible to large global actors.
- **Consumer trust:** Trust is increasingly important for businesses. Customers want to know what is being done with their data. MyData increases organisational transparency, helps build trust, and generates image benefits.
- **A smaller number of users is enough:** Today, the success of web services is often determined by the so-called “winner takes all” phenomenon, and many good services die because they do not reach a critical mass of users. MyData services are interoperable and interchangeable thanks to the easy portability of data. This also allows small niche services to benefit from network effects.

2.4.3 Society at large

The use of personal data is most often only talked about from the perspective of either individuals or organisations. However, the collective societal impact of personal data can be more than just the sum of the benefits and harms for individual actors.

Developments in human rights and information technology have largely occurred separately from each other. MyData is a way for a person to take control of their digital rights. A person must have the right and opportunity to control their own digital identity just like they have the right to freedom of thought and expression as a citizen. MyData provides a long-term model for the governance of personal data that serves as a sustainable foundation for digital societies.

At its best, the MyData model can support, for example, societal data collection, social justice, participation, and responsible collective action. In addition, people can receive services that make their everyday lives easier and new business opportunities would open up for companies. Important questions include, how collective action could be used to balance the current data economy and how people can, by sharing personal data, create societal and community value in addition to contributing to the production of economic value.

Risks:

- **Difficulty of discussions:** The realisation of MyData is a typical complex phenomenon. A complex phenomenon is one that is not possible to solve or define in principle, but in which order is created through self-organisation. There are many concepts associated with MyData discussions and the diversity of concepts makes discussions difficult. The lack of open dialogue marginalises the phenomenon.
- **Law and Technology in the Hands of Dr. Jekyll and Mr. Hyde:** Good-faith initiatives such as the GDPR force companies and organisations to open up their systems and make personal data available to people themselves. But as the developments in technology and legislation march on, quantum computers void current encryption technologies and security legislation begins to force companies to give the same access as is granted people also to public authorities and their subcontractors. What was meant for people themselves becomes available to others through technology and coercive regulation.
- **The Battle for the Power Ring:** In the data economy, the big players have an

advantage due to advanced technology and sheer user mass. People, smaller businesses, and even nation states are underdogs. MyData creates freedom of choice for people and balances competition. When the data is not locked, people choose different, small or domestic service providers. But the giants don't give up, but they buy up MyData startups, make supremely usable and free-to-use operator services, occupy standardisation forums, mine the field with technology patents, and lobby lawmakers. Eventually, the giants regain and further increase their power.

Benefits:

- **Societal use of data:** Solving difficult problems is made easier by researchers and decision-makers having tools for more comprehensive data collection. For example, people can give permission to use their own data for research purposes.
- **Legislation simplifies:** The need to legislate about use-specific access rights for different purposes is reduced when a person can authorise the disclosure of their data from public administration data resources to private services.
- **Freedom of choice and free movement of data:** People are increasingly using digital services in their normal everyday lives. Data portability enables the free switching of services and the movement of people and services globally without borders.

This chapter describes the technological vision of a global, human-centric infrastructure for sharing personal data, which is based on open ecosystems. The way this infrastructure is governed affects how easy it is to utilise personal data, how transparent the use of data is, how well its structures support open competition, and how personal data can be utilised in a human-centric way. The key elements here are that people have control over the data about themselves, that there are multiple providers of infrastructure-level services, and that services are interoperable and mutually substitutable.

This is one view of the kind of technical and governance solutions needed to implement MyData principles on a large scale and in a sustainable way. The goal is a reliable and simple service infrastructure which is open to new actors and new innovations. Services will be scalable and can be easily substituted for one another, as data will flow smoothly thanks to interoperability among the different operators and other roles in the ecosystems. For people, this means, for example, easy-to-use services for managing privacy settings, personal data storage services, and services for maintaining personal profiles. For organisations, this means ready access to high quality data and the ability to innovate based on it.

This vision is a synthesis of developments happening in the world, described in more detail in the “Understanding MyData Operators” white paper.²⁵ The vision will not materialise immediately or exactly as set out as such, as developments in this direction are still at an early stage and we have much still to learn. Some ideas, such as the ecosystem roles here described and the importance of distributed identifiers, are widely supported. Future developments in this field are being shaped by evolving legislation, technological and standards developments, companies operating in the field of personal data management and their business models, and the turning tide of public opinion and criticism of the current structures of the data economy.

The vision described here should not be thought of as one big whole that should be implemented in one go. Rather, different elements of it are and can be developed in parallel and largely independently of each other.

25 Langford et al. 2020.



PICK UP HERE



It is essential that people retain control over the data about themselves, that there exist multiple infrastructure providers, and that services are interoperable and interchangeable.

3.1 From platforms to networks

Do we even need an infrastructure to govern personal data? Wouldn't it be easier if APIs and applications simply communicated directly with each other without a layer of data-sharing infrastructure in between? Often, progress has been in the direction of such an organically expanding, infrastructure-free ecosystem where individual connections are made between actors as needed. On the other hand, platforms have emerged where a single actor collects and harmonises data from multiple sources and distributes it further. Both these models are fundamentally organisation-centric, built and operated based primarily on the needs of organisations. In contrast, the MyData model places the person at the centre of the ecosystem as the point of integration between organisations who hold and use personal data about them. In the MyData model, the providers of personal data management services are in competition with each other but form interoperable ecosystems and, together, provide the infrastructure for the transmission of personal data.

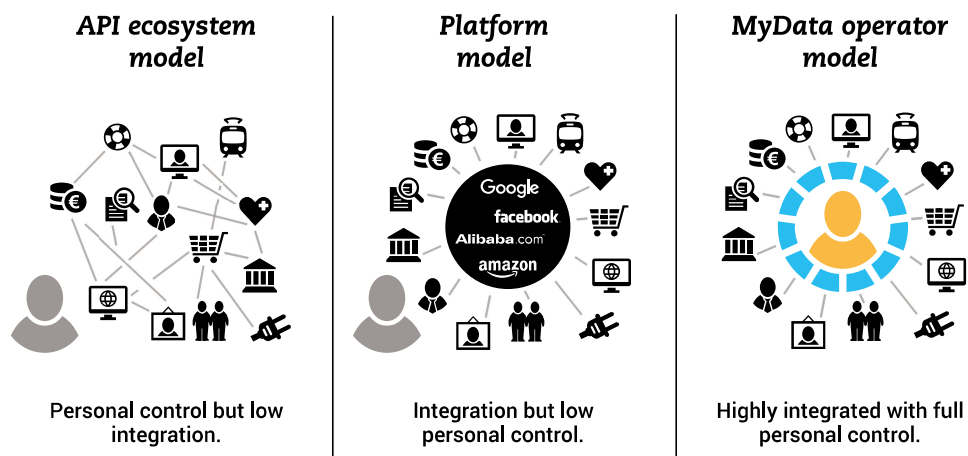


Figure 3.1 In the API ecosystem model (left), if the number of services increases, the number of connections will increase even faster. Centralising data management to platforms (middle) facilitates application development, but there is no incentive for different platform players to seek interoperability between platforms. Compared to the platform model, the MyData operators infrastructure (right) is robust and scalable because it is not dependent on any one organisation providing the infrastructure.

