MSc thesis in Complex Systems Engineering and Management

Creating value through opening up data by commercial companies

A case study at PostNL Data Solutions

Michael van der Wielen
2019

postnl

TUDelft

This page is intentionally left blank

CREATING VALUE THROUGH OPENING UP DATA
BY COMMERCIAL COMPANIES: A CASE STUDY AT
POSTNL DATA SOLUTIONS

Master thesis submitted to Delft University of Technology in partial
fulfilment of the requirements for the degree of

Master of Science
in **Complex Systems Engineering and Management**

by

Michael van der Wielen
Student number: 4252608

To be defended in public on 30 08 2019

First Supervisor:	Dr. M.E. Warnier	Systems Engineering
Second Supervisor:	Dr. A.M.G. Zuiderwijk-van Eijk	Information and Communication Technology
Chairperson:	Dr. M.E. Warnier	Systems Engineering
External Supervisor:	E. Beekenkamp	PostNL Data Solutions

TUDelft

PREFACE

In front of you lies the final product of six months of hard work. This master thesis thereby forms the last, individual project of my master degree in Complex Systems Engineering and Management at the TU Delft. This is also the conclusion of a number of very nice years being at the faculty of Technology, Policy and Management. I have had the privilege of having many smart but especially nice and sweet people around me during my journey in Delft. All in all, my student days have had many highlights and I am proud that I can end a great time with this thesis and start a new period in my life. I would like to take advantage of this by thanking a number of important people who helped me in the realization of this project, but who also assisted me during my time as a student.

First of all I want to thank everyone at the Data Solutions department of PostNL for the opportunity you have given me to write my thesis there. Special thanks to Cathérine Westerhuis and Patrick Vink who have made the first draft of the assignment, so I ended up at Data Solutions. In particular, I would also like to thank Ellen Beekenkamp, who guided me very actively as supervisor from PostNL. Our meetings were always very pleasant and very encouraging for my motivation during the execution of my research.

Next I would like to thank my second supervisor Anneke Zuiderwijk for the positivity of your way of giving feedback. Because I had little IT knowledge at the start of my thesis, I was sometimes unsure whether I was heading in the right direction. Your very substantive and detailed feedback helped me a lot in this.

Martijn, this preface is actually too short to thank you for all the help you have given me. That you wanted to guide me with the confidence you always gave me, made sure that I always continued my research in a positive way. Despite your full schedule you always took the time for me. I could not have imagined a better first supervisor.

Last but definitely not least, I want to thank my family and friends for all the support and enjoyment during my studies. Mom, dad and Matthijs, I could never have achieved this without you. Then of course dear Elsemiek, thanks for always listening to the ups and downs during my thesis. You are amazing.

This is a good end to a great time I had in Delft, on to the next challenge.

Please enjoy reading this report.

Michael van der Wielen
August, 2019

EXECUTIVE SUMMARY

In recent years the growth in the amount of available data has increased enormously. More and more companies, government organizations and citizens are becoming data dependent. Simultaneously with this growth in the availability of the amount of data, there is increasing interest in opening up this data. Open data is data that is freely available to everyone, can be reused by everyone and can be redistributed by everyone again. Up to now, it is mainly government organizations that make the data publicly available or use open data sources themselves. However, various studies have indicated that a great potential for open data is still unused. If a larger part of the current available data were to be opened up, much (economic) value could be achieved. However, these studies do not determine for commercial companies. These companies are currently lagging behind in opening up their data because they do not see the potential of this. Opening up data is seen as unnecessary insight for the competition and a loss of the market position and the annual turnover. To show companies the value of open data, this research aims to design a method that companies can apply. By applying this method it becomes clear to companies which data in the current structure of their business system has the potential to be opened up and subsequently what value this opened-up data yields for the company.

This methodology, named "Open Data Implementation Methodology (ODIM)" in this study, was designed using the design science research approach. The various phases in this approach made it clear which factors are important of open data at commercial companies and, among other things, offered a framework for setting up the method. By going through all the phases in the design science approach, the methodology could be designed, tested and evaluated.

In the first phase, the identification of the problem and the motivation, an extended literature review showed, among other things, that most research into open data focuses on government organizations. From these investigated projects the most striking challenges have been identified to be included in the design of the method. These challenges within open data projects are in the technical, social, legal and economic areas. Analyzing the literature also confirmed that conducting a case study is a suitable method to test open data projects. All these elements were then included in the design of the ODIM. Because analyzing the value of open data projects at companies has to do with the design and organization of new systems at companies, it was investigated what a suitable design method was for the ODIM. Due to the iterative nature and previously proven functions, the design engineering approach has been chosen according to Dym and Little. This method was used to subsequently adjust it so that it is suitable for open data implementation projects at companies. The method, the artefact in the design science aproach of this study, follows different phases that companies can

go through. By mapping the current system in detail, determining requirements of open data in the new system and analyzing for example the costs, revenues and use of the data by current users, a concept system design can be drawn up with a central role for open data. This design shows which datasets are suitable for opening up and what changes this entails within the business processes. By testing this new system with experts currently working with the system, the value of the open data in the new system can be determined and changes can be added. The results of the method for a company are ultimately insights into current systems and the value of open data in a newly designed system.

This open data implementation methodology was demonstrated in this study at the commercial company PostNL Data Solutions in the Netherlands. The new open data system and the ODIM have been validated through an expert panel workshop. The results were that open data was seen as a very valuable change compared to current business processes. The value of open data was primarily recognized in the economic, cultural and organizational area. According to the experts who participated in the workshop, opening the data ensures that innovation is stimulated and new and better products and services can be built. The transparency of the company and the accuracy of the data will also increase in this new system. Risks are that it can cost a lot of money and time to set up current IT systems for open data and that privacy can pose risks to users. To be able to apply this method to other companies as well, the benchmarks for value creation must be made more specific to companies. Value creation is very context dependent, which makes it difficult to apply the ODIM broadly. By means of the ODIM, it is possible to make clear how current data systems are structured at companies and determine which data sets are suitable for being opened up.

The conclusion of this research is that open data can create value at commercial companies by reorganizing current systems. The open data implementation method helps companies to gain insight into this and to determine the value of open data for their company. Further research is needed into the broad applicability of the method by examining multiple cases and into benchmarks for value creation. In addition, an implementation plan must also be developed in order to successfully build the developed system with open data.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

API	Application Programming Interface
AR	Action Science Research
BAG	Basisregistratie Adressen en Gebouwen
BPMN	Business Process Modelling Notation
BRPP	Basis Registratie Personen en Percelen
CBP	College Bescherming Persoonsgegevens (Data Protection Agency)
CBS	Centraal Bureau voor de Statistiek
DS	(PostNL) Data Solutions
DSRM	Design Science Research Methodology
EP	Expert Panel (Workshop)
EU	European Union
GDPR	General Data Protection Regulation
IS	Information Systems
KPI	Key Performance Indicator
NAB	Nationaal Adressen Bestand
ODIM	Open Data Implementation Method
OGD	Open Government Data
PCT	de Postcodetabel
PTT	Staatsbedrijf der Posterijen, Telegrafie en Telefonie

1 | INTRODUCTION

"Open data is data that can be freely used, re-used and redistributed by anyone - subject only, at most, to the requirement to attribute and sharealike."
- opendatahandbook.org

1.1 PROBLEM EXPLORATION

For many years it has been known that free and open markets and societies benefit from free access to information (Ho, Liu & Ramanan, 1997; Kareken, Muench & Wallace, 1973). Transparency of market prices and of governmental operations contribute to the efficient organization and integrity of the market. In addition, open information motivates and encourages market parties and citizens to innovate and thereby generate new ideas (Awazu et al., 2009). Today, with the enormous growth in computer power and data analysis methods in recent decades, it is possible to make this open information easily digital and shareable in the form of open data.

Open data gives great impetus to data-driven innovations and creates a new dimension within big data analytics. Businesses are given more options to offer services and products to consumers by combining their own data and open data (Beno, Figl, Umbrich & Polleres, 2017). In addition, companies also make applications based on open (government) data only. In the Netherlands, the mobile application "Buienradar" is a good example of this. By using the open data from the government organization KNMI, another company creates an application that displays the weather forecast to users. Another example where companies experience benefits by opening up data is at the company Geolytix in the United Kingdom. This company has chosen to open up data about the city of London (including locations of supermarkets, hospitals and cinemas but also the routes of all public transport) instead of selling the data as closed data. This had a positive effect on the number of users of the data and so the company was able to earn money from these new users with other products and data. The result was that the company achieved an estimated market value of 3.3 million euros in three years. These are examples of projects where it has been demonstrated that companies have benefits when data is opened up. Research from Manyika et al. (2013), among others, has shown that open data also has enormous potential (economic) value in various business sectors. Figure 1.1 shows how great this potential, measured in billion euros, can be when more open data

is available. The estimate is based on how projects can be better organized and coordinated and how new projects can be started more easily when more open data is available.

Potential value in open data, $ billion

Education	890–1,180
Transportation	720–920
Consumer products	520–1,470
Electricity	340–580
Oil and gas	240–510
Health care[1]	300–450
Consumer finance	210–280
Total	**3,220–5,390**

The values here are drawn from examples of open-data potential and are not a comprehensive sizing of potential value across the two sectors.
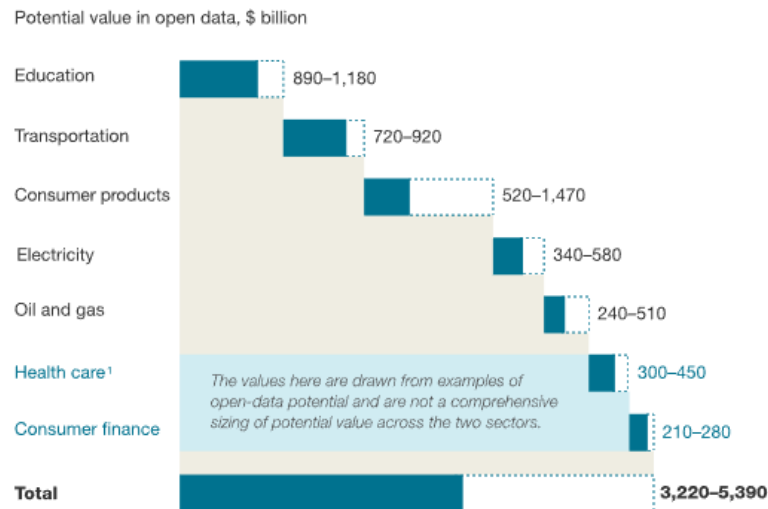
Figure 1.1: Potential annual value of open data in billion dollar. Source: Manyika et al. (2013) Page 9

The European Commission has also launched many studies into the potential value of open data. These studies were combined and further elaborated by Capgemini Consulting in 2017. This research showed that there is enormous potential not only in the economic field but also in many other areas when more data becomes publicly available. An overview of this potential as a result of research by the European Commission, reported by Capgemeni can be seen in figure 1.2

Market Size and value added as percentage of GDP	Number of jobs created	Cost Savings for the public sector	Efficiency Gains or productivity gains
• € 325 Billion direct market size for the period 2016-2020 • 36.9% increase in share of GDP from 2016 to 2020 • € 83,578 Million market for Public Administration in 2020	• 100.000 jobs in Open Data in 2020 • 7.3% average increase Open Data jobs • More than 2.500 jobs in at least nine countries	• € 1.7 Billion saved in costs EU28+ for Public Administrations in 2020 • Highest Cost Savings by Germany with € 296 Million • 0.22% of estimated Cost Savings for Denmark's Government Expenditure in 2020	• 7,000 lives can be saved due to quicker response • 2,549 hours can be saved that are currently wasted on finding parking • 629 million hours can be saved which is equivalent to € 27.9 Billion

Figure 1.2: Identified benefits and estimated numbers when available data is published as open data. Source: Berends, Carrara and Radu (2017) Page 4

And it is precisely in the availability of open data where there is a major problem in the current open data ecosystem (Janssen, Charalabidis & Zuiderwijk,

2012). The examples of open data projects mentioned in the previous section show that open data can offer benefits in areas such as economic and social areas for companies. In these examples, and in many other current open data projects, mainly open government data is used by companies or is it the government that makes the data publicly available (Janssen et al., 2012) and not the commercial companies themselves. However, the opening up of this data is still on a small scale and the business community must also make a contribution in order to be able to use the full potential of open data (this potential can be seen in figure 1.1). The maximum potential and therefore the maximum value of open data is currently not being realized in the business world (Kitchin, 2014; Ahmadi Zeleti, Ojo & Curry, 2016). Within companies, there is often the idea that opening up proprietary data can lead to unnecessary insights for the competition and a loss of market position (Hammell et al., 2012). For companies there is currently much uncertainty about the usefulness of open data and the value of open data for the company. This research further examines the possibilities of opening up data by commercial companies. To make this attractive, there must be a certain value to this data for the companies. This value for the companies is also investigated in detail in this study. In order to be able to start the research into opening up data within companies, it is useful to first sketch more context regarding the developments concerning data in recent years. The most important developments for the continuation of this research are: the increase in data availability and the risks of open data.

1.2 THE WORLD AROUND DATA

On the one hand, the growth in the interest in open data has to do with the fact that it has been demonstrated that there is much (economic) potential with regard to open data (figure 1.1). On the other hand, this interest also has to do with the fact that simply the data landscape in which we find ourselves has undergone enormous developments. The availability of data has increased significantly in recent years, which arouses the interest of many researchers (Borgman, 2017). The changes in the data landscape and the exponential growth in the amount of data are explained in more detail in the following two sub sections.

1.2.1 Increase in data availability

In recent years, more and more platforms have emerged where people leave their data (Netflix, Google, Amazon, etc.). In addition to personal data that users leave on those platforms, the online behaviour of people is also stored. Consider, for example, the spending pattern of a specific neighborhood on an e-commerce website. By storing this data centrally, companies can make targeted advertisements and the government has more insight into their citizens. Also because more and more devices (such as telephones, laptops and

tablets) are connected to the internet, data is collected and stored more easily. All these different forms of data collection form a so-called "data pool" within all kind of links and analyzes are possible.

As a result, data collection has risen explosively in recent decades, and particular in recent years. An example of this is that in 1988 the average Dutch person was registered in 30 to 40 government databases. In 2009, the former CPB (College Bescherming Persoonsgegevens) showed that this has grown to potentially hundreds or even thousands (Schermer & Wagemans, 2009). The government, companies and countless other organizations have more and more the availability to data. This growth entails not only the amount of data that is collected but also the ease with which collecting en storing is happening nowadays. (van Ooijen, 2014).

These days large-scale data storage is also referred to as Big Data. Figure 1.3 shows the position of open data within the entire data landscape. This research focuses on the open private data part and in particular how more companies can open up their data.
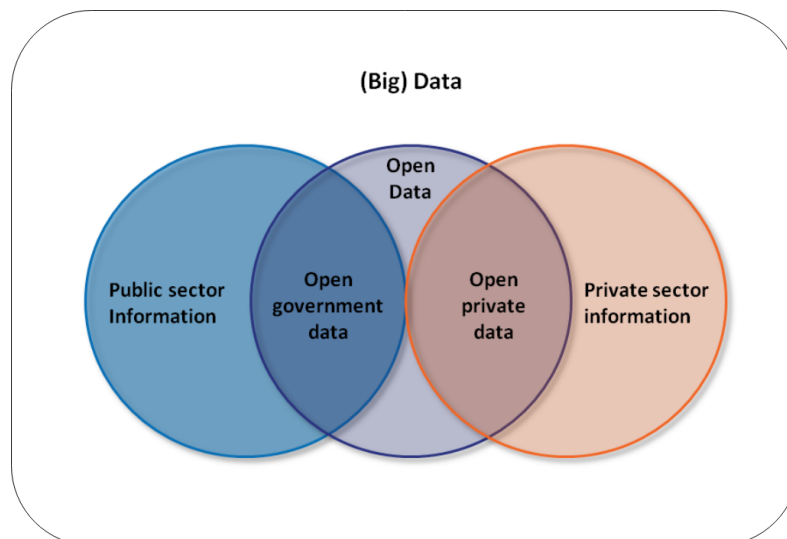
Figure 1.3: The data landscape

1.2.2 The development of open data

In addition to the fact that more and more data is available, there is also an increasing trend and demand for opening up the data (Frankowski, Van Der Steen & Meijer, 2015). The developments of this are somewhat the same but are not identical. There has been a discussion about ownership and access to data for many years, but that was independent of the large-scale collection of data only (Frankowski et al., 2015). Over the past few decades, initially more and more (government) organizations are starting projects to open up databases. The reasons that organizations do or would like to do this is mentioned in Susha, Zuiderwijk, Janssen and Grönlund (2015). This study summarizes the benefits of open data, also obtained from other studies:

1. Open data is appreciated for increasing transparency within organizations

2. Open data contributes to an increase in the participation of citizens

3. In democratic countries, more open data also contributes to higher democratic accountability

4. For citizens, industry, different research fields and other fields enhances and improves open data governmental and nongovernmental value-added services

5. Open data is also a powerful resource for business innovation because open data can drive the development of new business models grounded on collaboration between private and public stakeholders

1.3 THE OPEN DATA ECOSYSTEM

As has already been highlighted in the previous sections, there are two types of stakeholder groups in the open data ecosystem (in some literature known as the open data environment): data users and data providers. Within both users and providers it is easy to distinguish between three different groups: the government, businesses and citizens. All three groups produce data and deliver this to each other in a certain way and also make use of (each other's) data. The example of the Buienradar application (mentioned in section 1.1) is a good illustration of this: the KNMI (Dutch governmental weather institute) publishes open weather data which RTL Nederland uses to make the application and citizens can use this for free. All relationships between these three most important groups of stakeholders in the open data ecosystem are shown in figure 1.4.

In this study, the focus is on the data supplied and used by businesses, in figure 1.4 shown with a red dotted area. Despite the fact that this figure was compiled in 2012, nothing has changed as far as the current important stakeholders are concerned. The research focuses on companies that open up data themselves. These datasets can then be used by the government, citizens or other companies. By having the focus here, further exploration of the possibilities of the business potential of open data described at the beginning of this chapter will be carried out. In addition, the fifth advantage of open data as described in section 1.2.2 about the development of new business models through a collaboration of public and private stakeholders is further investigated on.