3.1.1 The API ecosystem model

Application programming interfaces (APIs) allow connecting different individual services in an agile manner. The resulting ecosystem promotes the flow of information, creates new business and accelerates digital service development. Generally, the motive for organisations to develop public APIs is to position their own service as central to the broader service portfolio as possible. However, the features and technologies of interfaces vary greatly, which makes the integration of APIs between different individual service providers time-consuming and tedious.

The ease of integrating different APIs will certainly improve over time. From the point of view of human-centric control of personal data, the problem with this kind of an ecosystem is the large number of services, the mesh of connections between them, and the following difficulty for people to understand how data about them is used. The only way to get a complete picture on how data about you flows between services is to log in to each service individually and look for a settings view that shows which other services have been granted permission to read data through the interface. Separate management of a few services is still possible, but with increasing digitalisation, every brand and service wants to establish its own digital customer relationship with the consumer, involving the collection and sharing of data. Suddenly people are involved in hundreds or thousands of managed relationships, and the need for infrastructure to manage data flows and digital relationships becomes evident.

3.1.2 Platform model

In the absence of infrastructure for managing and transmitting personal data based on open standards, individual companies operating globally are expanding their own personal data ecosystems and are seeking to become de facto standards through their sheer user volume. Common to the organisation-centric ecosystems created by these platform economy giants is that the data flows seamlessly within the ecosystem built around the central enterprise platform and the user identity it provides, but only to a limited extent outside of that ecosystem. There is a risk that new players will be prevented from entering the market altogether.

The platform model is also utilised in some sectors to enable the sharing of data collected by multiple actors. The health sector has many examples of this in various countries, including the British Patients Know Best and the Finnish Kanta Service. In such a structure, companies or public authorities establish a common platform for the transmission of data. Centralisation promotes the pooling of data and speeds up the development of new use cases for data, but at the same time the system becomes dependent on a single actor who defines the objectives and ways of doing things.

People's access to and control of the data about them may not be realised with platforms which are primarily designed to support the business models of the central organisation or to facilitate exchange data between organisations within a single sector. For example, online advertising companies operate in networks where few aggregator companies facilitate the movement of personal data, but with the objective is not to meet people's needs, but rather the needs of the ad tech companies in the network.

3.1.3 The MyData operator model

In the human-centric MyData operator model, a person acts as the point of integration regarding the data about them. In this model, a person can control the use of personal data about them across services by granting or denying access to data or by assigning services to act on their data. Some people could run the necessary technology in a secure way themselves, but most will want to rely on external service providers. These service providers offer tools for people and organisations to manage multiple types of personal data coming from multiple sources.

The key difference with the platform model is that, in the MyData model, there are many service providers that together form ecosystems in which personal data is shared with high levels of trust. The infrastructure is not based on centralising user information to a single service, as in the platform model. Instead, the participants in the ecosystems have common standards, policies, and governance that enable interoperability and data portability. This could be compared to, say, a network of banks. Instead of being able to transfer payments only between its own customers, the banks form an international network where payments can be made between customers of different banks.

3.2 Human-centric personal data ecosystems

The MyData vision, which highlights competition and open ecosystems, is based on the premise that there will be a large number of actors providing personal data transfer and management services, and that they must be interoperable and substitutable. Competing service providers work together to create globally interoperable ecosystems for human-centric personal data transfer in a similar way as banks create a network for payment transfers, telephone operators for call transfer, and e-mail messengers for e-mail transmission. This section describes the MyData vision for human-centric personal data ecosystems and the roles different actors can have in such ecosystems.

3.2.1 Ecosystem roles

Human-centric personal data ecosystem is composed of actors holding one or more of the following main roles:

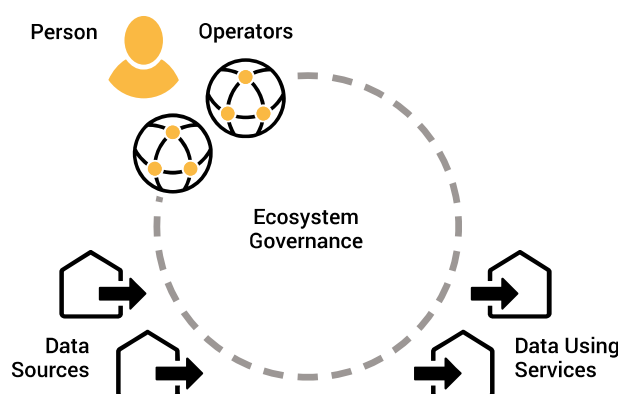


Figure 3.2 Roles in personal data ecosystems.

Person: The role of data subject as represented digitally in the ecosystem. Persons manage the use of personal data about themselves, for their own purposes, and maintain relationships with other persons, services, or organisations.

Operator: The role responsible for operating infrastructure and providing tools for the person in a human-centric system of personal data exchange. Operators enable people securely to access, manage, and use personal data about themselves as well as to control the flow of personal data within and between data sources and data using services.

Data Source: The role responsible for collecting, storing, and controlling personal data which persons, operators, and data using services may wish to access and use.

Data Using Service: The role responsible for processing personal data from one or more data sources to deliver a service.

Ecosystem Governance: This role is for actors that are responsible for managing, developing, and enforcing the governance frameworks for the ecosystem

Different kinds of actors like governmental organisations, private companies, and even individual people can take the roles of operator, data source, data using service, or ecosystem governance.

In addition to personal data, data transmission within the network may also include non-personal data related to companies or objects, for example. Technologically speaking, there is no significant difference in data transmission depending on whether the data is personal or not, but applicable regulation is different when it comes to the processing of personal data or mixed data sets containing personal data.

Naturally, other players, such as regulators, standards organisations, investors, media, etc., also influence how the ecosystem functions.

3.2.2 Specialised data operators

In such personal data ecosystems as described above, specialised MyData operators emerge who are able to build a sustainable and profitable business through offering infrastructure for personal data transfer, as opposed to utilising the data itself. People might have trusted MyData operators for different types of data: health data, property data, consumer profile, mobility profile, etc. Some operators may specialise in certain industries and others may be generic, while operators could also specialise in particular functional areas, for example, consent. If desired, a person could also manage all of their data through a single operator.

The role of an operator in ecosystems can be assumed by any actor, and at least in the first use cases, this role is most often performed by an actor with also the role of either data source or data using service. However, from both technological and business points of view, the management and transmission of personal data and the related services should be separate, or at least be separable from, the business and technology of data processing that results in the end service provided. This principle is referred to as Separation of Concerns (SoC) by which a modular approach to the development of a system is adopted. This approach entails each section addressing a different aspect (concern) of the overarching system. In the context of SoC in the personal data ecosystem, processing, storing, aggregating, displaying, governing data are concerns that need to be managed in a modular, transparent manner. SoC enables more opportunities for module upgrade, reuse, and independent development and allows for more transparency and good governance.

The MyData model will also facilitate the implementation of services that utilise personal data as application developers would not need to separately implement common features, such as permissions management, in their services. For developers, networks of operators offer a sufficiently large potential customer pool, and there will be no motivation for neutral operators to close the market, for example, by supporting only their cherry-picked applications.

High-level scenarios for organising personal data infrastructures

MyData operators serve a role in the creation of sustainable and human-centric data management infrastructures for personal data ecosystems. Different ways to organise personal data infrastructures exist and some of them are more aligned with the MyData principles than others. These are not to be considered mutually exclusive, as co-existence and hybrids are possible. In the MyData community, there is strong support for the competition-based scenario. However, it is possible for the last two scenarios (decentralised and competition-based) to co-exist without compromising the MyData principles.

Fragmented: Markets where many small operator-like entities compete to build small-scale use cases without interoperability between them.

Monopolistic data platforms: A few platforms provide connectivity and data sharing inside their ecosystems with little competition and no incentives for interoperability between the platforms.

Fully decentralised: A peer-to-peer world where standardised technical infrastructure and protocols enable data connections without any specific operator entities. In the decentralised model, the individual manages data flows directly from the end services or by having personal cloud-based applications on their own devices or hosted for them.

Competition-based interoperable operator network: Similar to the current network of telecom operators, energy providers, or banks where many mutually competing providers are interoperable and together provide global-level connectivity.

3.3 MyData operator reference model

The MyData model does not depend on or even recommend any single technology or type of technology. Rather, it is a model for how to deploy different and complementary technologies in a human-centric fashion. The MyData operator reference model describes nine core functional elements of operators. These elements affect how easy it is to utilise personal data, how transparent and human-centric the utilisation of personal data is, and how well the infrastructure supports open competition. The technologies and their deployment for each of the functional elements here described are being developed also independently of the MyData operators context.

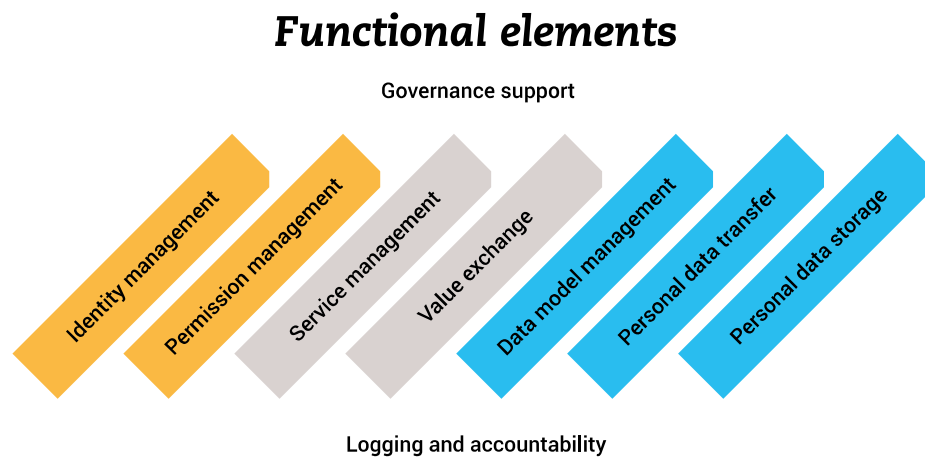


Figure 3.3 The MyData operators reference model describes nine functional elements of MyData operators.

3.3.1 Identity management

Identity management handles authentication and authorisation of individuals and organisations in different, linked identity domains and links identities to permissions.

Digital trusted and easy-to-use identity is an important enabler of a functioning personal data ecosystem. Managing the identities of individuals and confirming identities of other actors in the ecosystem makes it possible for individuals to act as the 'point of integration' regarding data about them. Individuals can have different identities, or profiles, with different data sources and data using services. For example, they can have public as well as private identities, or self-sovereign identities.

Decentralised identity solutions independent of a single identity provider enable MyData service substitutability and interoperability of authentication and authorisation solutions across services. One such solution, the concept of self-sovereign identity (SSI), is well aligned with human-centric personal data management.

3.3.2 Permission management

Permission management enables people to manage and have an overview of data transactions and connections and to execute their legal rights. It includes maintaining records (notices, consents, permissions, mandates, legal bases, purposes, preferences etc.) on data exchange.

People find it difficult to learn the terms of use and manage their data sharing and privacy settings. If in the future people also need to increasingly control the disclosure and authorisations of data about them, the user experience around privacy settings, contract relationships and permission control to personal data sharing must be significantly improved.

Permission management covers the technical functionalities required for human-centric control of personal data, such as the user interfaces and underlying data structures for individuals to view, understand, grant, revoke, and modify different kinds of permissions related to data flows.

The term 'permission' is used in a broad sense to cover the means that the individual has to take control of data flows. These means can be based on legislation (executing legal rights) or go beyond that. All the considerations of consent, notice, preferences, and permissions are captured by the wider definition of 'permission' used in this functional element. Part of the permission management functionality is that the operator only allows execution of such data transactions where the permission is valid.

MyData account as a standard operator service

As personal data ecosystems mature, progress towards increasing interoperability, and become more standardised, portable MyData accounts become possible. The account is a metaphor familiar from bank accounts, email accounts, and customer accounts, and embodies the concept of substitutability of service providers in the operator network.

A MyData account is a one-stop shop view and a place to manage data about one. It also provides an overview of connections with different applications and services. Currently, decisions about how to share and use personal data are made, for example, when taking into use new applications or later through changing the related privacy settings in those applications. However, through a MyData account, all permissions and various privacy settings you grant are stored in the user's MyData account, wherever they are made. In the same way, for example, online banking is a centralised view of one's own money and payment transactions, although most payments are made by debit card in a location quite different from the online banking user interface.

The goal of MyData infrastructure is for open standards, policies, and governance to allow interoperability between different personal data ecosystems, operators, and MyData accounts. It is technologically possible for a person to maintain their own MyData account and thus independently to connect to personal data ecosystems. However, with the MyData model scaling to a large number of users, the majority will rely on operators providing this service on behalf of end users.

One of the key design principles of the MyData infrastructure is that the MyData account should be easily transferable from one operator to another without losing the content of the account, or losing the permissions and connections made through the account. Account portability increases trust in the ecosystems and opens up markets to competition.



Possible features of MyData accounts:

- Account Statements: A clear view of the overall use of the personal data about the account holder.
- People can have multiple MyData accounts, and some of them can be shared by several people.
- Accounts have different levels of authentication and assurance levels, for example, some of them are strongly linked to a person's identity or electronic identity (eID), some are pseudonymous accounts that are identified only by their email address.
- A MyData account controls the disclosure of data and automatically maintains a record of what data has been disclosed.
- MyData accounts and their associated identities can be linked and combined to each other.
- A person can manage accounts directly or through separate interfaces (compare email clients that manage multiple email accounts).
- Remote management of privacy settings
- Revocation of permissions and deletion of data from services using it.
- Methods for individuals and organisations to connect to personal data API's and identify themselves as required.

3.3.3 Personal Data Storage

Personal data storage allows data to be integrated from multiple sources (including data created by a person) in personal data storage (PDS) under the individuals' control.

Although permissions management could allow data to be transferred directly from one service to another without the need for it to be stored, there are many use cases where it is useful for a person to store and archive information for himself and to act as a distributor for other services.