1.4 KEY CONCEPTS

This research will, as described earlier, deal with the possibilities of opening up datasets by commercial companies. To be able to fully utilize the potential
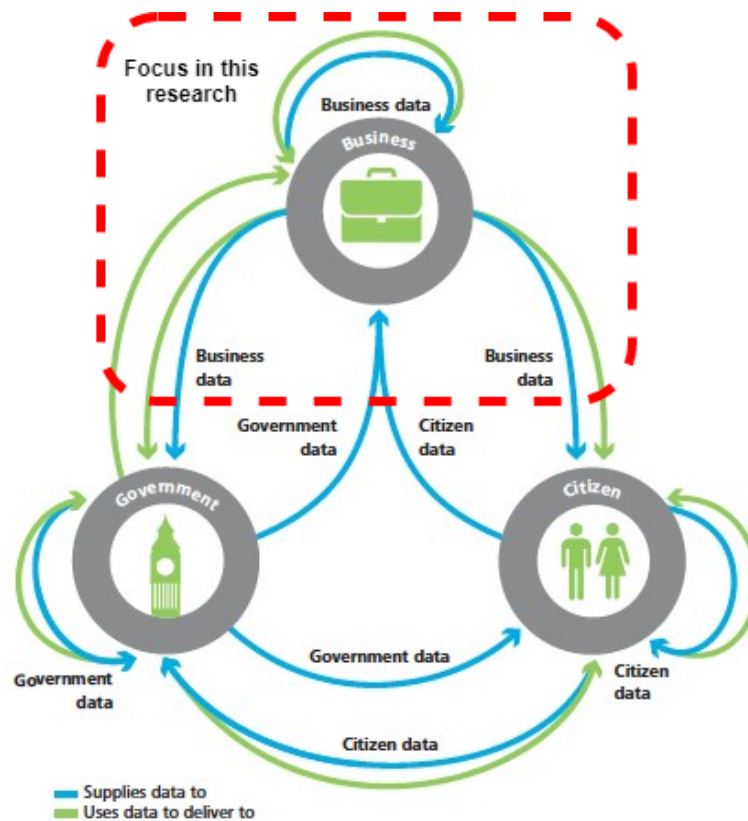
Figure 1.4: The open data ecosystem. The focus of this research is shown with a dotted area in red. Adapted and adjusted from: Hammell et al. (2012)

of open data, more attention needs to be paid by the business community to this. In the following two sub sections, these two important concepts that are used in this research are defined.

1.4.1 Definition of open data

It is important to use a general definition of the concept of open data in the remaining of this research. Open data is a concept which can be viewed in many different ways. Also in the literature, many different definitions are mentioned. In the book "The World of Open Data" written by Charalabidis et al. (2018) the most important characteristics of open data are described. The paper of Malamud et al. (2013) also mentions a complete definition with characteristics of open data. This study uses a combination of both studies (which each consist of several studies). This definition is as follows:

> **Definition open data:** "Data can be categorized as open data when the data is: complete, primary, timely, accessible, machine process-able, made online in persistent archives and as accurate as possible. Besides this the data access must be non-discriminatory, the data formats must be non proprietary data and the license must be unrestricted and bear no usage costs."

1.4.2 Commercial businesses

The benefits and value of open data are investigated in this study from the perspective of commercial businesses. These types of companies mainly have a profit motive and act on behave of the customer (user). In the literature that has been studied, much is already written about the use and opening up of data by government organizations (Zuiderwijk, Shinde & Janssen, 2018; Janssen et al., 2012; Ali Hassan & Twinomurinzi, 2018). That is why it is interesting to further explore the subject of open data from a business perspective. Because companies are still cautious about open data at present, the benefits of this must be well communicated. In order to lower the threshold and facilitate the step towards open data, open data for companies must meet most of the characteristics mentioned in the previous sub section. The ultimate goal is that the data meets all these characteristics.

1.5 THE OTHER SIDE OF OPEN DATA: THE POTENTIAL RISKS OF OPENING UP DATA

Opening up certain datasets are not only associated with advantages. One of the risks associated with open data is that it ultimately yields nothing. Actually opening up the data requires a lot of knowledge and skills. Organizations will do this from the point of view that they will ultimately gain certain benefits from this. But the effect can also be that the opened data does not produce a positive effect at all. Resources might be wasted on publishing the data that is not used or relevant. The risks of open data are described in Zuiderwijk and Janssen (2014a) and Zuiderwijk and Janssen (2014b). These risks from that literature are shown in figure 1.5.

1.5.1 The risks of open data

One of the risks mentioned in figure 1.5 has to do with privacy. The law in Europe that deals with this is called the General Data Protection Regulation (GDPR). Since this has been a relevant topic in recent years when data is being discussed and written about, this risk will be explained in more detail in the next sub section.

1.5.2 GDPR

Privacy is a concept that is of paramount importance to people and which is also legally established in certain ways. Because the growth in data collection has risen so rapidly in recent years, the regulations regarding data privacy lagged behind. This changed in May 2016 when the European Union introduced a regulation, a legislative instrument of general application, which standardized the rules for the processing of personal data by private com-

Category	Risk	Description
Legislation	Non-compliant	All kinds of legislation might be applicable from different domains. There might be unawareness of which legislation might be applicable.
	Privacy	The data protection act poses strict requirement on what can be published and what cannot be published. Although there are PEM, the privacy of persons can be violated unintentionally or data might be used for another purpose than what it was collected for.
Governance	Responsibilities	Difficulties with data ownership and stewardship. Unclear responsibility and accountability.
	Maintaining quality	Unclear if data is updated and what the data quality is.
Interpretation	Bias	Published data can be biased
	Ambiguity misinterpretation	Misinterpretation and misuse
Data quality	Poor data quality	Data might have different qualities (completeness, accuracy, timeliness). Decisions can be made on poor information quality or the wrong insights can be created.
	Timeliness	The most recent data might not be available. Embargo period prohibits the publication of recent data

Figure 1.5: Overview of risks of open data. Adapted from: Zuiderwijk and Janssen (2014a) and Zuiderwijk and Janssen (2014b)

panies and public authorities throughout the European Union (European Union, 2016). From that moment, companies were given two years, until May 25, 2018, to adjust their business operations to the new regulation called the GDPR (General Data Protection Regulation). One of the objectives of this regulation is to guarantee the protection of personal data within the European Union and also to guarantee the free movement of data within the European internal market (European Union, 2016). The consequences were, and still are, huge for companies dealing with the processing of personal data. A number of rules that must be followed from the introduction are, for example:

- Transparency to the person whose data is being processed. He or she must be aware of this

- The personal data may not be used for purposes other than an intended, legitimate, purpose

- Only the data that is necessary for the intended purpose may be collected

- The personal data may not be kept longer than necessary for the intended purpose

- The personal data must be protected against unauthorized access, destruction or loss

- The responsible organization must be able to demonstrate compliance with these rules

Companies that do not meet these requirements run the risk of being fined a maximum of 20 million euros or four percent of the annual worldwide

turnover. Research by Thomson Reuters (2019) has shown that more than half of the companies surveyed (1350 in total) almost a year after the official introduction still do not meet all GDPR requirements.

1.6 PROBLEM STATEMENT AND RESEARCH OBJECTIVE

The sections above describe a large part of the current playing field around open data. Open data is an topic which has increased interest in the current era because more and more data is available (which can be valuable to open up) and it has been shown that open data has (economic) potential. This potential is not being fully exploited at this time because mainly government agencies are using open data and opening up certain datasets. Companies are still lagging behind in this regard. This is also the central problem which is formulated as follows in this research:

> **Problem statement:** Currently, too little research has been done into how commercial companies can open up datasets and what the effects are for the company of these opened up datasets.

This problem is defined from the current literature on open data. An extensive literature review, to further define the knowledge gap, can be read in chapter 3. To address this issue, the following objective has been defined for this research:

> **Research objective:** To design an open data implementation method that can help commercial companies to find datasets that can be opened up and to identify which value open data can create for the company.

1.7 STRUCTURE OF THIS STUDY

In chapter 2 the main questions of this research are formulated and the methods are explained to answer these questions. In the next chapter 3 follow an extended literature review of important elements of open data for commercial companies. With the insights from this chapter, a design is presented in chapter 4 how the open data implementation methodology can be designed. This method is the artefact of this research and is described in detail in the same chapter 4. In chapter 5 the method is demonstrated on a case at PostNL to investigate whether the implementation of the method identifies potential of opening up data. In chapter 6 follows an evaluation of the method and explains the possibilities for a broad applicability of the method. Chapter 7 reflects on the choices made in this study, after which the answers to the research questions are given in Chapter 8.

1.8   SUMMARY CHAPTER 1

Research has shown that there is a lot of potential for open data. Because more data has become available in recent years and many advantages of opening up data have been demonstrated, the interest in open data has grown enormously. Nowadays, more and more government organizations make data available. In the open data ecosystem there are also two other important stakeholders, namely citizens and businesses. The businesses in particular are lagging behind with regard to opening up of the data. This is partly due to the fact that companies are not yet fully aware of the benefits of open data. Despite the fact that there are also various risks associated with open data, including with regard to privacy, it can be beneficial for a company to open up data. To give companies more insight into the value of opening up (parts of) their data, this research aims to develop a methodology that can help commercial companies identify which datasets can be opened up and which value this opened up data can create.

2 | RESEARCH APPROACH

The introduction has shown that open data has a lot of potential, but that it is not yet being fully utilized at the moment. To be able to use this potential, companies need to do more with open data. Because up to now government organizations are mainly concerned with opening up data, more research needs to be done into open data for (commercial) companies. The objective of this research, as defined in chapter 1, is to develop a method that can help companies with open data implementation projects. This method that will be developed aims to make it clear to companies which data can be opened up and what value this creates. In the remainder of this study, the method will be referred to as the open data implementation method. This study therefore aims to develop an artefact, namely the open data implementation method (ODIM) and therefore using the design science approach is suitable for this research (Hevner & Chatterjee, 2010).

First of all, the main research question of this research will be formulated in the next section. In section 2.2 follows an explanation of the design science approach and why it is suitable to apply within this research and to use it to design the open data implementation method. This is followed by an overview of all research questions that fit within the phases of design science and the methods used to answer the research questions.

2.1 MAIN RESEARCH QUESTION

This research focuses on answering a central main research question. This follows from the previously identified problem and is motivated by the goals, changes and developments previously set out in chapter 1. This revealed that the impact of open data should be investigated at companies. A method is being developed for this. To evaluate this method, it will be tested through a case study. This case concerns a company operating within the logistics sector, a relevant and useful sector for demonstrating the potential of open data (chapter 1 figure 1.1). Because the method will only be tested at a company within this logistics sector, the main research question will also be formulated to this. For a further explanation of the case and the relevance of the logistics sector, see chapter 5. The main research question of this research is:

"How can open data contribute in value creation of digital products or services for commercial companies in the logistics sector?"

This main research question is based on the hypothesis that companies can create certain value by opening up parts of their data and making them freely accessible. By making products that are unique to the consumer, the value can be increased in a changing digital landscape within the logistics sector. What exactly this value entails, and whether this is indeed true, is further investigated in this study.

2.2 DESIGN SCIENCE RESEARCH

This paper will further follow the Design Science Research Methodology (DSRM), proposed by Hevner and Chatterjee (2010). This method is chosen because it is usually employed in information systems (IS) which makes it applicable for this research as this research is related to Information Systems. DSRM combines the relevance of the project, the contextual environment, with the scientific knowledge base. Design science is concerned with "producing and applying knowledge of tasks or situations in order to create effective artefacts" (March & Smith, 1995). Offermann, Blom, Schönherr and Bub (2010) has investigated that different types or artefacts are distinguished in the literature. One of the artefact types is a methodology that demonstrates that design science research is a good approach to developing the ODIM in this research. Offermann et al. (2010) also describes that if a method is the artefact from the study, it should describe a series of activities to create a system or to interact with a system. This is precisely the purpose of this study (describing a series of activities to be able to implement open data within companies) and it has once again been demonstrated that design science research is the right approach to apply.

Various guidelines and methods have been proposed in earlier studies to apply design science research (Hevner & Chatterjee, 2010; March & Smith, 1995; Peffers, Tuunanen, Rothenberger & Chatterjee, 2007). Each guideline starts with identifying the problem and defines the purpose of the solution. Peffers et al. (2007) has made an overview of many previous studies where DSR has been applied in which all phases are highlighted. All these phases can be seen in figure 2.1. The DSRM is a nominal process sequence with some iteration between specific design process steps. DSRM consists of six different design process steps: 1) identify the problem, 2) define objectives of a solution, 3) design and development, 4) demonstration, 5) evaluation and 6) communication. The problem definition motivates the relevance of the research and the value of the solution which is partly done in chapter 1 and will be further elaborated on in chapter 3. The next step helps to understand what is possible and what are feasible solutions. In the third step the artefact will be designed which meets the proposed objectives. To demonstrate the artefact, the method will be tested by using it in a demonstration. In this research, this step looks at whether the ODIM works in a specific situation at a commercial company. Following this step, the artefact is evaluated and communicated in the final step.
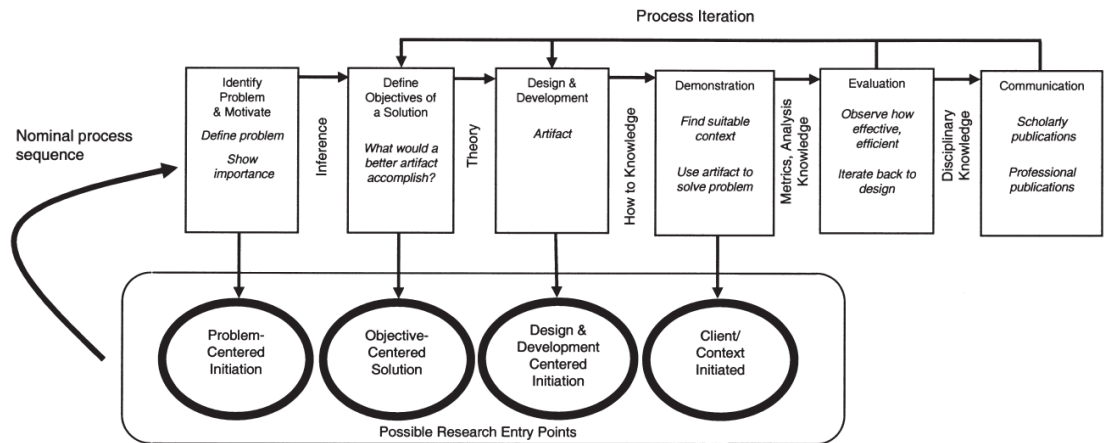
Figure 2.1: DSRM process structure. Source: (Peffers, Tuunanen, Rothenberger & Chatterjee, 2007) Page 54

2.2.1 Rigor and relevance in the design science research approach

Three closely related design cycles are connected within design science research. The relevance cycle, the first of the three, ensures that the relevance of the research also goes beyond the conditions set for the research. The application context is not only related to the requirements within the research, but above all this cycle ultimately defines the evaluation of the results. This cycle of the design science approach can be found in chapter 1, 6 and 7. The rigor cycle provides information from the past to support innovation in research. To prevent routine research and to ensure that the research adds something to the scientific world and is more broadly applicable, it is the task to make clear which research it builds on. This rigor cycle can be found in this study in chapter 3 and 8. The third and final cycle of the design science approach is the design cycle and discusses the core activities of designing and evaluating the artefact. These cycles are treated in this study in chapter 4 and chapter 5. These three important cycles within the design science research approach are reflected in much of the researched literature. Hevner and Chatterjee (2010) and Armstrong and Sage (2000) have laid the foundation for this research to build on.

2.3 RESEARCH DESIGN

The previous section describes the design science approach that is applied in this study to develop the artefact, the open data implementation method. To give a clear structure to the further course of this report, figure 2.2 has been compiled. Herein the phases from the design science approach are shown on the left in the green blocks. A research question has been formulated for each phase within this approach, which can be seen in the figure in the middle column (orange blocks). Each research question is tackled using a

different method which can be seen in the right-hand column, the purple blocks. The figure refers to research instruments which can be defined as "the specific methods that are used to execute a particular research strategy" (Gonzalez, 2010).
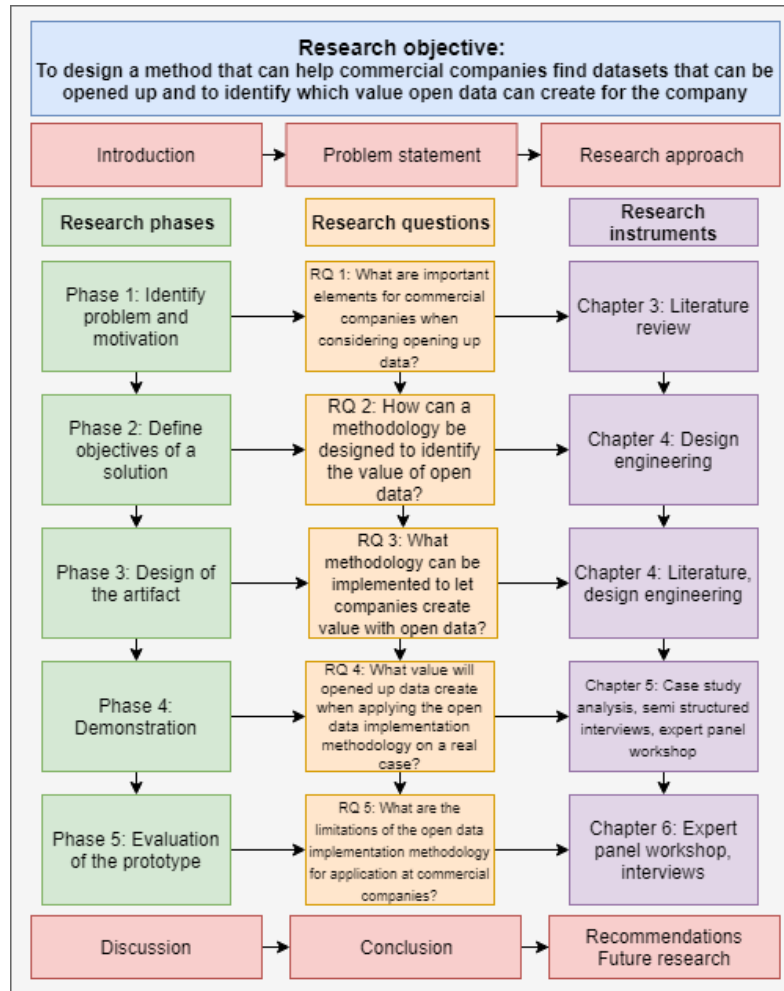
Figure 2.2: Research flow diagram

2.4 SUB QUESTIONS AND RESEARCH METHODOLOGIES

To be able to formulate an answer to the main research question, it is subdivided into five different sub questions in combination with the five different phases of the design science approach. The final phase of the DSRM, communication, will not be included in a sub question but is part of the conclusion and discussion at the end of this research.

2.4.1 Sub question 1

Sub question 1: What are important elements for commercial companies when considering opening up data?

The introduction has shown that open data has become an increasingly discussed topic in recent years. Much of the current research into open data mainly focused on open government data and governments that use open data. In order to provide insight into how companies can open up (parts of) their data , research must first be conducted into important elements for commercial companies when considering doing this. To answer this sub question, it is also worth investigating the conclusions from previous studies of government organizations that use open data. By also applying these elements within a business environment, this can also contribute to answering the main research question. By performing this analysis, it is clearly defined what the problem is and the motivation to tackle it. This fits exactly with the first phase within the DSRM.

METHODOLOGY: LITERATURE REVIEW   To get to know these elements, an extensive literature review will be conducted. An advantage of this is that a literature research uses the knowledge of others and therefore illuminates the subject from multiple perspectives. A disadvantage is that literature research has no direct contact with the research object (Wee & Banister, 2016). Because the first sub question is a somewhat more generic question, it can be answered well through literature research.

2.4.2 Sub question 2

Sub question 2: How can a methodology be designed to identify the value of open data?

The aim of this research is to develop a methodology that companies can use to implement an open data strategy within the company (the ODIM). Such a methodology is based on the assumption that the current business systems are designed differently to add value to the new system using open data. This method should be able to be used within multiple business environments and be both a practical addition and an addition to existing literature about open data. By focusing on companies instead of governments, this researches an as yet underexposed side of open data.

METHODOLOGY: DESIGN ENGINEERING   A method whereby the problem is identified, the current situation is mapped out and a proposal is made as to what a new system can look like has already been carried out in many previous studies. This so-called process design engineering can therefore serve as the basis for the steps that are designed in the open data implementation method. This research will use the design engineering method developed by Dym, Little, Orwin and Spjut (2008). In this sub question the steps that occur in this method are explained and adapted to a suitable method for analyzing systems and redesigned for open data implementation systems. A further explanation of the method of Dym et al. (2008) can be found in chapter 4.

2.4.3 Sub question 3

Sub question 3: What methodology can be implemented to let companies create value with open data?

The result from sub question two is a framework of a previously designed process design engineering method that can be adapted for open data projects. The global outline of the steps that must be followed are clearly highlighted. In order to make the method explicit and ultimately to be able to apply it in practice, the steps that occur in the method must be further elaborated. This sub question provides insight into how it is actually possible to identify which datasets can be opened up and how it can be investigated what value this creates for companies. The result of the sub question is the artefact to be designed according to the design science research approach (Hevner & Chatterjee, 2010).

METHODOLOGY: LITERATURE, DESIGN ENGINEERING   To answer this sub question the results are used from the previous sub question. That is why design engineering will once again be used here, where in this research the framework of Dym et al. (2008) will be used. In answering this sub question, the literature is also consulted in order to be able to apply existing techniques in the methodology. Examples of this are conducting interviews, knowledge acquisition management methods and conducting surveys.

2.4.4 Sub question 4

Sub question 4: What value will opened up data create when applying the open data implementation methodology on a real case?

Now that the previous three sub questions have revealed which factors are important within open data implementation projects at companies and what a methodology can look like to identify the value of this open data, it can be applied to an existing case. Where in the previous three sub questions it has remained mainly theoretical, this sub question will be more practically. By applying this, valuable results will follow for both the ODIM and the value of open data within companies. Results from this sub question are therefore twofold. First of all there will be a feedback cycle that reflects on the developed method and secondly this sub question provides insight into the area in which open data can create the most value within companies (or in this case specifically the company investigated in this case).

METHODOLOGY: CASE STUDY ANALYSIS, INTERVIEWS   A case study can be defined as "an empirical inquiry that investigates a contemporary phenomenon within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident" (Yin, 2003). A number of advantages of using the case study method is that this method is suitable for:

1. In-depth examination in complex real-life environments (Dubé & Paré, 2003)

2. Explorative research in order to investigate contextual factors over which the researcher has no or little control (Yin, 2003)

3. The research of Information Systems because a case study offers a holistic perspective on the interaction between the organization, people and technology (Dubé & Paré, 2003)

4. Analyzing circumstances where actors play a crucial role and the context is important (Benbasat, Goldstein & Mead, 1987)

Because the value of open data can differ for many companies, testing the methodology on a real-life case is important. In addition, many different stakeholders also play an important role in this because they too can assign different value to opening up data within a company. In the end, the stakeholders involved in an open data project are also the ones who acknowledge the value of open data. Applying the designed methodology to a real-life case is therefore a good way to investigate this.

Yin (2003) also argues that case studies can be criticized for a lack of rigor, their long duration and a lack of basis for scientific generalization. Yin (2003) argues that this criticism is misdirected because it is mainly because incorrect examples of the application of case studies in available literature and for not following specific methodological procedures. Because these procedures are developed in this study and then applied directly within the case, this will be less the case. Finally, Yin (2003) also argues that it is better to use multiple case studies to be able to draw a correct conclusion about the methodology applied to the case. Since the value of open data is specific for each case studied and given the time frame of this study, it has been decided in this study to limit it to a single case study. This means that this one case can also be discussed in great detail.

An interview technique will also be used within the case. Interviewing experts might help researchers find relevant and new information on topics that are relatively new or have a low amount of scientific literature. That is the reason that interviews can be seen as an empirical research method to extract relevant information (Bogner, Littig & Menz, 2009). Within this research field, experts will be interviewed to make a detailed design of the current system of the selected case study. The advantage of conducting interviews is that as a researcher you can get a good picture of the stakeholders involved and their underlying thoughts. However, a disadvantage is that interviews are time consuming and tend to be relatively small numbers which raise issues of representativeness and specificity (Opdenakker, 2006).

2.4.5   Sub question 5

Sub question 5: What are the limitations of the open data implementation methodology for application at commercial companies?

After the artefact has been demonstrated in the form of a case study, it will be evaluated. The result of this sub question is a possible adjustment, or general confirmation, to the methodology that has been tested. This sub question provides insight into how a subsequent researcher or company can best apply the method within another case.

METHODOLOGY: EXPERT PANEL WORKSHOP, LITERATURE   To be able to make substantiated adjustments to the designed artefact, it must be tested by real-life stakeholders. For the most part this has already been done in the previous sub question in the form of a case study. Part of the methodology is involving the stakeholders (chapter 4 for further elaboration). This is carried out through an expert panel workshop and interviews. On the one hand, this produces results that provide insight into the value of open data for the company from the perspective of the stakeholders involved, and on the other hand, these results are used as an evaluation of the method. The expert panel workshop is therefore conducted as part of the case study, but the results are also used to answer this sub question.

Expert panel workshops can generate discussions on predefined questions or case material concerning the value of open data and the use-fullness of the ODIM. This group-thinking can, therefore, generate different results than a one-on-one interview would. The other main reason for choosing an expert panel workshop as a research method is that it naturally provides the possibility to recreate real-world situations. Chapter 5 and appendix D contain a detailed explanation of how this expert panel workshop was set up and what the results are.

2.5   LINK COSEM PROGRAM AND THIS RESEARCH

To conclude this chapter, the question on how this research is linked to the master degree Complex System Engineering and Management (CoSEM) will be answered. The master degree, CoSEM, teaches its students to design solutions in a highly complex technical environment. The complexities arise when multiple stakeholders interact with a specific technical environment. The solution should enhance the preferences of stakeholders. This research focuses on the design of an open data implementation method. The design process takes place in the technological environment of the changing data landscape (see explanation chapter 1) in which the value of open data is identified. This leads to the execution of design science research whereby the current situation within a chosen case is mapped and a proposal is made for a new system design. This method of investigation, the methods used and theories used (such as Dym et al. (2008) for developing the open data implementation method), fit well within the program of CoSEM.

2.6 SUMMARY CHAPTER 2

To be able to use the full potential of open data, the business community must also become involved with the open data developments. This requires more insight for companies into the benefits and value of open data. To give companies more insight into this, an open data implementation method is being developed in this study, which should make it clear to commercial companies which datasets can be made public and what value this creates for the company. This methodology is the artefact that is being developed according to design science research. The phases identified in this research approach form the guideline for the continuation of this research and a sub question has been drawn up for each phase. These sub questions form the basis for formulating an answer to the main research question. Various methods such as literature review, a case study and an expert panel workshop are used for this. The final result of this research will be a methodology that can be applied within commercial companies for open data implementation en value creation.

3 | LITERATURE REVIEW: IDENTIFICATION OF OPEN DATA ELEMENTS

The problem is identified and motivated in the first phase of the design science approach. In this study, this section has already been briefly discussed in the introduction and will be further explained in this chapter by using an extended literature review. The problem at the moment is the fact that a lot of potential for open data has been demonstrated, but that companies are still not opening up their data mainly due to the fact that the value of opening up data is not yet realised. The literature review addresses this uncertainty by analyzing literature about opening up data by companies. From this review of the literature, important elements for commercial companies when opening up data can be included in the design of the artefact in the next chapter. This artefact will be the open data implementation method (ODIM) that should make it clear to companies which data can be opened up and what value this can create. This chapter will formulate an answer to the first sub-question:

What are important elements for commercial companies when considering opening up data?

These elements will form a good basis for the steps to be designed in the ODIM. In addition to the fact that this chapter provides an answer to the above-mentioned sub-question, this chapter also offers a further in-depth study of the current literature on open data. This chapter shows what the current state of knowledge is about open data research by performing the literature review and thereby identifies knowledge gaps for opening up data by companies.

3.1 APPROACH

A correct approach is needed to come to a systematic literature review. Dyba, Dingsoyr and Hanssen (2007) describes four important steps that are part of this process in his research.

1. Identify key research questions that represent the study

2. Identify the resources from which the literature will be drawn

3. Choose search terms and apply them to the resources defined in two

4. Use inclusion and exclusion conditions on the search results to select papers from the total result

What is immediately noticeable when looking at the literature is that when writing about open data, open government data (OGD) is meant. OGD is known as accumulated data which can be compiled into de-identified data-sets and which are released to the general public as open data (Huijboom & van den Broek, 2011). OGD has attracted much more attention in recent years as already mentioned in section 1.2.2. Because the availability of open government data (OGD) is constantly increasing, the pressure on other public organizations to also open their data has also increased (Janssen et al., 2012). Literature research on OGD therefore also provides many insights into the entire open data ecosystem and therefore into important elements which are important when companies consider to open up their data.

To obtain the best possible picture of the available literature on open data, the most widely used electronic libraries are used. These are: Science Direct, Scopus Database, ACM Digital Library, IEEE Xplore and Web of Science. It soon became clear that most papers that appeared on Web of Science, Science direct, IEEE Xplore and ACM are also available on Scopus. The following terms are used as search terms: *"Open data, Open government data, open data opportunities, open data challenges, open data value creation"*. Combining these search terms using the OR operator resulted in more than thousands results. In addition, use was also made of analyzing the reference lists of the identified papers to find more useful papers. The following selection criteria were applied to narrow the amount of results down to a manageable amount:

1. Only peer review articles written in the English language were considered

2. Studies that were only referring to open data without elaborate more on it were excluded. The study had to go deeper into the use of open data or the results achieved with open data to be of value for this study

3. The study must include a certain relationship between the use of open data in a business environment. A number of literature reviews have already been conducted on open data (Ali Hassan & Twinomurinzi, 2018; Safarov, Meijer & Grimmelikhuijsen, 2017) that explicitly only include open data in governmental context in their analysis. As this study focuses on value creation through opening up data at companies (explained in chapter 1), this literature review only includes papers that are related to the business environment. In concrete terms, this means that the users of the open data in the researched literature are business related or that it is companies that make the data themselves open.

4. Only literature written between 2013 and 2019, a period of approximately five years, has been included in this study. Since the amount of research in this area has increased considerably in recent years and the field of research is relatively new, only the most recently written literature is considered for this study. A period of five years was used as an assumption for this.

Reading the introduction, conclusion and abstract often provided a lot of insight into whether the paper was useful for this study. Given the timeframe

of this project and to provide a clear overview of the available knowledge, a literature review was carried out consisting of 15 papers. In order to be able to draw comparisons in the literature found and ultimately to identify a knowledge gap and to extract factors that are important within open data-related projects at companies, open data is viewed in the identified literature from different perspectives.

3.2 FACTORS AND PERSPECTIVES

Open data and OGD is a concept that have a role in many different contexts. As Zuiderwijk, Helbig, Gil-Garcia and Janssen (2014) already describes in a special issue in a journal, there are seven different perspectives that emerge in the literature when reading and analyzing open data literature. These seven perspectives are legislative, political, social, economical, institutional, operational, and technical perspectives. These perspectives are shown in figure 3.1. Many studies focus on one of the perspectives, but a combination of these leads to new insights. Since open data is viewed from so many different perspectives, it can also be concluded that research into open data is in a complex socio-technical system. In the remaining of this chapter, where the literature is being analyzed, all these perspectives have a certain role.
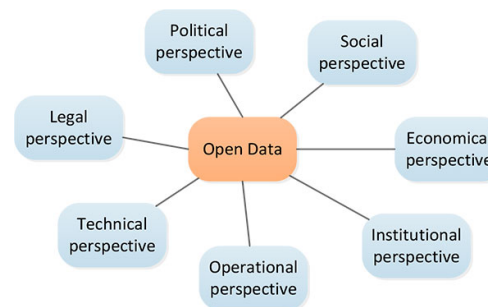
Figure 3.1: Perspectives on open data Adapted from Zuiderwijk, Helbig, Gil-Garcia and Janssen (2014)

Page 2

Now that research into open data can be viewed from many different perspectives and that it has been shown that it is in a complex socio-technical system, it is important to identify the factors in order to be able to compare the papers. These factors, or elements, also provide an answer to the research question proposed at the beginning of this chapter. By researching these factors, it can be identified to what extent the elements should come back in the open data implementation method that is being developed as an artefact in this study. In the literature review of Safarov et al. (2017) and Ali Hassan and Twinomurinzi (2018) a number of factors are discussed that are useful for comparing open data projects. These factors have been combined and based on own findings, factors have been added and removed. By performing this analysis, the following factors were compared in open

data literature: *"The context of the research, The method used in the research, The conditions of open data and The effects of using open data"*.

Table 3.2 shows an overview of the literature studied. To make it clear what is dealt with per factor in the various papers, sub-categories have been created. These sub categories fit into the framework prepared by (Amui, Jabbour, de Sousa Jabbour & Kannan, 2017) where numbers as symbols are used to classify the selected papers. The explanation of the numbers can be seen in table 3.1. A brief summary of each paper can be found in appendix A. In the following sub sections, the most important elements from the literature are explained for each factor.

Factor	Context	Codes
Context	Developed countries	1A
	Developing countries	1B
	Not applicable	1C
Method	Casestudy	2A
	Survey	2B
	Explorative	2C
	Qualitative	2D
	Quantitative	2E
	Design science research	2F
Challenges	Legal	3A
	Technical	3B
	Economical	3C
	Societal	3D
Effects	Participation / colloboration	4A
	Public services developement	4B
	Transparency and accountability	4C
	Economic development	4D
	Adoption	4E
	Innovation	4F

Table 3.1: Classification framework with codes composed by applying the framework designed by Amui, Jabbour, de Sousa Jabbour and Kannan (2017)

Number	Reference	Context	Method	Challenges	Effects
1	(Jaakola, Kekkonen, Lahti & Manninen, 2015)	1A	2A	3A, 3B, 3C	4A, 4B
2	(Susha, Grönlund & Janssen, 2015)	1A	2B, 2C	3C, 3D	4D, 4E
3	(Magalhaes, Roseira & Manley, 2014)	1A	2B, 2D	3B	4B, 4C, 4F
4	(Kalampokis, Tambouris & Tarabanis, 2013)	1A	2A, 2D	3D	4B
5	(Marin de la Iglesia, 2014)	1A	2E	3C, 3D	4B, 4D, 4F
6	(Zuiderwijk, Janssen, Poulis & Van De Kaa, 2015)	1A	2B, 2D	3B	4F
7	(Lakomaa & Kallberg, 2013)	1A	2B	3C, 3D	4A, 4F
8	(Radl, Skopek, Komendera, Jäger & Mödritscher, 2013)	1A	2A	3D	4A, 4D
9	(Andersen, Gür, Hose, Jakobsen & Pedersen, 2015)	1A	2A	3B	4B
10	(Vert & Vasiu, 2015)	1A	2A	3A, 3B, 3D	4A, 4C
11	(Maramieri, 2014)	1A	2A, 2C	3A, 3D	4A, 4C, 4E
12	(Desouza & Bhagwatwar, 2012)	1A	2A	3D	4A, 4B
13	(Volpi, Ingrosso, Pazzola, Opromolla & Medaglia, 2014)	1A	2C	3B, 3D	4A, 4E
14	(Wang & Lo, 2019)	1A, 1B	2B	3B, 3D	4A, 4E
15	(Canares & P., 2014)	1B	2A	3D	4C, 4E

Table 3.2: Analysis of selected literature using classification framework

3.2.1 The context of the research

In this subsection, the context of the research, a distinction has been made in the region where the research took place for the purposes of a systematic literature review. The collected literature shows that most research into OGD and other open data was conducted in developed countries (figure 3.2). Only two studies have been conducted in developing countries such as in Taiwan (which is still classified as developing country in Kim and Heo (2017) and in the Philippines (Wang & Lo, 2019; Canares & P., 2014)). Most of the research was conducted in Europe and in the United States where in Europe the research was often conducted in the Netherlands (sometimes also in comparison with another country). The conclusion that can be drawn from this is that research into open data takes place primarily in highly developed economies. One reason for this might be that in developed countries the access to data is much greater and the number of people connected to an internet network is also bigger. A similarity in the literature studied is that all researches have made use of open government data by companies or governments or that the government itself has made the data available (and not company data that has been made available). More research is therefore needed into the effects of open data in business. In addition, more research is needed in other countries with different institutional contexts to see whether current insights, and future insights, also apply to those areas.
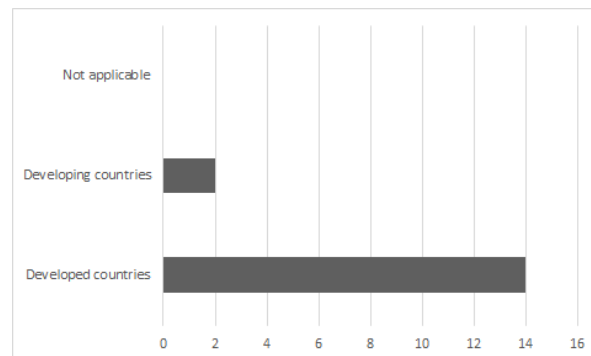
Figure 3.2: Number of times of open data projects in developed and developing countries in the literature researched

> **Proposition sub section 3.2.1:** "Research into open data has mainly been conducted in developed countries with mostly strong, stable economies. In the literature studied, it was mostly the government that made the data available. There is still little literature available on opening up data within commercial companies."