Many proto-operators provide personal data storage (PDS) for storing data originally created at the data sources and data created or asserted by the person. Such PDS functionality allows data to be integrated from multiple sources – harmonising, using and re-sharing it under the individuals control. The PDS can be implemented so that the operator does not have access to and does not know what data the person actually stores.

Using PDS as an 'intermediary station' for personal data configures the connections in the data ecosystem so that data sources and data using services can connect via the person, but not directly to each other. This configuration may simplify legal liabilities as well as the implementation of permission management.

3.3.4 Service management

Service management uses connection and relationship management tools to link operators, data sources, and data using services. Data can be available from different sources and can be used by multiple data using services.

Operators live in an ecosystem with data sources and data using services. Navigating this ecosystem requires the linking of actors through an operator: this is the purpose of the service management functionality. The human-centric manifestation of service management is the possibility for individuals to manage the relationships and connections to different data sources and data using services in the ecosystem.

Service management enables dynamic linking of data sources and data using services (permissioned by the person) so that data can be available at different sources and can be used by multiple data using services.

3.3.5 Value exchange

Value exchange facilitates accounting and capturing value (monetary or other forms of credits or reputation) created in the exchange of data.

Sustainable business models are a requisites for ecosystems in general.²⁶ This means that all of the actors in the ecosystem need to have more benefits than costs in the long run. Both benefits and costs can be also non-monetary in nature. For people, time and effort spent can be a big cost and benefits often come in the form of services. Personal data ecosystems exist to lower transaction costs and, in total, an ecosystem enables the creation of more value than the overall costs incurred by the participants collectively. However, value creation does not happen equally in all parts of an ecosystem and mechanisms for distributing value are needed.

As operators provide technical infrastructure for making multi-party data transactions possible, they are also in a natural position to keep track of such transactions for the purposes of payments and billing or creating other forms of rewards, such as loyalty and bonus points. Operators may provide a standard 'accounting' mechanism which transparently keeps a log of the data transactions so that the different parties in the ecosystem may use it as the base for payments between the parties.

3.3.6 Data model management

Data model management is about managing the semantics (meaning) of data, including conversion from one data model to another.

Semantic standardisation is the process of describing the meanings of data content so that data about the same subject originating from different sources can be easily combined and understood. Human-centric management of personal data cannot be achieved if all organisations ask for the same information, such as contact information in slightly different formats. If each service has different data models and concepts, then cross-organisational data sharing and related permissions management is cumbersome and expensive.

Harmonisation of personal data models strengthens options and potential for data portability and increases usability of data. Data models related to data transactions also need standardisation to achieve interoperability between operators in, for example, log data syntactics and permission models. Depending on the domain, semantic data standards are more or less evolved. Until widely adopted standards exist, translations between different data models are a necessity.

Data model management as an operator functionality facilitates translation of one data model to another. As many data standardisation processes are not human-centric in nature, data model management as an operator functionality can also facilitate interpreting standard data models to individuals. Personal data management without data model management is possible, but limited in terms of scalability, interoperability, and usability of the data.

²⁶ Haaker et al. 2006.

3.3.7 Personal data transfer

Personal data transfer implements the interfaces (e.g. APIs) to enable data exchange between the ecosystem participants in a standardised and secure manner.

Personal data transfer, through an operator or facilitated by an operator, is key to portability, access, and re-use of personal data. This functionality realises the interfaces to allow data exchange between data sources, data using services and operators in a standardised and secure manner. Data transfer can follow different models: data can flow through an operator, or an operator can facilitate the direct transfer from data source to data using service under a valid permission.

'Data sharing' is a catch-all term that hides a multitude of variations. In most cases, there will be an original or 'master' version of the data that may be held by an organisation or an individual. Operators need to manage the transfer of personal data in line with permissions to ensure that data is not unnecessarily duplicated and can be updated easily across any copies when required.

3.3.8 Governance support

Governance support enables compliance with the underlying governance frameworks to establish trustworthy relationships between individuals and organisations.

Human-centric governance helps mediate the relationships between people and organisations. This dedicated functionality in an operator can guarantee that MyData principles are followed and enable compliance with underlying governance frameworks.

All operators, to some degree, need to operate within a framework of governance in order to be transparent about assurances to individuals concerning the quality and trustworthiness of their services. Operators may be able to select governance frameworks within which to work or they may have to respond to mandatory requirements within their sector and jurisdiction. Governance requirements translate into responsibilities for the operator which can then, in a well governed ecosystem, result in liabilities. In a governed transaction, a specific liability can be agreed upon or transferred. The governance support element contains the functional counterparts of the ecosystem governance frameworks discussed later. Operators may enhance and deliver this functionality.

3.3.9 Accountability and logging

Logging and accountability entails keeping track of all information exchanges taking place and creating transparency about who accessed what and when.

Transparency and accountability are important principles and prerequisites in many legislations. Accountability can enhance assurance and logging can mitigate risks of misuse or unintended use. Logging is not the sole responsibility of the operators and has counterparts in data sources and data using services.

In general, governance implies some accounting obligations; but if no explicit governance applies, logging and accountability are still needed for auditability and transparency.

Interoperability

Interoperability provides overall system benefits at different, distinct dimensions that can and should be developed concurrently: technical (connectivity), semantic (informational), and organisational (governance, business models etc.).²⁷

Technical level: Definitions of connectivity, syntactics, and protocols for data exchange (e.g., APIs) and data storage that underpin basic integration. The first objective here is to enable the easy connection of new data sources and data using services to an operator and their mutual interoperability, where operators can work with each other technically.

Semantic level: Harmonised information with shared data models and mutually agreed content. The pragmatic approach here is to identify the categories of data where common data models are most essential for MyData. These could be semantic models for data control and governance (e.g. data transaction records, consent records purpose categories) or widely used attribute data types and domain specific data models.

Organisational level: Interoperability in more mature ecosystems goes beyond the technical and semantic levels, encompassing shared objectives and policies between organisations. These objectives and policies will cover issues such as responsibilities, liabilities, business models, and governance structures.

²⁷ Tolk 2010.

4. Examples of MyData application areas

Implementing and maintaining MyData requires investment and it changes the business of utilising personal data. Who pays and benefits? Are people really interested in MyData services?

The overall benefits of mature MyData markets have been fairly widely recognised and acknowledged. For example, the reduction in duplicated data collection, the emergence of seamless digital services, an open competitive environment, and the stronger legal position of people are spurring the development of MyData solutions in organisations around the world.

It is important to note that half-way MyData doesn't generate half the benefits of a full implementation. Markets work if data sources, data using services, infrastructure providers, and people themselves each benefit more than their costs are in the long run. If one of these parties is not involved, data will not move and no one will benefit. On the other hand, if incentives are successfully implemented and ecosystems emerge, then network effects can accelerate their growth rapidly.