3.2.2 The method used in the research

In addition to the environment where the research was conducted, it is also important to draw a comparison in the method used to conduct the research. Research into open data is primarily qualitative research, whereby it is explicitly mentioned in a number of papers (Magalhaes, Roseira & Manley,

2014; Kalampokis, Tambouris & Tarabanis, 2013; Zuiderwijk, Janssen, Poulis & Van De Kaa, 2015). The other methods mentioned in the literature are often also qualitative. In the literature researched in this study, a quantitative method is only used in a single study (Marin de la Iglesia, 2014). In this research, use was made of a cost indicator to be able to map processes of open data and to easily display the direct cost savings that arise from it. Quantitative research is difficult to perform on open data because it can have different effects that are not measurable. Research from Marin de la Iglesia (2014) only looked at the costs and benefits of open data projects, making it easier to apply.

In the literature studied, in most cases use was made of either a case study or a survey as a method for conducting open data research. This shows that both case study research and survey are good methods when research is carried out into opening up data. For example, Jaakola, Kekkonen, Lahti and Manninen (2015) sets a number of conceptual requirements for open data services in its research, which are then assessed on the basis of a case study. This has shown that certain requirements that were set for an open data service also prove to be correct in practice (or at least in this one specific case investigated). Andersen, Gür, Hose, Jakobsen and Pedersen (2015) shows through a case study that opening up data alone is not enough to achieve set goals. According to this study, it is important to place links in the data that has been opened up in order to gain insight. Further explanations of case studies and surveys done on open data can be found in appendix A.

A study that is not mentioned in the earlier mentioned literature but is important to mention here is the dissertation of Zuiderwijk (2015). The researcher here also uses the design science approach in combination with a case study. This method worked well for this study. However, the focus was on OGD, so it is interesting to see whether this method also works with open data in a business environment. Zuiderwijk et al. (2014) also notes this in the study of 143 papers not one include a theoretical foundation for the design of open data infrastructures in business environment.

> **Proposition sub section** 3.2.2**:** "Previous research has shown that conducting a case study in combination with a survey is a good method to investigate whether data at organizations has the potential to be opened up."

3.2.3 The challenges of using open data

To be able to make validated statements about open data, it is important to also investigate the challenges of open data that emerge in the literature. In the 15 papers that were investigated in this study, all challenges were mainly traceable to four specific areas: legal, technical, economical and societal. These are also four of the seven perspectives that emerge in the aforementioned analysis of Zuiderwijk et al. (2014) (figure 3.1). Figure 3.3 shows that most of the challenges in open data research are in the field of societal and technical. For example, Susha, Grönlund and Janssen (2015) en-

dorses the most important motives for companies to adopt open data. This paper continues on the trend that open data can generate a lot of potential for companies. However, this study also shows that this adoption can vary greatly per company and that no generic conclusions can be drawn here. If companies want to do something with open data (open themselves or use open data databases), extensive research must first be done into the (potential) effects of this for the company. Other research has also shown that users' wishes often do not match the products made with the open data or the initial reasons why the data is made available (Vert & Vasiu, 2015; Desouza & Bhagwatwar, 2012). Solutions suggested and tested for this include citizen participation through private addition to open datasets. In this way, data becomes increasingly accurate, which can then be used for broader purposes. The conclusion is that research has shown that open data, supplemented by third parties, better meets the needs of the market. This, too, just like the point indicated above, depends on the situation and needs to be worked out further for each case in order to be able to properly chart the effects.

In addition to the societal impact and challenges that open data ecosystems entail, the technical preconditions and requirements are also essential for the proper functioning and outcome of open data. Open data projects should, for example, take into account the language and terms used and the user-friendliness of the open data (tools). Account must also be taken of the user-friendliness of the produced APIs (Application Programming Interface) for optimum potential of the open data projects / products (Andersen et al., 2015; Volpi, Ingrosso, Pazzola, Opromolla & Medaglia, 2014).

Economic challenges within open data projects often have to do with the uncertain costs involved in opening up the data or the uncertainty in the turnover of the products that is achieved after opening up the data (Jaakola et al., 2015; Susha, Grönlund & Janssen, 2015; Marin de la Iglesia, 2014; Lakomaa & Kallberg, 2013). Open data can also create many economic opportunities for entrepreneurs, as demonstrated in the research of Lakomaa and Kallberg (2013). This research shows that open data provides an acceleration in the process of data product creation which leads to innovation and can bring economic opportunities.

Open data projects must also comply with numerous legal requirements. For example, in the case study conducted by Vert and Vasiu (2015), the mobile applications that use open data must meet certain requirements to be able to make real-time adjustments. The legislation on GDPR that was introduced a few years ago (discussed in paragraph 1.5.2) also has a lot of influence on the open data ecosystem.

> **Proposition sub section 3.2.3:** "Projects that deal with open data should primarily take into account societal, legal, technological and economical challenges. For each project it applies that the size and the type of challenge may differ but also the impact of the challenge"
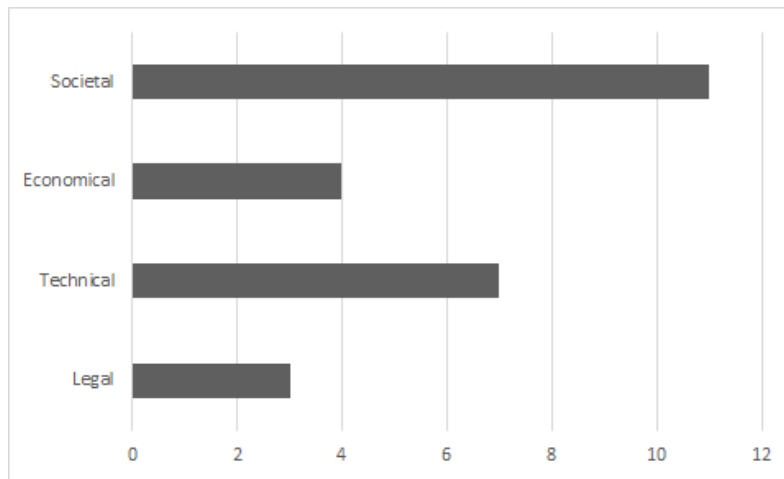
Figure 3.3: Number of times a challenge of open data is researched in the literature that has been studied

3.2.4 The effects of using open data

In this final subsection there has been looked at what was achieved after or during the study of open data. These effects are summarized under the categories listed in table 3.1. The categories are, among other things, compiled on the basis of Zuiderwijk et al. (2018), which describes various goals / effects of using open data in her paper (figure 3.4). These effects can all be linked to each other in certain ways. These effects with additionally effects added based on own insight were used in this analysis. The effects that emerge in the literature have a number of similarities and differences (figure 3.5).
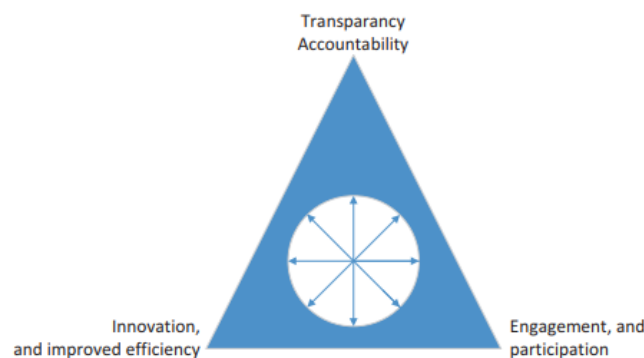
Figure 3.4: Objectives of Open Government Data Adapted from: Zuiderwijk, Shinde and Janssen (2018)

Opening data can bring many possibilities. For example, new markets can be opened and innovative IT solutions are made possible (Frankowski et al., 2015). In addition, much is written in the literature about the generation of social, environmental and economic value when using open data (Jetzek, Avital & Bjørn-Andersen, 2014). Open data could be a source of innovation in this, which, according to the European Commission, can generate an enormous economic boost in the form of billions of euros. The reuse of

geo information, statistical, weather and research data are good examples of this. All this can also provide new opportunities and business for many companies (Chan, 2013; Gurin, 2014).

In addition, Jaakola et al. (2015) shows that open data can achieve a lot as long as there is collaboration. Good communication is important for both users and developers to use the data properly. Users are sometimes not aware of the possibilities that developers create with the open data by developing various APIs. In addition to the possibilities that the APIs can create for users, the needs also differ per organization. For example, Susha, Grönlund and Janssen (2015) has demonstrated that in open data research it is difficult to draw conclusions that can be generalized. This is one of the reasons why research into open data is still at the reasonable beginning and much is still possible.

Research also shows that most effects are achieved with open data projects by creating citizen participation and company collaboration (Lakomaa & Kallberg, 2013; Jaakola et al., 2015; Desouza & Bhagwatwar, 2012). Because data becomes public, citizens and companies can create more business and thereby contribute to economic and social progress within a country. The involvement of companies and citizens in the government is also growing as a result. However, the literature lacks the next step that can be achieved with the effects that open data projects bring. Little is known, for example, about the (commercial) value that is additionally achieved when open data is used. On the other hand, almost nothing is written about the effects of projects where companies themselves open their data. The studies of Magalhaes et al. (2014) and Marin de la Iglesia (2014) also state that it is interesting to conduct more research on this.
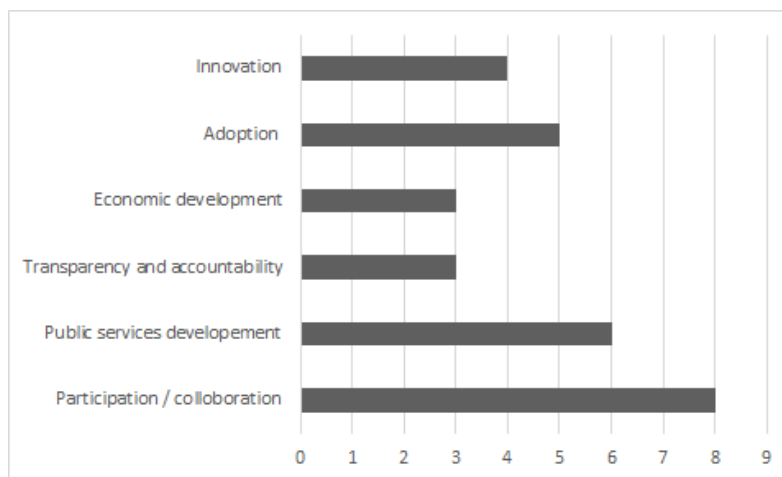
Figure 3.5: Numer of times an effect of open data is mentioned in the literature researched

> **Proposition sub section 3.2.4:** "The effects of open data projects for companies are broad and differ per case. Open data has positive effects on citizen participation, public service development and adoption"

3.3 KNOWLEDGE GAP

In the section 3.2, a total of four different factors were selected that are often highlighted in the literature on open data. The conclusions that follow from this include that it has already been demonstrated that open data projects can have (positive) effects on companies and that there are challenges in many areas. By open data projects at companies is meant here that companies use OGD in their processes to make it run more efficiently or to organize it differently. However, little research has been done into what the effect is if companies themselves make (parts of) their data public. This may have to do with the fact that companies prefer to keep their own data secret and do not have a good understanding of the effects if they publish their data publicly. At present there is too little insight into the value creation of open data by companies. In addition, research has also shown that too few open data studies have been conducted in developing countries and that most studies have used a case study approach or a survey of proven working methods.

3.4 SUMMARY CHAPTER 3

This chapter on the one hand gives an impression of part of the available literature on open data and on the other hand provides insights of important elements of earlier research into open data. These two points have been investigated on the basis of an extended literature review. This review confirms what has already emerged briefly in the introduction: current research into open data focuses primarily on government organizations that use open data sources or the government that makes data available itself. In addition, it was confirmed that the design science approach in combination with a case study is a good way of doing research into open data. In the current literature, studied in this research, there is little to be found about (commercial) businesses open up data themselves. This is probably due to the fact that there is no clear method available to identify the value of open data because many cases are context dependent. The answer to the research question posed at the beginning of this chapter, *What are important elements for commercial companies when considering opening up data?*, can now be formulated. Important elements when companies consider opening up data are mainly economic, technical, legal and social elements. In earlier research into opening up data, these elements emerged as important factors for organizations. These elements can be included in the compilation of the open data implementation method . By incorporating these elements in a certain way, companies get a better understanding of how they can open up data within their company. These elements mainly contribute to this by identifying the area in which this open data creates value for the company. The design of this so-called open data implementation method will be discussed further in the next chapter.

4 | DESIGN OF THE ARTEFACT: OPEN DATA IMPLEMENTATION METHODOLOGY

"The potential value of open data can only be unlocked if the data is made available in the first place"

It emerged from the previous chapter that currently too little research has been done into how companies can create value if their proprietary data is opened up. Companies are generally reluctant when talking about open data. A first thought that comes in mind to the person to whom the concept of open data is presented is that opening up data is at the expense of the turnover and gives the competition an unnecessary insight into the company's internal data. To stimulate an open data strategy at companies, a framework must be established in which the identified elements from the literature review from chapter 3 play a central role. In this study, this framework will take the form of an open data implementation methodology (ODIM) to provide insight into the value of open data for companies. The development of this method, the artefact according to the design science approach, will take place in this chapter. The first phase in this approach that comes after the identification of the problem is the phase in which the objectives of a solution will be defined. The phase will formulate an answer to the second sub question:

How can a methodology be designed to identify the value of open data?

By answering this sub question it becomes clear how this open data implementation method can be constructed and what this designed artefact will accomplish (the central question in the second phase according to design science research). If an interpretation has been given to this, the methodology can be made specific for open data projects in order to be able to apply it in practice. This is the next phase in the design science approach: the design and development of the artefact. This will also take place in this chapter, as a result of which an answer will also be formulated to sub question three:

What methodology can be implemented to let companies create value with open data?

As a result, the first three steps of the design science approach have been completed at the end of this chapter and the artefact, the ODIM, can be demonstrated in a real life case in chapter 5.

4.1   OBJECTIVES OF A SOLUTION: A METHODOLOGY

The solution presented in this study to implement open data within commercial companies is an open data implementation method. The purpose of this method is to make it clear to companies which datasets can be opened up and what value this creates for the company. In this way, commercial companies can apply this method and it contributes to the development of open data and can contribute to the overall potential of open data that is not yet fully utilized. When it comes to open data implementation, companies will have to look at how current systems can be arranged differently. And this can be a big step for many companies since data is of great value in the current era. In an article in the Financial Times, written by Foroohar (2019), it is shown, among other things, that the growth in the value of proprietary and personal data is not nearly stopped and that the economic value could double in the coming years. Such investigations will not exactly motivate companies to make their data public. This therefore requires a methodology that provides insight, to contribute to the open data strategy of a company, what the value is if certain datasets are opened up.

Companies will have to organize their systems differently when data is opened up. This requires a system design approach whereby the current situation is analyzed and a proposal is made for a new design. In the context of this research, the methodology must therefore consist of an analysis of the current data situation at a company and a new data system design will follow in which open data plays a key role. Methods for designing these types of systems are common in the available literature. In this research it will be decided to choose one of these methods and to adjust it so that the steps that are included here are suitable for open data implementation projects. In the following section a number of these standard system design approaches are named and it is explained why this research will continue according to the system design approach according to Dym et al. (2008).

4.2   SYSTEM DESIGN APPROACHES

There are many methods for designing a system in the literature. In the sub sections below, three of the methods will be explained with associated advantages and disadvantages. The choice to compare these three, and not for example anymore, was made because this creates a clear enough picture of the possibilities of system design. In this research, the most important result is the design of the open data methodology, which focuses on that. The following three system design approaches are explained: the waterfall method (Royce, 1987), the spiral method (Boehm, 1988) and the engineering design process method (Dym et al., 2008). The following sub section first explains why it was decided to compare these three approaches. This explanation is based on the book written by Armstrong and Sage (2000): An introduction to systems engineering.

4.2.1 Systems engineering elements

Much has been written about systems engineering in the available literature. Many different definitions are mentioned and many different approaches are described how systems engineering problems could be tackled. To design the open data implementation method, an approach must be chosen. Given the scope and time frame of this study, not all available approaches can be compared. However, the book written by Armstrong and Sage (2000) does offer a lot of information about the available suitable approaches. Before discussing the suitability of the approaches listed in the following sub sections, it is good to first explain system engineering as an overall approach. This is useful because it makes clear why a systems engineering approach is of value for the development of an open data implementation method. The most relevant goals of systems engineering are listed below that also serve as important objectives within the open data implementation method. These objectives are literally reflected in Armstrong and Sage (2000):

1. System engineering processes should encompass all phases of the system life cycle, or life cycles as the case may be, including transition between phases.

2. System engineering processes should support problem understanding as well as communication among all interested parties at all phases in the process.

3. Systems engineering processes should support quality assurance of both the product and the process that leads to the product.

4. Systems engineering processes should be based upon methodologies that are teachable and transferable and that make the process visible and controllable at all life-cycle phases.

5. Systems engineering processes must support both system product development and system user organization.

6. Systems engineering processes should support quality, total quality management, system design for human interaction, and other attributes associated with trustworthiness and integrity.

The above-mentioned elements of systems engineering processes are the most important elements that must also be reflected in the open data implementation method. By explaining this here, it shows that a systems engineering approach is a correct method to develop the open data implementation method.

Many different methodologies can be distinguished within systems engineering. According to Armstrong and Sage (2000), a good system design approach must meet at least the following few conditions:

1. In the method of systems engineering it is generally desirable to allow for learning as the life cycle process is iteratively repeated

2. The phases in the method are maximum seven in total

3. Problem formulation, analysis of alternatives and interpretation and selection are important elements in the systems engineering method

The three most important system design approaches that meet the systems engineering criteria mentioned above include: the waterfall method, the spiral method and the engineering design process method (Armstrong & Sage, 2000). To make a choice between these three methods, they are first explained in the following sub sections.