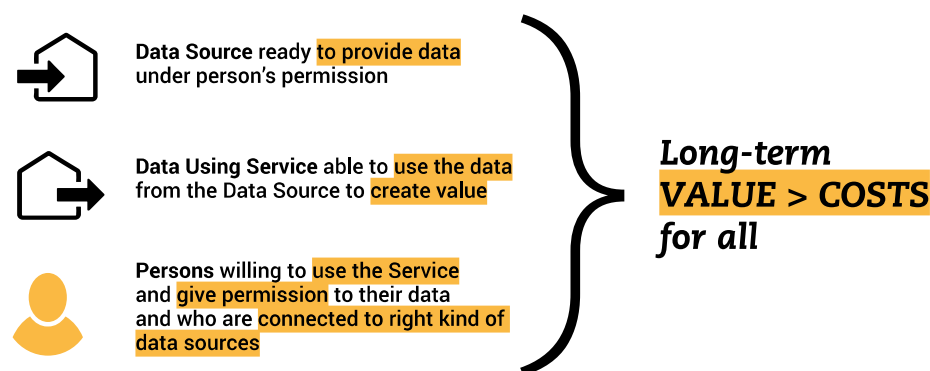
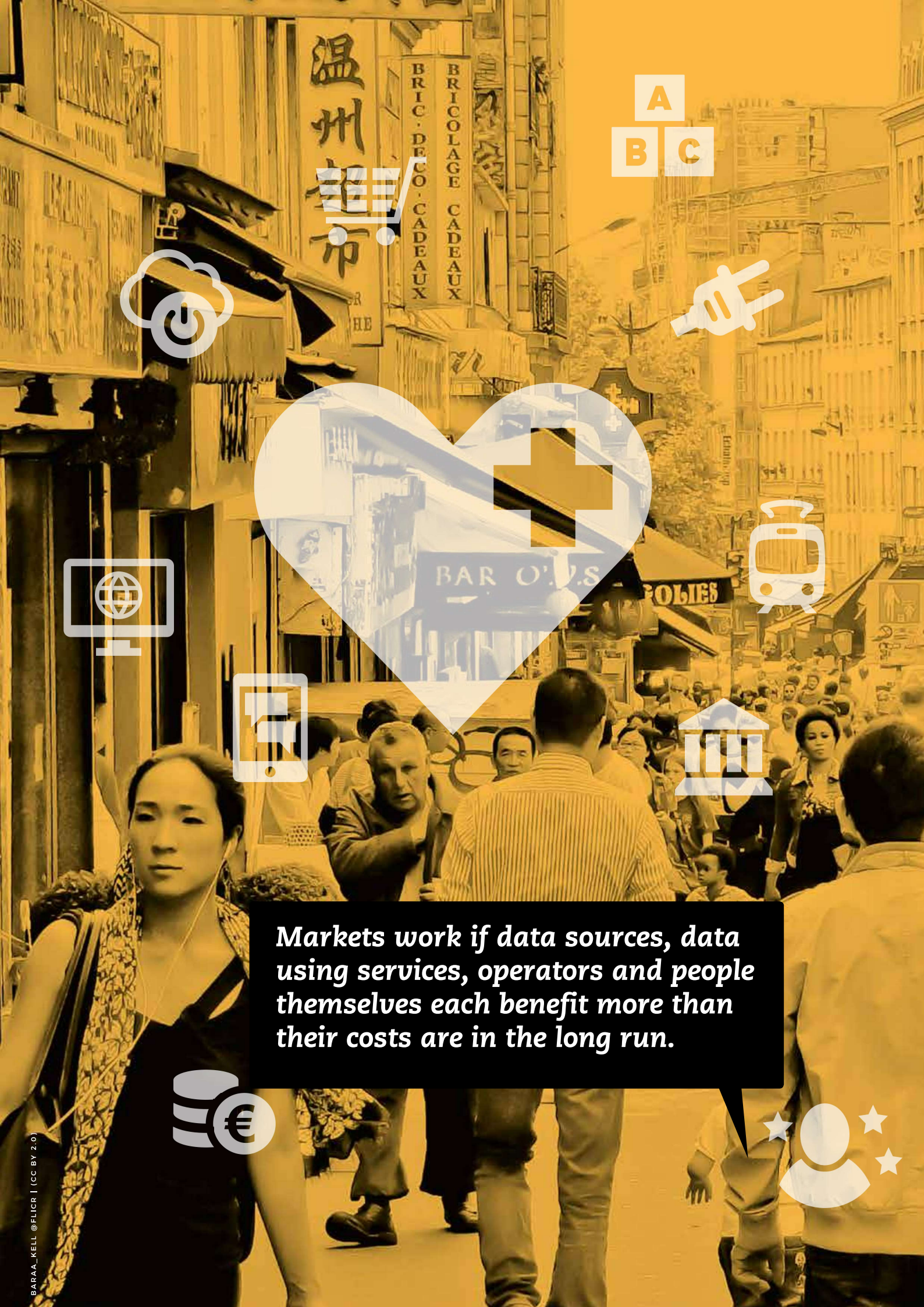


Figure 4.1 The sign of a viable MyData use case is that all parties involved benefit more than they incur costs in the long run and profits are distributed across the different parties. That is to say, even if the total benefits of a use case outweigh its total costs, it is further necessary that benefits and costs are distributed in such a way that it is profitable for all parties to participate. The networks of MyData operators described in the previous chapter can, among other things, also act as networks for the exchange of value and sharing benefits for example by reimbursing data sources for the costs they incur.

Costs and benefits are not automatically distributed evenly among different roles and actors. There is a strong assumption that there is a demand for personal data to begin with, i.e., that there exist services based on the use of personal data and that there would be more of them if such data were more easily available. On the other hand, for data sources, moving towards MyData can mean costs and risks. For example, the data contained in a given public register may be in high demand, but the agency responsible for the register might not significantly benefit from, or may even lose revenue by, providing access to the data for people, as it requires also investment and effort to maintain quality control of and interfaces for the data.



A
B C

Markets work if data sources, data using services, operators and people themselves each benefit more than their costs are in the long run.

It should be reiterated that the MyData infrastructure described in the previous chapter is not only a network for data transmission, but also a network for services, value generation, and exchange. MyData operators could maintain smart contracts for data transmission that allow for automatic balancing of costs and benefits in the network. For example, if a data using service pays for data, this revenue can be seeded among the operators and the data sources who maintain that particular interface. These payment models of the operator network have not yet been tested with MyData, but other networks, such as banking and telecommunications, can be used as models.

For people, MyData promises, among other things, a better user experience in digital services. The same information does not need to be entered and updated in many places, services are more automated and personalised, and so on. Successfully creating a better user experience depends on the usability of the implementations. Operator network payment models should also act as an incentive to develop the usability of data management. Operators with easy-to-use services will gain customers and a larger market share. The precondition, of course, is that switching operators itself is sufficiently straightforward and does not become a blocker for competition.

The following sections describe six examples of ways in which MyData can change the relationships between people, data sources, data using services, and operators and bring benefits to different parties. These examples have been chosen to provoke thoughts and to evoke new ideas of the various different opportunities for personal data use which MyData thinking and implementations enable.

4.1 Self-assembled profiles: Portable media profile

Poorly targeted recommendations and online ads are commonplace, but sometimes recommendations fit surprisingly well. From a customer's perspective, the very same fitting recommendation can come across as great digital service or as scary surveillance that undermines trust. This is the so-called creepy-cool line. On which side of the line the profiling service lands depends in part on whether the customer understands what data the recommendation is based on and how the service provider knows those things about them.

Profiles that are self-managed by people themselves allow for more accurate recommendation systems and the transparency they bring contributes towards trust in companies who use or provide targeted advertising and recommendations. Such profiles are potentially much more versatile and accurate than those produced by a single company, as they can combine data from different sources. With MyData, a person can also use the same, rich profile data in multiple services instead of each service devoting resources to creating a separate customer profile of them.

For example, a portable media profile will allow the transmission and utilisation of one's media preferences and behaviour data in several services. A user could continue to watch a video that was interrupted in one media service from the same point in another service, and the data gathered by both services would enrich the same media profile, allowing both services to provide more appropriate recommendations in the future. Further, different types of media providers could access a more holistic and rich profile of their customers by allowing for data of different types of media consumption to be combined. The books I borrow from the public library, the articles I read in the daily newspaper, the music I listen to while doing both, and the video content I watch online, all combined will allow for much better customer insights and thus better recommendations for the person in question.

Similarly, a mobility profile could be shared with Mobility as a Service (MaaS) providers and a health profile would facilitate interaction with different health and wellbeing service providers. Other possible custom profiles could be, for example, a contact profile and a privacy settings profile.

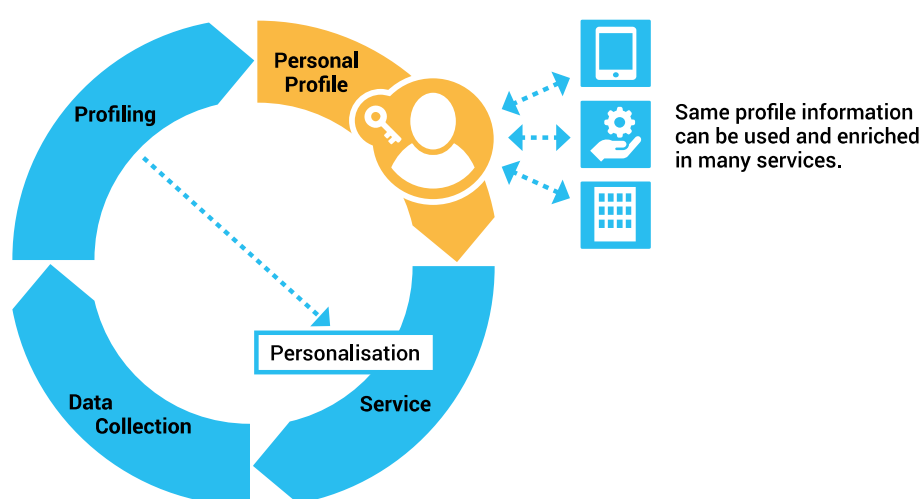


Figure 4.2 Traditionally, profiling and service personalisation based on the collection and analysis of user data takes place within a single service (blue arrows), and as such the accumulated profile data cannot be used elsewhere. At the same time, user data generated in other services is not available for enriching or refining profiles so created. The self-managed profile information (orange), built according to the MyData model, would allow the compilation of data from several sources and the utilisation of the same profile in different services.

4.2 Verified data: CV 2.0

MyData allows people to access and use the data about themselves regardless of the original data collector or provider. This increases flexibility when data using services do not need to be in direct contact or contract with the data sources. However, in many use cases, the flexible transfer of data alone is not enough, and it is also necessary to ensure the provenance and veracity of the data. Such certification of information is possible with the help of electronic signatures.

For example, future recruitment and HR services could work with competency profiles provided by people according to the MyData model. We can think of such a competence profile as a new type of digital CV, in which one's own competences are not only described in machine-readable format, but it is also possible to verify academic and other qualifications electronically. A recruitment service could, in this way, confirm that information about a degree has actually come from the records of an educational institution and has not been tampered with. In addition to academic records and degree data, it would also be possible to verify driving license status, language certificates, hygiene passports, work permits, various courses, and the like electronically if the MyData infrastructure for it is easy enough for data sources to adopt.

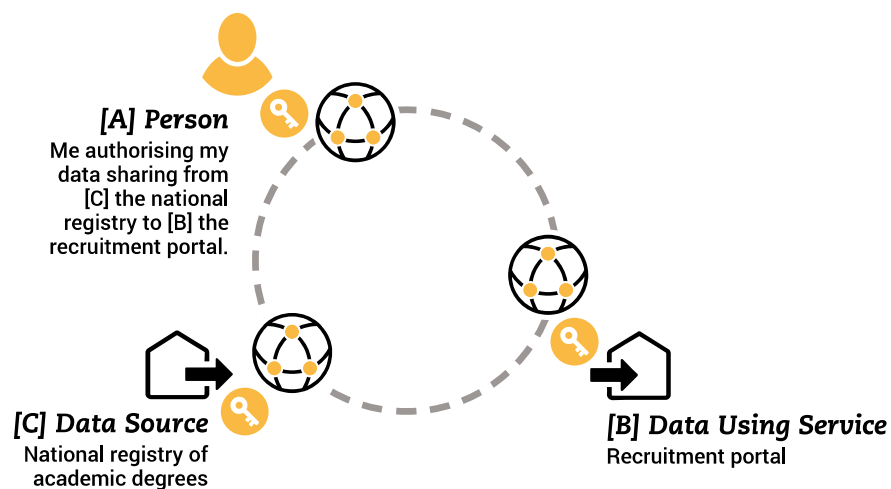


Figure 4.3 In this example case, the jobseeker [A] completes their competence profile (e.g., CV information) in the recruitment portal [B], where it is possible to import, in machine-readable format, certified data from the national register [C] about one's study history. Instead of a centralised platform solution, according to the networked MyData model this could be implemented in such a way that both the data source, the data using service, and the person themselves connect to the network through different service providers or MyData operators.

4.3 Distributed utilisation of user data: Data from a retail loyalty card

Four out of five Finnish households have a retail chain loyalty card. People benefit from using them primarily in the form of bonuses and discounts, but possibly also in the form of more relevant services. Retail chains, like many other services, use the data they collect to optimise logistics, to target supply, and in marketing. Retail chains are developing new e-service offerings that also rely on purchase data, but at the same time they act as gatekeepers as to who can take advantage of the data they collect. Loyalty programme data is only usable in the companies' own services.

MyData includes the idea of decentralising or distributing service chains, where different actors specialise in different parts of the service chain. If people could easily transfer their purchase data from loyalty programme systems to other applications, this would enable more agile service development. For example, one company could create applications that guide groups of customers suffering from allergies or chronic illnesses in their food purchases and another company could compile customer purchase data extensively for personal financial management purposes. Serving these detailed customer needs may not be very high in the priorities of the retail chain's loyalty programme development, but the implementation of these services would increase the usefulness of the loyalty card for customers and thus also be in the interests of the retail chain.



Figure 4.4 With the permission of the user, companies can obtain rich profile information from the person on the basis of and in return for which the company is able to provide the customer better services and service-related communication.

Business and customer information about each other

Customer data management is part of any business. Data about corporate customers is usually somewhat manageable and customer data can be bought from directory companies. The smaller the individual customers are in terms of their purchase volume, and the more there are of them, the more challenging customer data management becomes. In consumer retail, maintaining an up-to-date customer registry is already nearly impossible. Many companies might even want to get rid of managing and maintaining their own customer registry. For them, it would be enough to receive up-to-date data electronically from the customer themselves at the start of a customer relationship, either directly to their own system or even so that the customer data is not even recorded but rather processed and used only at the relevant moments in the customer relationship.