4.2.2 Waterfall method

The waterfall method is a method whereby the development of the new system (the method was originally devised for the development of software systems) regularly runs down (Royce, 1987). This process involves a number of phases: definition of the problem, basic design, detailed design, design building, testing integration and management and maintenance. The intention of this way of working is that a project is built up in different phases. In the implementation of this method, the second phase is only started once the first phase has been completed. Advantages of this method are that when errors are discovered it takes little effort to fix them (less time and money because a phase must first be completed perfectly) and the method is widely known, so that many people who work with it have experience with it and therefore work with it more easily. A disadvantage is that the method does not assume that the requirements of the system change and with the influence of external factors this can happen. Another disadvantage is that a lot of time is wasted because people cannot continue with the next step if the previous one is not yet finished (while this is often the case in practice). Figure 4.1 shows a schematic representation of the waterfall method.
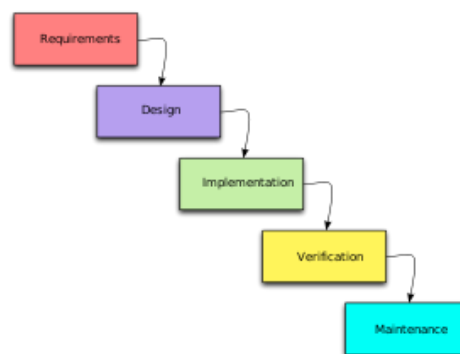
Figure 4.1: Simplified representation of the waterfall method. Adapted from Royce (1987) Page 329

4.2.3 Spiral model

The spiral model has a total of four phases that are executed in succession. After the implementation of all four of these phases, it can be decided to stop the project or repeat the four steps (Boehm, 1988). The four phases are as follows: determining the purpose of the system including the alternatives and the limitations, evaluating the alternatives where the one with the least risk is chosen, developing and verifying the system or product and lastly carrying out a review. A major advantage of using this method is that the most risk-bearing solutions are removed at the start of the project, so that there is great confidence in the project. On the other hand, planning is almost impossible because it is often not sufficient to estimate the risks in advance. So there is much uncertainty about the completion of the project. Figure 4.2 shows a schematic representation of this spiral model.
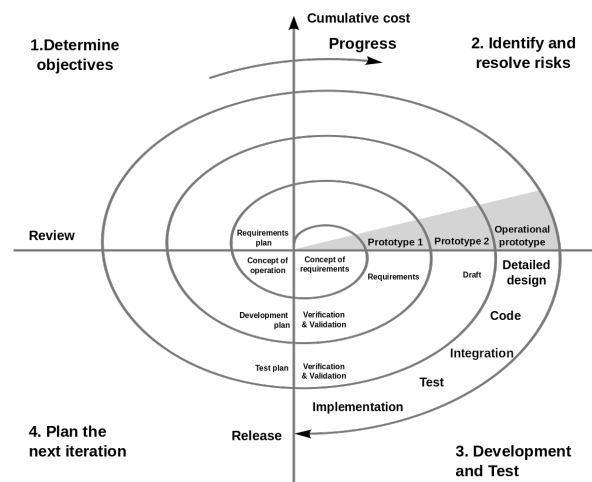
Figure 4.2: Simplified representation of the spiral model. Adapted from: Boehm (1988) Page 4

4.2.4 Engineering design process

Engineering design is a method used by many engineers to identify and solve problems within systems. There are many processes available within the engineering design method, but they can often all be traced to a number of basic steps to break down complex problems into smaller parts. These steps are: problem definition and framing of the problem, conceptual design stage, preliminary design, detailed design and design communication. Advantages of using this method is that it is a systematic and iterative method because the steps that occur in it are repeatable and there are clear constraints (in the form of requirements or functions, for example) that are bound to the design. A disadvantage is that this method can cost a lot of time and manpower. A schematic representation of engineering design process compiled by Dym et al. (2008) is shown in figure 4.3.
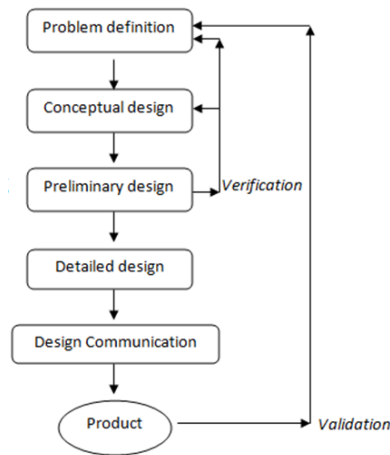
Figure 4.3: Engineering design process based on Dym, Little, Orwin and Spjut (2008)

4.2.5 Chosen design approach for compiling the methodology

Three system design approaches are explained in the previous three sub sections. To develop the open data implementation method, it was decided in this study to apply the engineering design process according to Dym et al. (2008). The main reason for this is the iterative nature of this process. There is a chance that companies using the method are not immediately sympathetic to the process of making their data available for free. This makes it useful if there are interim evaluation moments to determine whether the data will continue to be opened or whether there will be a stop or whether another data set will be chosen to be opened up. By designing the open data implementation method in such a way as Dym et al. (2008) did in the engineering design process, the chance of using and further developing the method is increased. The following section describes which steps will occur in the open data implementation method and how they differ or correspond to the standard engineering design method.

4.3 HIGH LEVEL OVERVIEW OF THE OPEN DATA IMPLEMENTATION METHOD

It has emerged from the previous sections that the engineering design process is used for the development of the artefact by Dym et al. (2008). With this, the second phase of design science research (the way in which this research is structured) can be completed. The objectives of a solution is a methodology according to the steps of the engineering design process that are adjusted to give insight in which data can be opened up and which value this will create for companies. The interpretation of this method on a high level can be seen in figure 4.4. One of the advantages of this method is that there are interim evaluation moments for companies that apply the method. For example, before a lot of money or effort has been put in, they can choose

to make another dataset public or to stop the open data strategy. This section also explains the differences between the phases that occur in the engineering design process and in the ODIM. From section 4.4 the ODIM is filled in completely so that it can be used in practice in a real-life case.
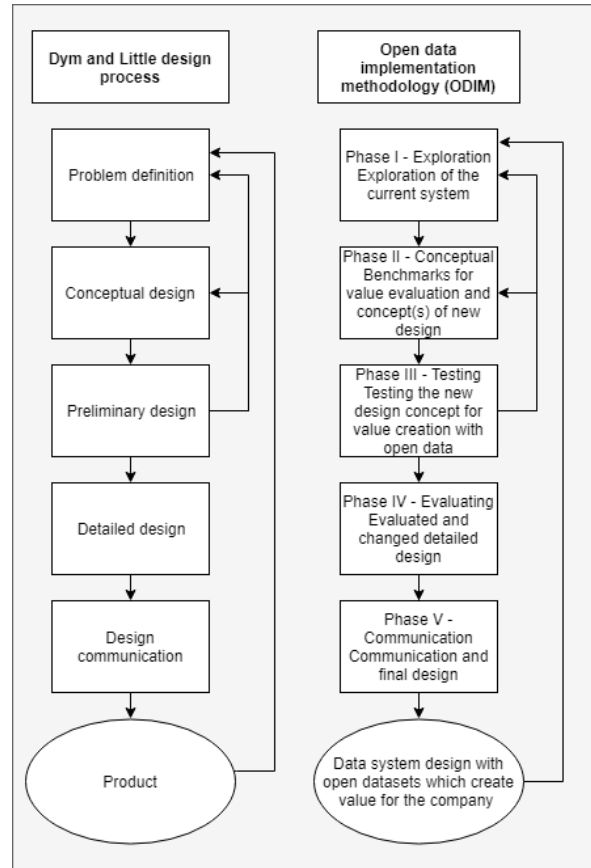
Figure 4.4: Left: classic engineering design process by Dym, Little, Orwin and Spjut (2008). Right: the open data implementation methodology (ODIM), the artefact in the design science research approach of this study

To indicate the structure, the ODIM uses phases with corresponding numbers. These are parallel to the steps in the engineering design process. An overview of the differences is shown in table 4.1

4.4 DETAILED DESIGN OF THE OPEN DATA IMPLEMENTATION METHOD

In the previous section a high level overview of the open data implementation method, the artefact in this research, is presented. To be able to use this artefact in the next phase of the design science approach, the demonstration phase, the steps must be explained much more explicitly. From the previous section it has only become clear what the outputs are of each phase that are used in the next phase. To arrive at an open data system that provides value for commercial companies, the ODIM will be further explained in this part of chapter 4. Each section has a title that corresponds to the naming in

	Outputs of the phases according to Dym, Little, Orwin and Spjut (2008)	Outputs of the phases in the Open Data Implementation Methodology
Problem definition / Phase I	Customer requirements Problem statement Objectives Functions of the system	Requirements for open data systems Objectives of stakeholders current system Analysis current system Design of current system
Conceptual design / Phase II	Multiple designs Metrics for objectives	Benchmarks for open data value evaluation Desings of new system including open datasets
Preliminary design / Phase III	A chosen design Analysis of chosen design	Measurements of value of open data in new system
Detailed design / Phase IV	Proposed design and design details	Evaluated design New or same open datasets as previous design
Design communication / Phase V	Final report including drawings Fabrication specifications	Final design of an open data ready system which has the potential to create value in certain areas which is tested

Table 4.1: Differences between the output of the phases proposed by Dym, Little, Orwin and Spjut (2008) and the open data implementation method

figure 4.4 (phase I, phase II etc). At the end of this section it will be clear for each phase in ODIM why it occurs in the method, what it contributes to designing an open data system and many other details.

4.5   DETAILED DESIGN ODIM PHASE I – EXPLORATION

The first phase of the ODIM is one of the most extensive phases in the entire design of the method. Because it is important to map the current systems of the company and to be able to make a good proposal later for a new system with open data, various analyzes must be carried out. All of these steps are described in the following subsections and can be performed simultaneously to each other when the method is performed in a team with more people.

4.5.1   Process models

Almost all companies in the 21st century have to deal with data. Within almost all business processes, data flows play an essential role for communication and the structure of the processes. The amount of data is only growing and the complexity to analyze this data is also increasing (Ben Ayed, Ben Halima & Alimi, 2015). Due to the growth in the amount of data and the increase in complexity, the first step in opening up data is finding a suitable dataset. Certainly at large institutions, this can be difficult to identify because the data structures here are often even more complex. Opening up data is an example of a radical innovation within the system of existing business processes. Changes in systems are accompanied by failures and success (Maidique & Zirger, 1984) which may have the consequence that the first dataset that is chosen to be opened is not exactly the right one. Later in the process it is possible to go back to this first step to choose a different data set. This demonstrates the iterative nature of opening up data in business processes and arguments for the importance of process models in the first phase. An important element in this first step is therefore to map these business processes in great detail. In this step it must be clear where the

input data comes from, what the actions are that happen with the entered data and what the output data is.

A suitable method for charting these processes in detail is by visually displaying them in the Business Process Modeling Notation (BPMN). This notation is a well-known standard for capturing business processes in early phases of system development (Dijkman, Dumas & Ouyang, 2008). By placing the processes in so-called "swim lanes", it is clearly shown how the data flows through the system and what happens with it. To obtain this information for these process models, use can be made of interviews, own observations or other analysis techniques.

> Map and describe all processes in detail so that it becomes clear what the input data is, what the actions are that happen with the data and what the output data is. The Business Processes Model Notation is a suitable tool for this.

4.5.2 Involve stakeholders

To guarantee the relevance of opening up the data, it is important to involve the stakeholders involved in the process as early as possible. Because a real life case is conceptualized in this phase of the methodology, knowledge acquisition by the stakeholders involved is of great importance (Kotiadis & Robinson, 2008). To make a choice as to which dataset can be compiled, an inventory must first be made with, for example, the users or the developers of the data. In this step, an analysis must be made of how the data is developed and how users use it. By already doing this in the exploration phase, the technical aspects of the data become clear early in the process. Also involving the stakeholders involved at such an early stage increases the chance that they will remain satisfied when the data is opened up. This step can also be used to discuss common fears and misunderstandings with those involved. Since opening up data can result in fear for many people (for example, loss of market position or turnover), this is a good time to make this known (Beno et al., 2017). Once again interviewing is a good method for this.

> Ask the stakeholders involved about the development and use of the data. Detailed (technical) knowledge about the data becomes clear after this step.

4.5.3 Analysis of current system

In order to substantiate why certain data sets must or can be opened up within a data system, several analyzes can be performed. Given the iterative nature of this methodology, it is possible to return to this step later to perform the analysis again and to choose another dataset to open up or stop the process. Four examples of how to choose a dataset are the following:

1. **Asking the community**
   Prepare a list of data sets that have the potential to be opened up. Publish this list on a webpage or other platform to receive feedback from the community. Ensure that it is possible to provide feedback on the datasets in an easy way. This contributes to the quality and amount of feedback. Mailing lists or interviews with interested parties can also be used here. Organizing a consultation event can also be a useful way of using this method. By consulting the community you gain insight into the needs of potential users and you can better estimate the value of opening up the data.

2. **Cost basis**
   Choosing a data set can also be based on the costs of compiling and maintaining the data set. When these costs are high, there is a greater chance that others would like access this. The chance of a (too) large loss on sales is considerable, since the data sets with the highest costs often also result in high sales. This can cause restraint with the owner(s) of the data. Based on this argument, an analysis can also be carried out of the data sets on which the least turnover is achieved. By first opening up this data and actively promoting it, it is possible to analyze well what the reaction of the users are and the value that is achieved with this. If this is experienced positive, the datasets with higher turnover or higher costs can be opened up.

3. **Ease of the release**
   Instead of viewing which data has the most value when it is opened, it can also be viewed which data is easiest to open. Easy, and therefore often small, data sets can act as a catalyst for larger changes within systems (Huijboom & van den Broek, 2011). A danger that comes with this way is that small data sets can be of lesser value for users.

4. **Observe peers**
   The fourth and final way to choose a dataset to open up is to look at other companies. By observing what these do, a lot can be learned from the strategy and adopt or adjust it as desired.

> Choose a dataset to open up based on one or more of the following criteria: community based, cost based, ease of release and/or based on observations of peers

4.5.4 Requirement analysis

To implement open data within a current system, this data must meet certain requirements. In this step in the method it is therefore important to conduct research into the requirements of open data. This is best done through literature research because there is already a lot available about the requirements of open data. Some requirements will also be case specific. These will become clear during the interviews in which the process is being analyzed or in later phases of the method. These functional requirements describe how

the new system, including the opened data, must behave. These requirements refer both to the way in which the data is opened up and how the entire new system should behave. Functional requirements define the specific functionality that shows how a system can be used (Stellman & Greene, 2005). The focus in this phase of the method is on functional requirements because the study is about what kind of data users find functional and usable to be opened up. This study does not elaborate on non-functional requirements such as sustainability, accessibility and scalability. The reason for this is that this study examines the potential value of open data and therefore the functionality. Because this study does not elaborate on non-functional requirements, one of the assumptions is that these requirements are met.

The functional requirements are described on the basis of the chosen case. These requirements will for the most part be widely applicable. This means that the requirements compiled can be applied to multiple cases. This also contributes to the reproducibility of the method for other cases at companies. Nevertheless, there will also be case specific requirements that can or will be unique for each company. These must be described separately so that this distinction is clear.

All case specific functional requirements can be identified in various ways. Good methods for this are the use of a case, conducting interviews, drawing up on the basis of observations, analyzing existing processes and analyzing existing literature.

4.6 DETAILED DESIGN ODIM PHASE II – CONCEPTUAL

After an extensive analysis of the current situation has been carried out, a proposal can be made for a new system design in which open data plays a central role. To test this design in phase III, benchmarks are drawn up in this phase II to determine the value of open data in the new system. These two steps are explained in the following sub sections.

4.6.1 Composing Benchmarks

Benchmarks must be drawn up to ultimately be able to measure the new system on the value op the open data. This is also the step where the elements identified in chapter 3 appear. In this chapter the important elements for commercial companies when considering opening up their data are discussed. Various challenges and effects of open data were named in this, which are converted into value evaluation elements in this step. These benchmarks are widely applicable to open data initiatives. By testing open data projects based on the same benchmarks, a good comparison can be made between these projects. It can also be determined in this way when an open data initiative is successful and when it is not (the available literature

sometimes refers to KPIs and sometimes to benchmarks, both terms aim to measure the value of the open data in the new system (Parmenter, 2015)).

The research of Susha, Zuiderwijk et al. (2015) compares different open data benchmarks in a meta-analysis. One of the studies mentioned in his study (to give an example) is the preparation of open data KPIs by the European Union. The data portal that the EU has developed has been compiled by the European Commission in collaboration with a consortium consisting of Capgemini Invent, including INTRASOFT International, Fraunhofer Fokus, con terra, Sogeti, the University of Southampton, Time.Lex, 52 North and the Lisbon Council . When evaluating the success of an open data initiative, it is important to not look at the quantity of the data published but focus on the demand for the data. To make this measurable, the consortium has developed an open data barometer for this.

Ultimately, the study of Susha, Zuiderwijk et al. (2015) investigated how open data benchmarks differ in measuring open data maturity and what can be learned from this. Among other things, this research shows that many models used in the literature to compare open data projects can be traced back to the general definition of open data. In this study, this definition has already been included in the requirement analysis (phase I) and serves as a starting point before the data can be opened up. In other words, this study mainly tests the so-called open data readiness, which conditions must be met in order to be made public. The actual impact of the data opened up, and therefore the potential value for a company, is not explicitly mentioned. Therefore, other factors are needed to determine the value of such an open data project. To make this impact explicit, (Susha, Zuiderwijk et al., 2015) shows that the measurements for this should cover various levels. A number of examples of these impact levels are: increasing efficiency and transparency of business, economic growth driven by open data, environmental benefits, educational support available and much more.

THE CHOSE FOR SPECIFIC BENCHMARKS FOR EVALUATION OF OPEN DATA
The previous sub section has shown that measurements for open data benchmarks should cover various levels. To be able to determine the value of the data that has been opened up, several aspects must be considered. In this study the choice was made to base these aspects on the most important considerations that companies make when opening up data. This was investigated in sub question three in chapter 3. The result of this is that important elements for opened data are economic, technical, legal and social elements. To test the new system, with the data that has been opened up, benchmarking must take place within at least these domains. This is also confirmed in the research conducted by Attard, Orlandi and Auer (2016). This research adds a fifth benchmark: the organizational dimension. When the open data is tested within all five dimensions, the potential value of the new system is correctly estimated. That is why it was decided to use these five dimensions as benchmarks in this study. Largely due to the results from sub question three, supplemented by the study of Attard et al. (2016). Each benchmark is explained in more detail below.

- Technical dimension: value creation within this dimension is often in the dataset itself. Because open data must comply with a certain format (machine process-able and non-proprietary), it must make it easier for users to use the data. In addition, the data ambiguity is also reduced because the format in which the data is delivered makes the data set easier to understand.

- Policy / Legal dimension: to make datasets not open but closed, a lot has to be arranged. Licenses must be requested and kept up to date. Certain laws and regulations can also ensure that valuable data sets are not used at all and therefore do not create value.

- Economic / Financial dimension: this dimension can form a barrier if no money is made available within the company to identify which dataset can be made public. But in this dimension, the benefits, and therefore the value that can be created, lie mainly in the innovations that are possible. By opening up data, companies can build new data products with this. In addition, it also offers the company value by selling products that work with the data that is opened up. By opening up the data, it becomes clear to users what the dataset consists of. However, this does not mean that the user can actually do something with it. (Economic) value can be achieved here by companies and other organizations.

- Organizational dimension: the value creation within this dimension focuses primarily on increasing efficiency within the organization. By opening certain datasets, users can detect errors and adjust them themselves or make a proposal to adjust them. By partially outsourcing this task to the users, the data set will be more accurate and therefore provide more value for the users.

- Social / Cultural dimension: this dimension regards the feeling of the public towards open data in your organization. An important factor is the transparency of the company. It could reduce corruption by enabling greater transparency. It also promotes participation of citizens and other organizations by providing more insight into the data. The accessibility and openness of the company is improved by being more transparent, which can provide added value for the company.

4.6.2 Concept design of the new system

The next step is to design the new system. In phase I it became clear which dataset has the potential to be opened up and the requirements of the open data. The combination of these offers possibilities for a new system. By talking to those involved and based on own observations, the best assumptions can be made about what such a new system should look like.

In this phase of the method it is useful to visually display this new design in a clear drawing. This contributes to the structure and overview of the case. This drawing must make clear which dataset is opened up in the entire

system and what the interaction is between subsystems and stakeholders. A detailed description of this design is required. In this phase it must also be made clear whether all requirements have been met and in what way. Dym et al. (2008) describe that it is best to develop multiple concept designs for comparison. In this way, the best design ultimately emerges after completing the entire method.

4.7 DETAILED DESIGN ODIM PHASE III – TESTING

Now that it has become clear from the previous two phases of the methodology what the new system can look like, it is time to test this design. This is done on the basis of the benchmarks drawn up in the previous phase. A good way to test this design is to facilitate an expert panel workshop (EP) with different people who influence or are involved with the system. This is a good method because the workshop participants are familiar with the current system and can estimate the value of elements that change in this system. These experts can also be the ones who will eventually build the system, so involving them at this stage is a good choice. The new system can be properly evaluated with these experts. The following framework in figure 4.5 is applied within this EP workshop to test the value of the open data. The first part consists of a generic brainstorm about the value creation of open data and the second part about the value creation of specifically the new system. For a detailed explanation of the expert panel workshop and its setup, see appendix D. Other ways to test it include creating a simulation model or actually opening up the data in a pilot.

The iterative nature of this methodology also returns in this phase. When the results of the EP workshop are that the wrong dataset with too little value has been chosen in the entire system top open up, we can go back to an earlier phase before the system is further developed. It is also easy to decide after this step to stop using an open data strategy and not to open up datasets.

4.8 DETAILED DESIGN ODIM PHASE IV – EVALUATING

It emerged from the previous phase that changes had to be made to the system that was evaluated or not. When the user of this method has arrived at this step, the design can be worked out in detail in its entirety. This also includes the changes from the previous phase. In this phase, almost the same steps can be carried out as described in Dym et al. (2008) in this phase (detailed design): "We now articulate our final design in much greater detail, refining the choices we made in design down to specific part types and dimensions. We use detailed design knowledge and procedures expressed in specific rules, formulas, and algorithms that are found in design codes "(Dym et al., 2008). In this phase, the results of the design evaluation, the
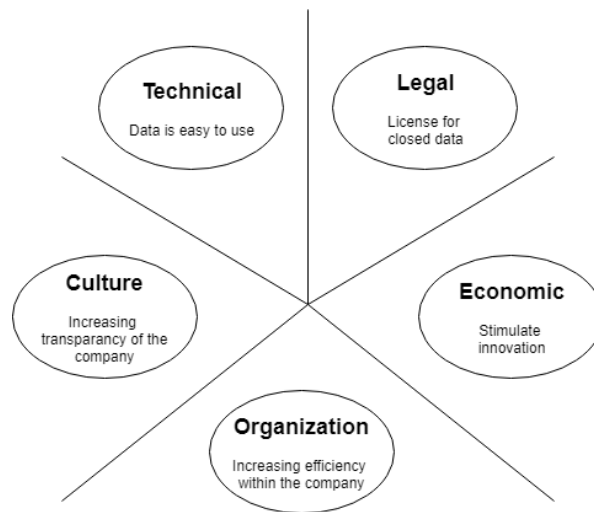
Figure 4.5: The five value creation dimensions of open data

points at which open data in the new design can be of great value, are also discussed.

4.9 DETAILED DESIGN ODIM PHASE V – COMMUNICATION

The steps that are followed in this final phase also correspond to the steps proposed by Dym et al. (2008): "We now spell out and present our design process, the resulting final design, and its fabrication specifications. In practice, the designer will usually have already developed much of the documentation along the way, and this communication phase will be more about tracking and organizing prior work products than writing a "new" report from "scratch" (Dym et al., 2008). ' The final design is therefore communicated in this final phase of the methodology. In the case of an open data system at a commercial company, it can also be of value to use a business model on this. One of the earlier steps shows what the most important benchmarks are in the new open data system. According to Susha, Zuiderwijk et al. (2015), one of the most important conclusions is that these benchmarks must be converted to the wishes of the users and made explicit. To exploit this value that emerges after measuring the benchmarks of open data, to maximize the benefits and to enable the creation of innovative products and services, a business model must be designed before the data is opened. This is required to ensure that the products and services generate the necessary value proposition and meet the needs of customers / users and eventually generate substantial value. The 6-Values Open Data Business Model Framework designed by Ahmadi Zeleti et al. (2016) ensures that businesses are taking into account all aspects of an effective and efficient business model and understand the effect of the different aspects on each other.

In figure 4.6 the 6-V business model is shown schematically.
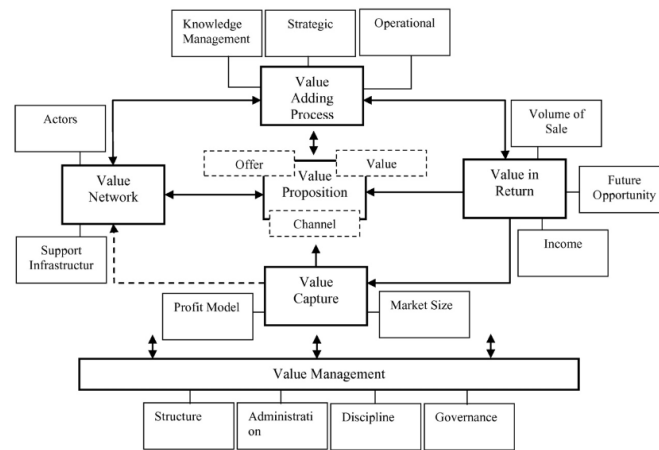
Figure 4.6: The 6-V business model conceptual framework designed by Ahmadi Zeleti, Ojo and Curry (2016)

4.10 SUMMARY CHAPTER 4

This chapter answers two of the prepared sub questions that fit phase two and phase three from the design science approach. To create value with the help of open data at commercial companies, this study proposes the development of an open data implementation method. This is the objective of the solution. The first sub-question asked in this chapter is *How can a methodology be designed to identify the value of open data?* The answer to this question can be formulated as follows: Since a new system will be designed for companies in which open data has a central role, this methodology can be constructed from a process design approach. A lot of research has already been done on this and in this study it was decided to adjust the engineering design process approach by Dym et al. (2008) to make a suitable open data implementation method. So to identify the value of open data the methodology can be designed using a engineering design process approach. This was then used to formulate an answer to the following sub question *What methodology can be implemented to let companies create value with open data?*. The answer to this question is the artefact in the design science approach, the open data implementation methodology. By implementing this methodology companies can create value with open data after following five phases that are of an iterative nature (figure 4.4: exploration, conceptual, testing, evaluating and communication.

5 | DEMONSTRATION OF THE ARTEFACT: A CASE STUDY ANALYSIS

In the previous four chapters, the first three phases of the design science approach have been completed. It has been identified that there is a lot of potential around open data and that this potential is far from being fully utilized and that companies (rather than just government organizations) will have to cooperate in the open data environment to be able to use this potential. For this, values of open data have been identified that have been processed in the developed artefact of this study: the open data implementation methodology (ODIM). This is (conceptually) described in its entirety in the previous chapter 4. The expected result of this method is that when it is used by commercial companies, these companies gain insight into which datasets can be made open and what kind of value this can create. This chapter discusses the next phase in the design science research approach: the demonstration phase. The ODIM is applied to a real life case to test whether an open data strategy is (or can be) valuable for a commercial company. This chapter provides an answer to the following sub question:

What value will opened up data create when applying the open data implementation methodology on a real case?

The results of this chapter will be two-fold. First of all, the case gives a lot of insight into how the methodology can be used and whether value is created for the chosen company in this specifically chosen case when an open data strategy is used. Secondly, this sub question also contributes to the overall evaluation of the proposed methodology. The results of this chapter will therefore serve as input for the next phase in the design science approach. In this next phase, the evaluation phase, the method will be adjusted where necessary.

First, the first section describes the basis on which the case was chosen. The context of the case is then explained and described in detail. This contributes to the information required before the method is applied to the chosen case. After it is clear on the basis of which criteria the case has been selected and described in detail, the method is applied to the case based on the five different phases described in the ODIM.

5.1 CASE STUDY SELECTION

An important aspect within case study research is the selection of the case. The limits of the generalization of the research results are made clear by a

correct explanation of the selection of the case (Eisenhardt, 1989). According to Eisenhardt (1989) cases can be selected in two different ways: statistical sampling or by theoretical sampling. With statistical sampling, statistical evidence is collected on the distribution of variables within a specific population and the case is selected on the basis of this information. With theoretical sampling a case is chosen because it is expected that it will repeat an earlier case and thereby validate this case, expand an existing theory or because the case provides a good example. Because research in the field of open data is still relatively new and this study contributes to theory building concerning open data, the selection of the case in this study is based on theoretical sampling. Also due to the fact that a case study can evaluate whether the applied methodology is valid, this research method is also suitable for this study. To test this designed methodology it would be better to test it on multiple cases. On the other hand, when a single case is done, it can be worked out in detail and these results can also be valuable. This further explanation why a single case was used in this study is explained in chapter 2.

5.1.1 Criteria for selecting a suitable case study

To test whether the ODIM is suitable for developing a new system at companies where open data plays a central role and to test if this provides added value for the company, a suitable case must be selected. To select this case, a number of criteria have been drawn up which are listed below.

1. There must be enough data available for the researcher to be able to make detailed statements about the existing processes. People involved in the current design of the system must be willing to help and provide information.

2. Within the case there is already experience with the use of open data or already with opening up datasets. This contributes to the fact that those involved want to contribute in thinking about what a new system can look like in which open data plays a central role. Certainly in the first case in which the ODIM is tested is this useful.

3. The case relates to a commercial company, preferable in the Netherlands. As discussed in chapter 3, little research has been done into open data within companies, which is why this is an important selection criteria. The researcher of this study lives in the Netherlands, which facilitates access to information.

4. Several stakeholders are involved in the case that represent different or corresponding interests. This ensures that the case can be viewed and analyzed from multiple angles.

5. Within the company there must be a kind of motivation to switch to a different strategy with regard to certain data sets or data products. Since the research field around open data is relatively new, and the method will also be tested for the first time, this selection criterion has been chosen. When companies earn a lot of money or otherwise get

a lot of value from their current (closed data) systems, it can become difficult to introduce an open data strategy.

The case that was chosen to apply the methodology is a case at PostNL Data Solutions. The researcher of this study was given the opportunity to test his methodology within this commercial company. The context of the chosen case follows in the following section. This is followed by an explanation as to why this case meets the above criteria.

5.1.2 Context of the chosen case – The Postcodetabel

In the Netherlands, mail delivery is for the most part done by one company. The letters and parcels are delivered to individuals and companies throughout the country. The way in which these letters are delivered has had quite a development in the Netherlands. Where first the mail deliverer walked from city to city, nowadays the mail is delivered with cars, trucks and electric bicycles. Mail sorting is an important part of the mail delivery process. Before 1977 this was done by employees who were educated to do this. By memorizing all the city names in the Netherlands, they were able to perform their work as good as possible. By linking the city name to a street name and a house number, the letter was delivered to the right person in almost all cases. In the fourth quarter of the twentieth century, this process changed as postal sorting companies started using computer-controlled machines for sorting mail. This made the process much more efficient and careful. However, this renewed process required an additional element in addition to the city name, street name and house number. This element to ensure that the sorting process went smoothly was the postal code. Simultaneously with the introduction of the postal code, all households and businesses in the Netherlands were sent a large red book with the postal codes linked to the correct addresses. This book was published by the PTT (the owner of the postal code and the largest postal delivery company in the Netherlands) who ensured that all parties within the logistics market were provided with the correct information regarding address data.

In the years that followed, however, a lot has changed in the logistics market and specifically within the address and postal code market. Due to digitization, far fewer letters have been sent in the Netherlands in recent years. On the other hand, due to this same trend, and with it the rise of the e-commerce market, the growth in parcels in the Netherlands has increased enormously (see figure 5.1. People in the Netherlands order their products such as clothing, shoes and their daily groceries more and more online, so the expectation in this market is that the decline in letters and the increase in the demand for parcels will only continue to increase (Ministerie van Economische Zaken, 2017). This is putting increasing pressure on logistics companies that have the task of bringing these products to consumers, which is one of the reasons why a great deal has changed within the PTT. Where it used to be a state-owned company, since 1989 it has been completely privatized to respond to these changes and is nowadays known as PostNL.
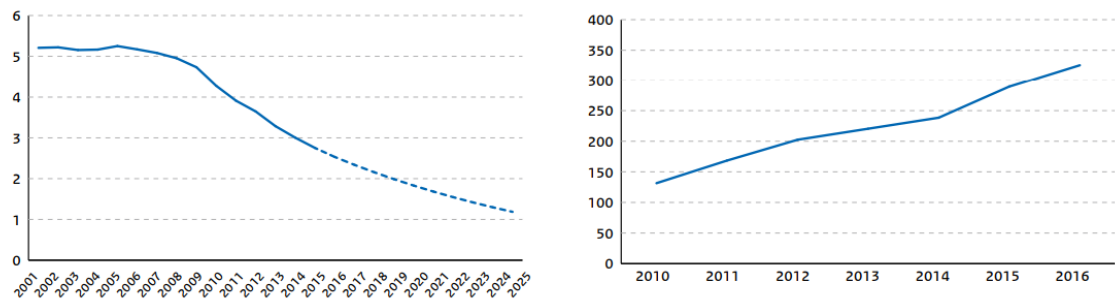
Figure 5.1: Changes in mail volume in billions (left) en parcel volume (right) in millions

Due to the growth in the parcels market and the decrease in the mail market, the interest of (e-commerce) companies regarding postal code and address data is also changing. Even more than before, it is important for these companies to have the correct data. This ensures less return of parcels, which contributes to a better environment and a reduction in costs. PostNL has set up its own department to assist companies in this: PostNL Data Solutions (PostNL DS). This department makes digital products for customers. The establishment of this department within PostNL also has to do with the major changes that have taken place in the "digital world". Whereas previously households and companies were sent the red book with all the correct address and postal code data, everyone now has much easier access via the internet. This red book was replaced from the rise of the internet (around 25 to 30 years ago) by (to this day) the largest product (in terms of sales and turnover) within this department: the Postcodetabel (PCT). The PCT is sold to companies with the aim that these companies have the correct postal code and address data of their current and new customers. By providing a dataset in which all names of the cities are correctly linked to other properties such as the postal code, street name and house number, value is created with this product for customers.

> **Example for which purpose users are using the PCT:** If you order something on an e-commerce website, you must enter your personal details when paying. If you only enter your postal code and house number, the street name of where you live often appears immediately. If this is the case, this is due to a user who has purchased the PCT and implemented this in their own system.

Because PostNL is the legal owner of the postal code, and the brand of this company is good in the Netherlands, many customers came to PostNL to have this correct dataset. However, because more and more people and companies can also obtain this information online (often for free), the value of this product has fallen dramatically over the years. In fact, a few years ago the Dutch government came up with a free alternative to this database in the form of BAG data (Basisregistratie Adressen en Gegevens). PostNL then filed a lawsuit for this because they felt that they had exclusive rights to these data because they are the legal owner of the postal code. They then lost

these, whereby the Dutch government and PostNL agreed on the "Convenant inzake postcodes" (Ministerie van Infrastructuur en Milieu, 2014). This covenant states that there is a collaboration concerning the allocation and issuing of postal codes. Because PostNL had lost this court case, interests arose for both parties for having their own database concerning postal code and address data. For PostNL this is the PCT and for the government this is the BAG. Having two databases concerns all addresses with the correct information from all over the Netherlands is inefficient (this has been confirmed by several employees at PostNL B). It is therefore possible to investigate whether this is not more valuable if this is a central database. Since BAG is open data and the PCT is closed data, the ODIM from chapter 4 could be applied to this case. The aim is to test whether it is valuable to open the PCT (or parts of it) and thereby can be combined with the BAG data. The focus on the first part, opening up the PCT. This study does not elaborate further on the possible creation of a central database in which the PCT and the BAG are merged. Another reason why the PCT is suitable for testing for open data is that a new strategy is needed concerning this product due to the falling turnover and the falling number of customers. The result of this is that PostNL is open to innovation for this product, which means that they are open to ideas such as open data.

Despite the fact that this study will not elaborate on the possible creation of a central database with regard to postal code and address data, it is nevertheless useful to further specify and explain the BAG data of the government. By providing insight into the choices made by the Dutch government to open up BAG data (which is almost the same as PostNL's PCT data), well-founded choices can be made and explanations can be given about the value of opening up data.