Today, consumer data about a particular person exists in multiple companies' customer relationship management (CRM) systems. Correspondingly, people could have tools to manage their relationships with companies – from auto shops to hairdressers – in vendor relationship management (VRM) systems. Such a vendor registry (whether vendors of goods or services) could contain contracts, warranty receipts, and a contact history with each company. For example, when a person moves from one apartment to another, they could automatically share their new contact details with all companies they want to stay in contact with. Customer-managed VRM and the company-managed CRM systems could complement each other and exchange data as necessary.

Self-managed customer profiles and vendor registers will also make it easier for consumer-customers to compare companies better than currently. Instead of the customer spending time on hunting for the best deals, they can announce their intention to purchase by attaching it to their public profile and receive personalised offers directly from companies. Such systems based on customer purchase offers exist in some individual industries. For example, in the Finnish transport providers' *tilausajot.net* service, the customer fills in the information about the time and route of the coach ride they need and receives offers via email from transport companies. Such an inverse model is called the intention economy.²⁸

28 Searls 2012.

4.4 Societal data collection: MyData and research

Societally significant research often requires the collection of data from large quantities of people and often also from multiple data sources. However, legislation governing the use of public data pools for research varies and the means of obtaining this data are tightly regulated in most countries in the EU. Data resources useful for research, such as data from telecom operators on people's movements and geolocations, are also accumulated in registers kept by the private sector. In the future, research and other methods of addressing shared problems will require new ways of collecting data.

These new ways must also include citizens' participation by way of permitting the use of their data or contributing to joint decisions regarding the use of that data. For example, the Finnish Biobank²⁹ collects samples and data for health research on the basis of individuals' permissions. There can also exist a number of different ways in which data-transmitting organisations can share the added value generated through research. For example, the Swiss Midata.Coop³⁰ is a cooperative whose members jointly decide on the allocation of the revenue generated from data sharing, e.g., for further research or education.

²⁹ <http://www.biopankki.fi>

³⁰ <https://midata.coop>

4.5. Data as part of interaction: Transparency of public services

By utilising personal data about people, services can be automated and provided at the appropriate time to those who need them. When automating services, it is important to make them transparent. Users must have effective means to correct any errors or omissions in the background data used that could lead to erroneous conclusions in the automated process.

Traditionally, data is in the back-end systems of organisations, which is why customer and customer service views of the data differ from each other. Due to the asymmetry in both the amount of data and its visibility, it can be difficult for the customer to understand the decisions that affect them and the reasons for those decisions (e.g., whether a bank loan application is approved).

In the MyData approach, data is part of the interaction between a person and an organisation and the person has equal access to the data about themselves as the organisation. This allows people easily to check the accuracy of the data about themselves and also, according to MyData principles, to transfer that data for other applications. In the case of public services, the citizen could have the same view of the data as the agency that makes decisions about them. For example, an applicant for student support could see directly from the relevant agency's online service what income data the current student support decision is based on and how additional income would affect their situation vis a vis receiving further support.

One of the possibilities of utilising the MyData model is proactive service offering where services are targeted, made more proactive, and automated by using personal data. From the service provider's point of view, proactive service provision can increase customer satisfaction and streamline processes, for example by distributing demand more evenly. In Finland, the transition from tax returns forms, which each taxpayer had to fill for themselves and submit for tax authority review, to taxation proposals automatically generated by the tax authority and submitted for taxpayer review, is a good example of proactive service provision. It utilises the data about people already held by the taxpayer and, as an operating model, both makes people's daily lives easier and makes the tax authority's operations more efficient. Similarly, people whose passports are expiring could be actively provided with an electronic service for passport renewal. This would smooth out the peaks in pre-holiday periods for passport renewals.

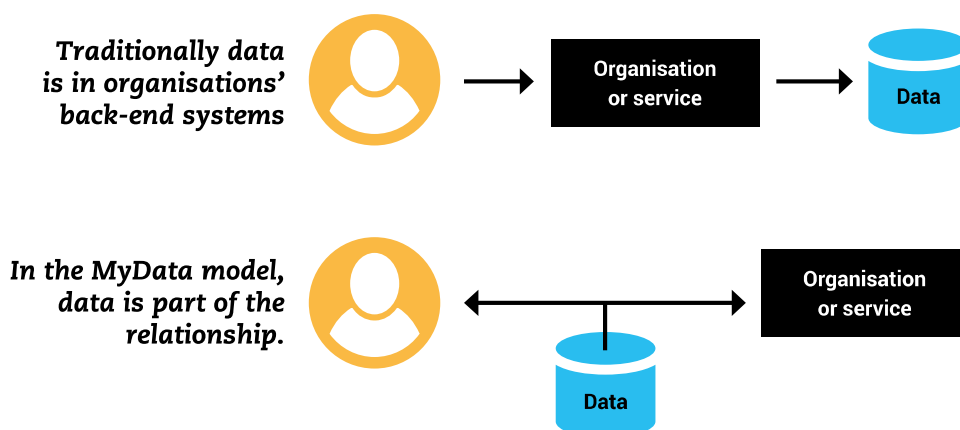


Figure 4.5 In the MyData model, personal data is a shared resource and part of the interaction between a person and an organisation.

4.6 Data about things: Is the data about my car about me?

Data is not only collected in web services and mobile phone applications. More and more devices we use in our daily lives also store data, are connected to the internet, and communicate with each other: an Internet of Things (IoT) is no longer only in the future. IoT devices generate a wealth of real-time data, some of which can be classified as personal data. The proliferation of sensors and data flow-based functionalities in consumer products raises urgent questions about the management and use of that data. Is it acceptable that the data collected by IoT devices is only available on the manufacturer's platform?

For example, new cars store and share a wealth of data about the location and operation of the vehicle as well as the driver's driving style. Currently, this data is controlled by the car manufacturers, who determine who has access to the data. Restrictions thus imposed have practical implications, among other things, for who can service the car. Implementing MyData principles for vehicles would mean that it would be clear to people what data the vehicle collects and for what purposes it is shared. When data is passed on to service providers according to people's own choice, they will be able to choose for themselves how to maintain their car.

Data from multiple people, devices, and spaces

MyData thinking is easily confined to natural persons. Data is, however, usually generated and utilised in and by the interactions of two or more parties. People, organisations, and increasingly also objects and physical spaces can be involved. In such cases, different people and other parties must have their appropriate rights regarding the same data. It is often difficult to identify the only one person about whom some data exists. For example, the data of a sensor installed in a car applies to everyone driving the car, shopping is often done for the whole household even if there is a single payer, and the actual user of a telephone line can be someone other than the owner of the subscription.

The identity with which data is linked may be a person or organisation, but it might as well be an object or physical space. For example, a device or a space, say an apartment, could carry its own history and data on repairs made to it. It would also be justified for access to the device's or space's "own" data to be transferred when the device or space is bought or sold or otherwise changes owners.

Simple cases where it is clear that the occasion involves the personal data about a single person are not and will not be the most common ones when it comes to personal data. Cases involving multi-person data, organisation data, and device and space data, with their particular requirements for data sharing and management infrastructure, are more challenging to solve. On the other hand, generalisable and scalable solutions can spread even faster and wider.