5.1.3 BAG data and PCT data

The case discussed in this chapter focuses on the data within the PostNL DS Postcodetabel. The data contained herein corresponds in certain parts to the BAG data made open by the Dutch government. The data contains information about all addresses and buildings throughout the Netherlands. Municipalities in the Netherlands are obliged (but they do not always do this) to pass on new addresses to the Land Registry. The Land Registry in the Netherlands is a government service that maintains data about land, buildings and addresses and other geographical information for the Netherlands. In addition, municipalities must also pass on the new addresses to PostNL. PostNL then links a postal code to this new address to ensure that mail can be delivered to this address. This combination of data, postal code, street name and house number, ends up with the Land Registry which subsequently places it in the BAG and makes this dataset public. This information is contained within both the BAG and the PCT. Within these datasets, much more information is also available about addresses and buildings such as: unique identification number of a building, the year of construction, accommodation object (the case when, for example, two addresses 1A and

1B fall under one building), user area, purpose and location on the map by means of XY coordinates. A function for which users of the PCT use the data is given in the previous paragraph. There are many other examples that users use BAG data and PCT data for. This includes (based on observations within PostNL see appendix B DS and "BAG bronhouders - Kadaster zakelijk" (2019)):

- Energy companies: use of data from the BAG for asset management

- Housing associations: use of data from the BAG for property management and policy development

- Banks, mail order companies, utilities and other organizations with many customers: use of data from the BAG to improve the quality of customer files

- Insurance companies: use of data from the BAG for risk assessment and claim settlement to real estate objects

- Suppliers of geo-information and address files: enrich data from the BAG by combining it with other information

So there are many similarities between the type of data that can be found within the BAG database and within the PCT database. The problem lies in the fact that this data does not exactly match and two national "truths" therefore arise concerning postal code and address data in the Netherlands. The reason that these databases do not exactly match has several causes. PostNL, for example, adjusts the data when they themselves find that this is incorrect. There are postmen throughout the country who can check the data in real time. The Dutch government does not have this capacity to keep the BAG data up to date. In addition, municipalities pass on their address changes to the Land Registry Office (because this is required by law) but not always to PostNL. These are two examples that may cause differences between the BAG data and PCT data. A third difference is that PostNL adds to the PCT whether it is possible to deliver mail (this is, after all, the core task of the company), which is not included in the BAG data.

5.1.4 Meeting the criteria

The chosen case study has been extensively described in the previous sub sections. Before this, the selection criteria have been set to choose this case (5.1.1). Table 5.1 shows an overview of why the chosen case meets the set criteria. Now that the context of the case study is clear and it has been stated that the chosen case meets the set criteria, the open data implementation method can be applied to this case. This is done in the next section.

Criteria description	Case study meet criteria?
1. Enough data available	Many different data sources have been used as input within the entire current PCT in recent years. Many different data products and services were created on this. The researcher has access to all this data and details of all data within the PCT. This criterion is met here.
2. Company already has experience with using open data or even already opening up data	One of the input sources of the PCT is the BAG data. BAG data is open data as described in the previous sub section. The awareness among PostNL DS employees is also that an open data variant of the PCT is available, which means that an open data strategy is being looked at with great interest.
3. Commercial company in the Netherlands	PostNL used to be a government owned company. Nowadays this is a commercial company. The specific department where the case is carried out is Data Solutions, which used to be a commercial company under the name of Cendris. Nowadays this company is called PostNL Data Solutions and is a 100 percent subsidiary of PostNL.
4. Multi actor environment	The PCT is currently the product that earns the most money and the product that is purchased by most customers. As a result, many employees within the department have to deal with this product and also a vision or an idea about a possible new strategy for this product. Because the government also has an important role within this process with BAG data, all three of the aforementioned stakeholders from chapter one (open data ecosystem) appear in this case: the government, businesses and citizens.
5. Benevolence for a changing strategy	The turnover and the number of customers on the PCT has been falling drastically in recent years. The demand from the department is therefore to apply a different strategy to this PCT portfolio. An open data strategy can be a solution for this.

Table 5.1: An overview of all the established criteria for selecting a case study (first column) and if the chosen case study meet this criteria (second column)

5.2 APPLICATION OF THE ODIM ON THE CASE OF THE POSTCODETABEL

From this point on in chapter 5 the open data implementation method (ODIM), the artefact of this study, is applied to the chosen case of the Post-codetabel at the commercial company PostNL Data solutions. This methodology was designed in the previous chapter 4. For the structure of this study, all the steps that are followed are once again schematically represented in figure 5.2. In this figure, the steps are also made more specific for the chosen case. All details for the elaboration per phase follow in the following sections.

5.3 APPLYING ODIM ON PCT – PHASE I EXPLORATION

In the first phase of the open data implementation method, it is important to map the complete current system. As explained in the ODIM in chapter 4 there are several methods for doing this. The steps that are taken in this phase are involving stakeholders, establishing requirements, making process models and analyzing the data in the current system (4.5.3). All these steps are explained in the following sub sections. The step in which stakeholders are involved is not further explained in this case in this chapter. The interviews with the people from the PostNL DS department and the observations made by the researcher can be found in appendix B. The results of this were used in the preparation of the analyzes and a detailed elaboration of, among other things, the process models of the PCT.
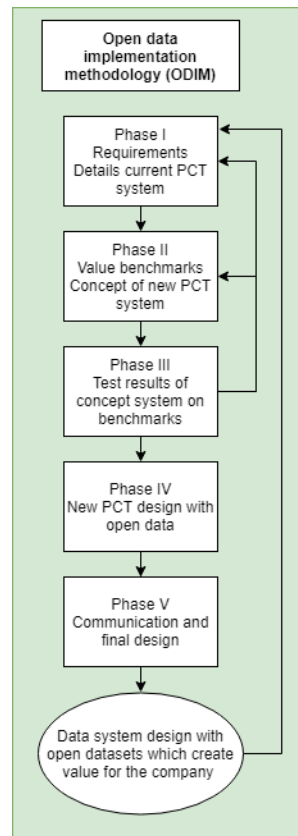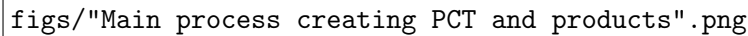
Figure 5.2: Applying ODIM on the case of the Postcodetabel

5.3.1 Process models of the PCT

The PCT is composed of many components and many processes that follow each other or run at the same time. In order to be able to make statements about a new design in which open data plays a role, it is important to first map all current processes in the PCT in detail. This was done by talking to a lot of people within the PostNL department (appendix B) and by researching, reading and trying out a lot. Ultimately, the PCT can be summarized in three main processes: build up PCT, create products and check mutations. These three main processes are explained in detail in the following section. For the overview, this sub section first explains the process of the PCT at the highest abstraction level. This is also done on the basis of a BPMN model. This model is shown schematically in figure 5.3

This model does not yet show the data that serves as input for the PCT but only the processes. The input data follows in the BPMN models in the next sections. The datasets that end up with the users are the products created in the second main process (create products). These are combinations of datasets from the PCT. The PCT therefore consists of thousands of records in an excel file with all kinds of information about an address. Users can then choose the product consisting of datasets that they want as required. This does not have to be all data that is within the PCT. Over the years, this has resulted in dozens of different types of products, each specified according to customer requirements. The PCT is updated every Monday morning so

figs/"Main process creating PCT and products".png

Figure 5.3: BPMN model of the main process of building the PCT and creating data products (CONFIDENTIAL)

that users have the option of accessing the most recent data every week. See appendix C for a more detailed explanation of all the products, different data sets and analyzes that occur within the complete PCT system. The three main processes are described in more detail in the following sections.

SUB PROCESS: BUILD UP PCT   This sub process is visually represented in figure 5.4. The PCT is based on three different data sources. This is first of all the data from the Basis Registratie Personen en Percelen (BRPP data). This is internal PostNL data that is obtained, among other things, during the sorting, coding and delivery of the parcels and letters in the Netherlands. This data is also supplemented by external sources and creates a dataset that contains all addresses in the Netherlands according to "the PostNL truth". This data is placed next to the address data of the Dutch government in order to create a uniform address file of the Netherlands. This data is also used in the PostNL sorting machines to ensure that all mail arrives in the Netherlands properly. This address data is supplemented by a third source, namely data from the Central Bureau of Statistics. This fills in address specific data such as the average income or the average age of a certain neighborhood in the Netherlands.

All this data is combined and then structured into two different data tables: the sequence table and the housenumber table. Both tables consist of the

same data, but the user of the final product that follows from one of the two tables can extract less detailed information from the sequence table than from the housenumber table. For example, the sequence table shows the data of a row of houses in the sequence from housenumber 2 until housenumber 100. The housenumber table, on the other hand, specifically displays the data per housenumber. This last table is therefore much more accurate and more useable for the user. From these two tables a number of data products are made that are ultimately delivered to the end users. These processes are explained in the following paragraph.

figs/"Sub process build up PCT database".png
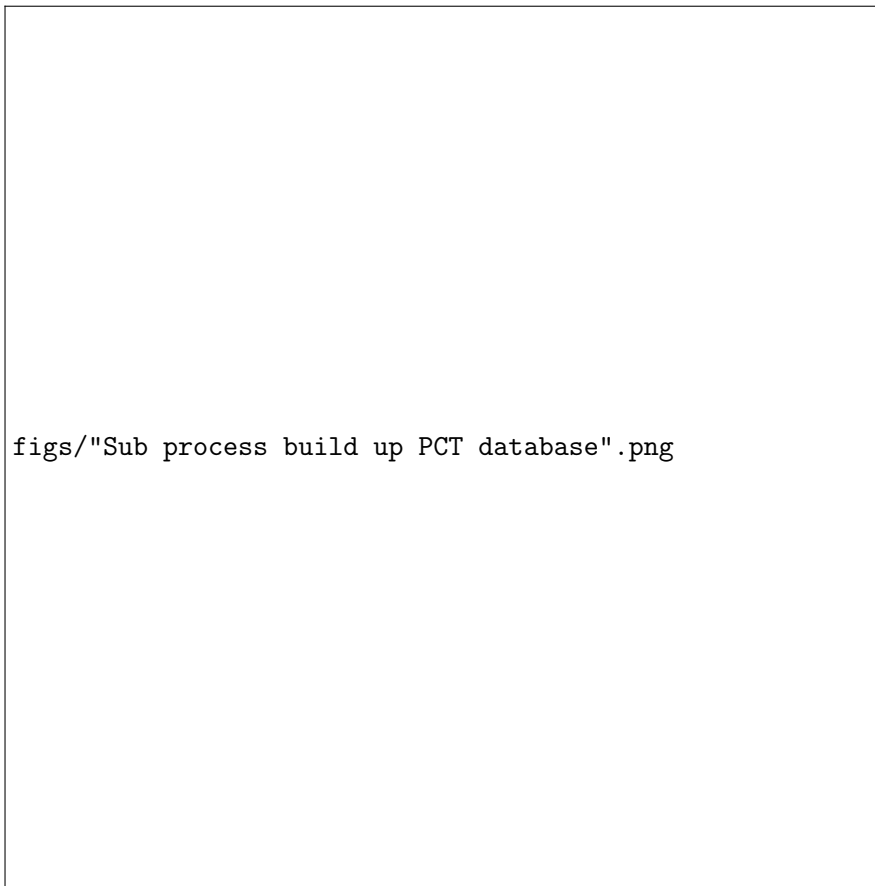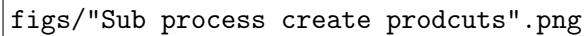
Figure 5.4: BPMN model of the sub process building up the PCT (CONFIDENTIAL)

SUB PROCESS: CREATE PRODUCTS Now that the PCT has been created in the first sub-process, it is up to the end user to indicate what he wants to use the data for. The product is put together based on the wishes of the user. One of the most important elements that distinguishes is the period in which the user receives the updated data. These are all separate data sets that must be made. The customer can receive these updates on a weekly, monthly or annual basis. This section only shows the processes of making the products (figure 5.5). A list of all technically different products can be found in appendix C.

figs/"Sub process create prodcuts".png

Figure 5.5: BPMN model of the sub process creating the products out of the PCT (CONFIDENTIAL)

SUB PROCESS: CHECK MUTATIONS   To keep the PCT up to date, the mutations are checked every week. In this way it can also be guaranteed that the users have the most recent address data. The third main process, checking these mutations, is represented in a BPMN model in figure 5.6.

5.3.2   Requirements of open data in the new system

There are many techniques available in the literature to achieve requirements of the new system. In this case the focus is on value creation using open data in the newly developed system. The change that therefore occurs in the system will only have to do with this open data source or multiple sources, so that not much will change about the system. For this reason, it has been decided in this case to prepare the requirements that this open data must meet in the new system. This also is one of the ways suggested by Dym et al. (2008) that asks the question: "What features or behaviors would you like the ... to have?". This question is asked with regard to open data in this new system to be designed. The question is answered in this case on the basis of extended literature research. This is a good method in this case because much has been written about the requirements of open data. Many of the requirements also follow from the general definition of open data that can be

figs/"Sub process check mutations".png
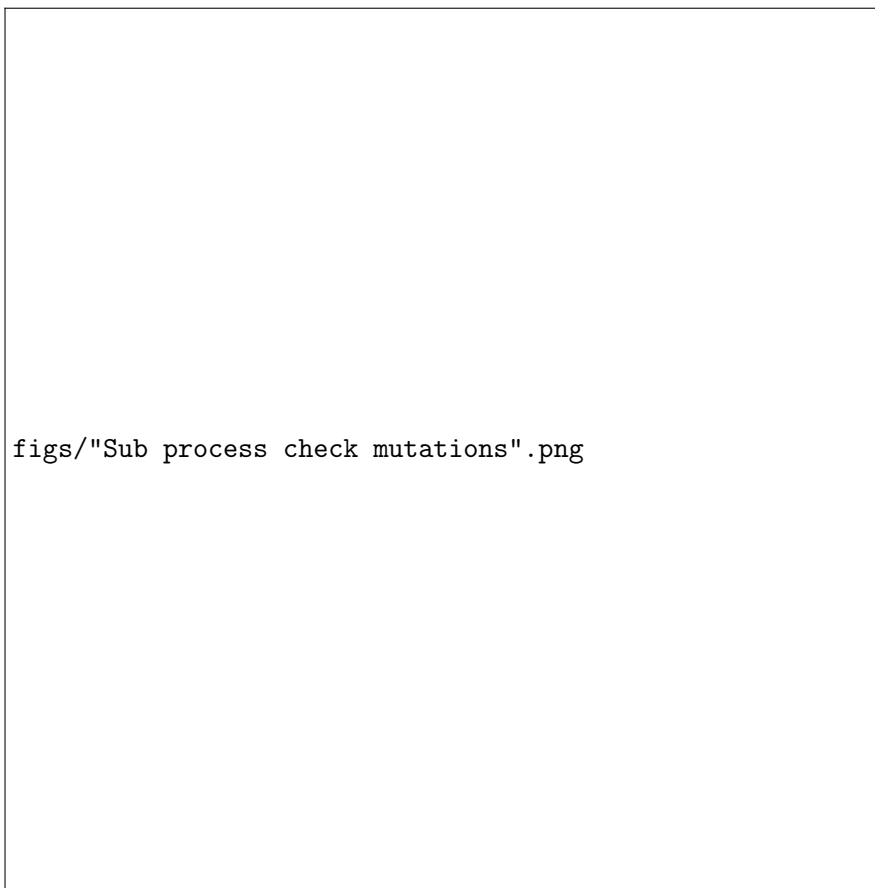
Figure 5.6: BPMN model of the sub process checking the mutations of the PCT (CONFIDENTIAL)

found in sub section 1.4.1. The complete list of requirements that the open data must meet is shown in table 5.2. This list of requirements is mainly based on research by Zuiderwijk (2015), which has researched on detailed information on what open data in open data infrastructures must meet.

Users are mentioned several times in the table in which the requirements are written, but also in other parts of this chapter. It is therefore useful to first define the user of this data. This definition of the user applies to all times when the user is mentioned in this chapter.

When speaking about the user of the open data within the system, all types of end users are meant. This can be government organizations, companies but also individuals. What they then do with the open data is independent of the requirements for open data in the new system. Users can use it to build (paid) products but they can also spread it further as open data.

5.3.3 Analysis of situation of current system

In the ODIM it is discussed in sub section 4.5.3 that there are four different methods to identify which datasets are suitable to be made public within the complete current system. Within this case about the PCT, it was decided

Requirements of the open data in the new system	Description when the requirement do not speak for itself
The data infrastructure should be a single place for datasets from a variety of other infrastructures/places	The open data user can go to a single place to receive the data which they want to use. This will create easy access and potential improved usability
The open data should allow the users to integrate the data with other datasets retrieved from other sources	
Use easy to learn terms to describe the data in the dataset	Improves the way of understanding the dataset for users
The data must be easy to find by using specific keywords	
The dataset should facilitate filtering, sorting and ordering the search results	It must be easy for users to identify what they are looking for in the dataset
The dataset must be delivered in one easy to understand interface, this interface can not change over the years	It must be possible for users to implement the open dataset in their own systems. Changing your own system to make it possible to integrate the open data can be a hard task. Therefore it is necessary that this interface won't change. A risk of changing the interface is losing many users.
The dataset should provide a context in which the dataset has been created	This creates transparency for the users and possibly improves the amount of users
The dataset should provide examples for which it can be used	
The conditions for publishing the data should be clear to the users	
It must be easy to make changes in the dataset	This creates real time, up to date, data which will have effect on the amount of users
It must be easy to add new attributes to the dataset	
The dataset should provide interaction between the providers and the users	This interaction creates a dataset exactly to what the users want
The dataset should assist in viewing who has viewed the data and for what reason	
The system should provide tools for interactive communication between providers and users	
The system should provide tools to make interaction possible between users	
The system should provide insights in the real time quality of the data	
The system should provide information of the context for what the users used the data	

Table 5.2: Open data requirements for the new system. Based on own assumptions and previous work of Zuiderwijk (2015)

to apply two of these methods, namely: choosing a dataset based on the costs and by observing peers. A detailed explanation of this, a complete analysis in which the costs and revenues within the current PCT system have been mapped in detail, can be found in appendix C. The most important conclusion from this is that on all products that are formed within the PCT (see figure 5.5 for details of this process) there is a loss of revenue and a decrease in the total number of customers. In the products that are still being purchased to a large extent and where there is a lot of revenue, customers are often concerned with the datasets of address, house number and postal code. The conclusion is that this is seen as the most important data within the entire PCT system. This is partly also the data that is already made public in the BAG by the Dutch government.