Antikainen, Janne, Jarmo Eskelinen, Marc de Vries, Heli Koski, Tommi Niemi, Mika Pajarinen, and Pyykkönen Sinikukka. 2016. "Massadatasta Liiketoimintaa Ja Tehokkaita Julkisia Palveluja." <http://tietokayttoon.fi/julkaisu?pubid=10701>.

Cavoukian, A and Castro, D. 2014. "Big Data and Innovation, Setting the Record Straight: De-identification Does Work".

European Commission. 2017. "Guidelines on the Right to Data Portability - Article 29 Data Protection Working Party." https://ec.europa.eu/newsroom/document.cfm?doc_id=44099.

European Commission. 2017. "Staff Working Document on the Free Flow of Data and Emerging Issues of the European Data Economy." <https://ec.europa.eu/digital-single-market/en/news/staff-working-document-free-flow-data-and-emerging-issues-european-data-economy>.

European Commission. 2019. "Guidance on the Regulation on a framework for the free flow of non-personal data in the European Union". <https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=COM:2019:250:FIN>.

European Commission. 2020. "European Strategy for Data." https://ec.europa.eu/info/sites/info/files/communication-european-strategy-data-19feb2020_en.pdf

FTC. 2014. "FTC Recommends Congress Require the Data Broker Industry to Be More Transparent and Give Consumers Greater Control Over Their Personal Information." <http://www.ftc.gov/news-events/press-releases/2014/05/ftc-recommends-congress-require-data-broker-industry-be-more>.

Goode, Lauren. 2017. "Carnegie Mellon Researchers Want to Fix App Permissions Once and for All." The Verge. <https://www.theverge.com/2017/2/10/14562514/cmu-privacy-assistant-app-mobile-app-permission>.

Kantara. 2017. "Consent Receipt Specification - WG - Consent & Information Sharing - Kantara Initiative." <https://kantarainitiative.org/confluence/display/infosharing/Consent+Receipt+Specification>.

Knuutila, Aleks, Vesa Kokkonen, Heikki Sundquist, Ossi Kuittinen, and Salla Thure. 2017. "MyData Muutosvoimana: Julkishallinnon Henkilötiedon Ihmiskeskeisen Hyödyntämisen Mallit Ja Vaikutukset." Valtioneuvoston kanslia. <http://urn.fi/URN:ISBN:978-952-287-446-7>.

Langford, J., Poikola, A., Janssen, W., Lähteenoja, V. and Rikken, M. (Eds.). 2020 "Understanding MyData Operators". MyData Global. <https://mydata.org/wp-content/uploads/sites/5/2020/04/Understanding-Mydata-Operators-pages.pdf>

Narayanan, A and Felten, AW. 2014. "No Silver Bullet: De-Identification Still Doesn't Work."

Newman, Nathan. 2013. "Taking on Google's Monopoly Means Regulating Its Control of User Data." http://www.huffingtonpost.com/nathan-newman/taking-on-googles-monopol_b_3980799.html.

Ohm, P. 2010. "Broken Promises of Privacy." 57 UCLA L. Rev. 1701.

Pitkänen, Jyrki. 2018. "Itsehallittavan Identiteetin Sääntely EUn Yleisessä Tietosuoja-Asetuksessa." Lapin yliopisto, Oikeustieteellinen tiedekunta.

Pitkänen, Olli. 2014. "Sinun Tietosi Eivät Ole Sinun: Rekisteröidyn Oikeus Hyödyntää Omia Henkilötietojaan." Oikeus, no. 2/2014: 202-14.

Poikola, Antti, Kai Kuikkaniemi, and Ossi Kuittinen. 2014. "My Data - Johdatus Ihmiskeskeiseen Henkilötiedon Hyödyntämiseen." Ministry of Transport and Communications. <http://urn.fi/URN:ISBN:978-952-243-418-0>.

Salesforce. 2019. "State of the Connected Customer Report Outlines Changing Standards for Customer Engagement." <https://www.salesforce.com/company/news-press/stories/2019/06/061219-g/>

Searls, Doc. 2012. *The Intention Economy: When Customers Take Charge*. Harvard Business Press.

Sirkkunen, E., and P. Haara. 2017. "Yksityisyys Ja Notkea Valvonta: Yksityisyys Ja Anonymiteetti Verkkoviestinnässä-Projektin Loppuraportti." [tampub.uta.fi](http://tampub.uta.fi/handle/10024/100510). <http://tampub.uta.fi/handle/10024/100510>.

TEM. 2017. "Suomen Tekoälyaika - Suomi Tekoälyn Soveltamisen Kärkimaak-si: Tavoite Ja Toimenpidesuosituksset." 41. Työ- ja elinkeinoministeriö. http://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/80849/TEMrap_41_2017_Suomen_tekoalyaika.pdf.

Valtioneuvosto. 2015. "Ratkaisujen Suomi - Pääministeri Juha Sipilän Hallituksen Strateginen Ohjelma 29.5.2015." http://valtioneuvosto.fi/documents/10184/1427398/Ratkaisujen+Suomi_FL_YHDISTETTY_netti.pdf.

Viitanen, Jukka, Reijo Paaanen, Valto Loikkanen, and Aki Koivistoinen. 2017. "Digitaalisen Alustatalouden Tiekartasto." Tekes. https://www.tekes.fi/global-assets/global/ohjelmat-ja-palvelut/kampanjat/alustatalous/alustatalouden_tiekartasto_web_x.pdf.

World Economic Forum. 2013. "Unlocking the Value of Personal Data." <http://www.weforum.org/reports/unlocking-value-personal-data-collection-usage>.

Ympäristöministeriö. 2014. "Sähköisen Asunto-Osakerekisterin Toimintamalli." <http://www.ym.fi/download/noname/%7B76F44244-DE60-4B65-95F2-634D6A857096%7D/99026>.



This report is published under a Creative Commons Attribution 4.0 License (CC BY 4.0). <https://creativecommons.org/licenses/by/4.0>.

In connection with re-use, the authors Antti Poikola, Kai Kuikkaniemi, Ossi Kuittinen, Harri Honko, and Aleksi Knuutila (Open Knowledge Finland), the third edition editor Viivi Lähteenoja (MyData Global), and the publisher Ministry of Transport and Communications, Finland, must be mentioned.

The original Finnish report and its English summary (Poikola et al., 2014; Poikola, Kuikkaniemi, & Honko, 2015) were implemented by the Open Knowledge Finland Association's MyData working group and funded by the Ministry of Transport and Communications, Finland.

The second, updated Finnish edition of this report (2018) was implemented by the Open Knowledge Finland Association's MyData working group and funded by TEKES' strategic opening of the Digital Health Revolution, in the framework of which researchers from Aalto University, Tampere University of Technology and the University of Oulu have participated in the work.

This third, updated and translated English edition of this report (2020) was implemented by MyData Global association and funded by the Ministry of Transport and Communications, Finland.

Graphic design: Kirmo Kivelä

ISBN 978-952-243-617-7: MyData – an introduction to human-centric use of personal data (3rd, revised edition)