5.3.4 Conclusion of the first phase

All steps carried out in this phase contribute to the development of the new system in the next phase. In order to show the changes in relation to the current system in this next phase, a figure has been composed of the current system of the PCT which clearly states what are currently closed data sources and what open data sources are being used when the PCT is compiled. This figure shows that the PCT is delivered as a batch product. PostNL DS also has other systems and resulting products in the form of APIs (Application Programming Interface). For the sake of completeness, and to demonstrate how the data flows all run in the current system, these are included in this figure. It is furthermore clear that products, both the PCT and other APIs,

are put together from open data sources and from closed data sources and that subsequently all these products are sold as closed data products. End users have to pay for it. End users also cannot make changes to the data themselves because there is no feedback loop in the system. This overview of the current system design of the Postcodetabel can be seen in figure 5.7.
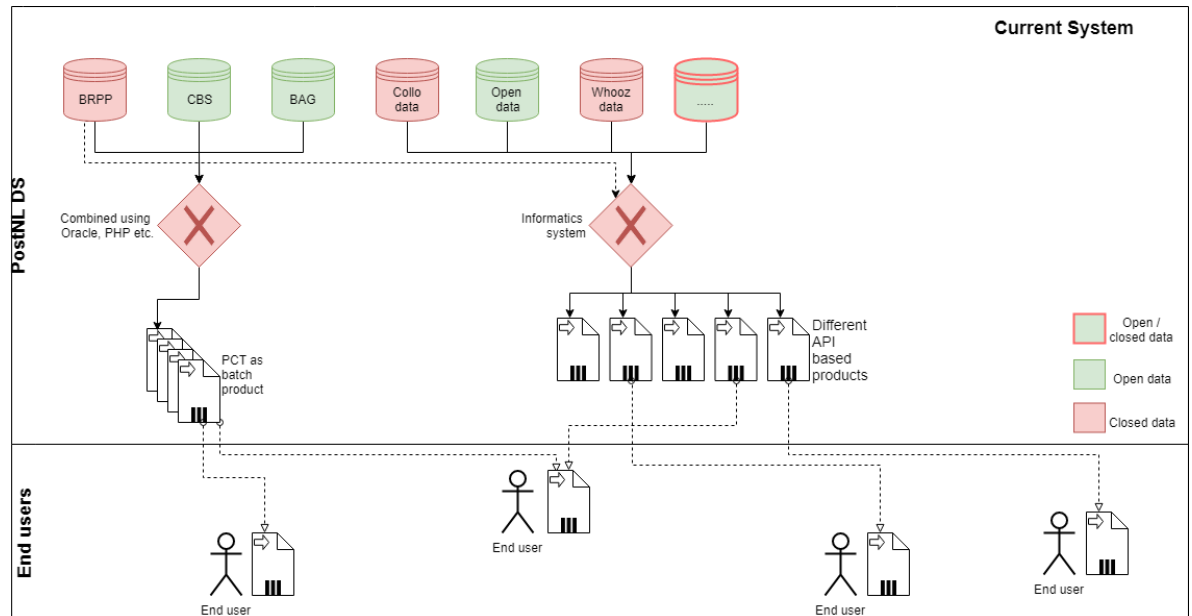
Figure 5.7: Design of the current system of the Postcodetabel (a batch product) and other products (API's). Overview of alle data sources (closed and open) which are used to compile all the different products at PostNL Datasolutions

5.4 APPLYING ODIM ON PCT – PHASE II CONCEPTUAL

In the previous phase it became clear how the entire current system around the PCT works. Problems within this system have been demonstrated in the BPMN models and an overview of the current open and closed data sets used to compile the PCT has been showed. In this phase, a proposal is made for a new system and benchmarks are drawn up to test this new system for the value of the open data. These two steps are further explained in the following two sub sections.

5.4.1 Benchmarks to measure the value of open data in the new PCT system

An extensive explanation has already been given in the ODIM of how benchmarks have been drawn up in a systematic way to be able to measure the value of open data in new systems (sub section 4.6.1). These benchmarks are a result of the literature study that was previously done (important factors of open data at commercial companies in chapter 3), combined with the

work of Susha, Zuiderwijk et al. (2015) and Attard et al. (2016) to determine the value of open data in the new system. These benchmarks are: technological, legal, economic, organizational and cultural. See sub-section 4.6.1 for a detailed explanation of this.

In this part of performing the ODIM on the chosen case, it could also have been decided to set up other benchmarks that are, for example, case specific. This is not done because a good analysis has already been performed on the above benchmarks and case specific benchmarks can also create a certain bias when applied. For example, if the company determines this for itself and starts appreciating this for itself, this is kind of subjective. This choice is further discussed in chapter 7.

5.4.2 Concept design of new PCT system

According to the design engineering process method by Dym et al. (2008) this is the phase in which multiple designs are made in order to be able to choose one. This was not done in the elaboration of this case because it did not fit in the context of this research and by making one design of the new system it could be properly validated in the next phase. This choice is further discussed in chapter 7. The new design is based on the conclusions from phase 1 (by analyzing the current process and conducting interviews, a well-founded choice can be made for a new design with open data in it). This design is explained below and is shown schematically in figure 5.8.

There are three important stakeholders in this new design: PostNL DS, the end users and other third parties. One of the most important differences compared to the current system is that all data sources now come together in a central database called the "Nationaal Adressen Bestand" (NAB). All datasets are stored here, part of which is then released as open data and part as closed data within this NAB. In this new design it was therefore decided to open up part of the data and not all data immediately. Following the cost and revenue analysis and observing the BAG data (the open data of the Dutch government), it was decided in this design to have the name of the city, street name and the housenumber as open data. Other data within the NAB, such as delivery point (is a certain location suitable for delivering mail or parcels) and size mailbox, remain closed data. This choice was made because the latter two aforementioned datasets PostNL DS have a unique position and that data is just being acquired. The willingness to open up this data immediately is therefore estimated to be small in advance. End users can then purchase products based on open data and closed data sources from the NAB. However, users can now also choose to download the open data for free and get started with it (for example by building their own API). Third parties have also been given this option now that this data is available. An additional consequence of this new system is that both end users and third parties can submit mutations to the NAB database, something that was previously not possible (see figure 5.6). So there is a feedback cycle in this new system that makes the data more accurate and must make and keep

users satisfied. In the following paragraph, changes in the dataset, a set is presented how this feedback on the data can work.
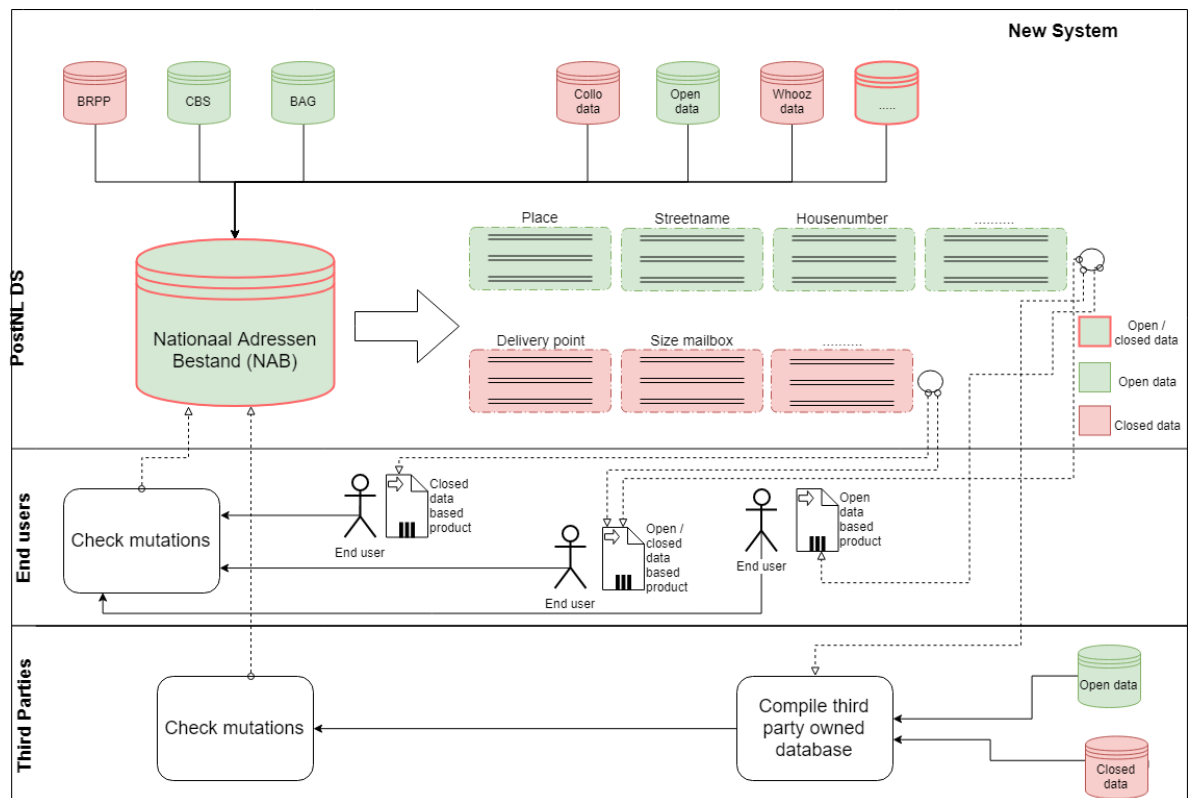
Figure 5.8: Conceptual design of the new system of the Postcodetabel where open data is the key element

CHANGES IN THE DATASET   One of the requirements for open data is that it must be able to be adjusted by users. In the new system setup of the PCT, this must therefore also be made possible. The opened data can be viewed by everyone, which means that it can happen that an individual sees an error in the dataset. An example of this is that the user sees that an incorrect house number addition is stated at his address. Since it is not efficient to have every citizen in the Netherlands adjust the data individually, this will go through an intermediary. With the BAG data from the Dutch government, the municipalities are this intermediary. A condition for making changes to the data is that the person who wants to make the change must provide proof of this. This can be for example in the form of a photo or another document. Municipalities can then check this and choose whether or not to implement this change in the dataset. This system appears to work well, making it a good option to apply for the PCT database when it is being compiled. Another important condition is that users must be able to easily communicate this change. This should, for example, be possible via a website with a form that clearly states how the change must be entered. This is user-friendly for both the importer (the person who communicates the change) and the processor (the relevant municipality where the change is).

5.5 APPLYING ODIM ON PCT – PHASE III TESTING

Now that the conceptual new system for the PCT has been designed in the previous phase, it is time to test this system in this phase. To test the system, it will be investigated what value open data can yield within the five benchmarks established. There are many ways to perform this test, or evaluation / validation, of the new system. Ways to do this include creating simulation models that vary with different variables, conducting a pilot in which part of the data is opened up or testing the system with experts. In the elaboration of this case, it was decided to evaluate the system with experts by organizing an expert panel workshop (EP workshop). Chapter 7 discusses further the choice for the workshop compared to the other ways to test this system. Detailed explanation of how this workshop was set up can be found in appendix D. The most important results are discussed in this section.

5.5.1 Setup Expert Panel Workshop

The workshop is organized at the DataSolutions department within PostNL. A selection has been made of people who know a lot about the PCT system and are open to think about a change within this system. Firstly, an explanation has been given of the concept of open data, what the potential is and examples of companies where open data implementation has been a success. This contributed to the thinking process of the workshop participants in such a way that they were open to reflect on the potential value of open data within the new PCT system. The workshop is divided into two parts: a brainstorm about the potential value of open data and an evaluation on the new PCT system. The brainstorm was intended to make the participants think openly about open data and to be forced to assess benefits and risks. Chapter 1 has shown that companies are often negative about the possible implementation of open data within their company. Questions like: *What will this bring us? Does this not only cost money? And why data is given away for free?* are often questions that arise for employees of commercial companies when open data is being discussed. The introduction about the added value of open data and the organization of a brainstorm session as the first part of this workshop ensured that these questions were answered for the participants. This contributed to the second part of the workshop in which the participants were asked what value of the opened data they saw in the newly designed system. This design (figure 5.8) was presented to the participants and using the framework (figure 5.9) containing the most important benchmarks of open data, the participants were able to estimate the potential value of the system. Questions that were central during the workshop were:

1. What do you think are the *benefits* of open data implementation at commercial companies if you have to put it in one of the five dimensions?

2. What do you think are the *risks* of open data implementation at commercial companies if you have to put it in one of the five dimensions?

3. Do you agree with the data currently being opened up in the new PCT system? If this is not the case, what data would you open up?

4. What value, within one of the five dimensions, does open data yield within the new system of the Postcodetabel?

5. What are the most important requirements for you with this new system of how the PCT is set up?
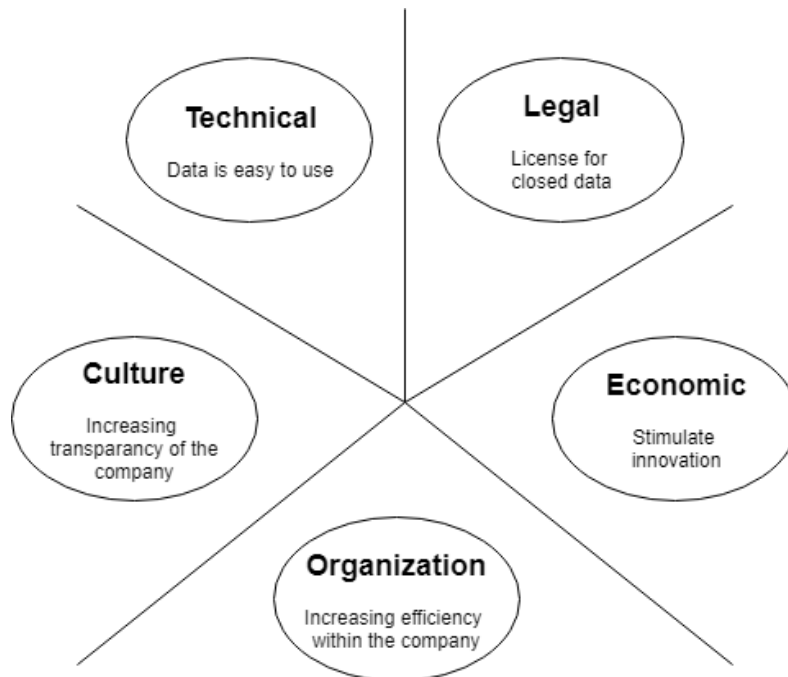
Figure 5.9: Framework of benchmarks for value creation in open data systems

The workshop was held with a total of seven different PostNL DS employees. For both the brainstorming and the evaluation on the new system, the group is split into two groups. A more detailed explanation of the setup of the workshop, the slides that have been used and the timetable can be found in appendix D. The results of the brainstorming session are not relevant in this section of the case and for this chapter. The following sub section discusses the results of the evaluation on the new system and the value of open data.

5.5.2 Results value creation open data in new PCT design

The results of both groups that participated in the workshop are discussed based on the questions asked. This sub section is also classified in this way to provide structure and to clearly show where the value of open data in the new system design comes from.

What are the most important requirements for this new system?

- The self-service portal where users can submit changes for changes in the data must work well. This feature of the new system is seen as a key element which should therefore function properly for the overall value of the system.

- Following the court case (section 5.1.2) with the Dutch government, the Postcode covenant has been agreed. Before the choice is made to make the data publicly available, it must be carefully considered whether this is legally permitted. Perhaps this conflicts with the BAG.

- Another important requirement for this new system is the proper design of the security system. Making open data available can also cause you to encourage more parties to do malicious things and to hack the system. This must therefore be properly sorted out before the data is made public.

- Another requirement of the new system is that there must be a kind of teaser in the opened data to lure users to other (closed) data within PostNL DS. End users get a lot of insight into what is possible with the data, but may not be able to do this themselves. A teaser then serves to attract more users to the open / closed based data product or the APIs that are built on these open data sources.

Which datasets should you open and why?

- This proposal of the new system states that only the combination of city , streetname and housenumber will be made public. Both groups indicated here that when datasets are made open, all data that also occurs in the BAG should also be opened up. End users therefore see the difference in quality and are more likely to choose for the PostNL DS data. This provides a lot of value for the new system and for PostNL.

- One group from the workshop also added that all data sources that we use ourselves must be made public in the NAB. These are like open data sources and by adding a teaser from PostNL DS, users see the added value of this data compared to the already open available data sources from, for example, the CBS.

In which area does open data create the most value for you in this system?

- Both groups that participated in the workshop recognize the great economic potential of open data in this new system (economic dimension). This is mainly due to the possibilities for innovation within existing products and services. By giving users open insight into the available data, it can lead to these users realizing that they need more than just this open data. Or that they would like to see this open dataset enriched with one of the (still) closed datasets. The new APIs and other data services that can be built on this new system offer enormous potential.

- The second point is the transparency that is created for the company as an important factor that adds value to this new system (cultural dimension). By being more transparent in the market, the PostNL brand will only become stronger.

- As a third and final point, both groups recognize that open data in this system provides a lot of value creation in the organizational area (organizational dimension). Current IT systems will be viewed more

critically because the data that is made public must meet certain requirements. This is not the case when the data remains closed. The data will therefore become more accurate and the organization of the department will improve.

5.5.3 Comparison of results between groups

The results of the expert panel workshop are shown in the previous sub section. In this sub section the differences and the similarities in the results between the two groups are already partly highlighted. For the structure of this study and because it is more relevant to state separately, these differences and similarities are shown in the table 5.3. The results are based on the discussions in the workshop. The survey also indicates differences between individuals. These results can be found in appendix E.

Differences	Similarities
The groups did not entirely agree on the value that would be achieved within the cultural dimension. The brand name of the company would go up, but this could also have a downside. If the data is not accurate enough, it appears that many users are driving away. The other group did not entirely agree with this point and thought it was only positive branding for the company.	Both groups suggested opening up all data that is also shown in the BAG database.So not only the combination of postal code, address and housenumber.
In addition to opening up all the data in the BAG database, one group also said that other datasets should be opened up. They also proposed opening up the data collected from other open data sources, but then under the PostNL brand name. The other group did not agree with this because it is not their own data and little is being changed about it.	In both groups, the recognition of value within the economic domain was most present. In particular, the possibilities for making new products played an important role in this. According to both groups, open data would lead to innovation within the company.
The groups did not entirely agree on the transparency that the open data would be achieved. Both acknowledged that this would go up, but they did not agree that this was only positive for the company.	The groups also agreed on the value potential within the organizational dimension. Open data would contribute to an increased quality of IT systems.
A group suggested having a freemium subscription to the open data. That only a part is opened and if the user wants to see more, they have to pay for this. The other group did not think this was the purpose of open data, it stated that everything must be open and that users can be tied to the company in a different way.	Open data works great as a teaser to attract potential customers to other products within the company.

Table 5.3: Differences and similarities between the results of the two groups in the expert panel workshop

5.5.4 What if the PCT data is not opened up?

Another point that was presented to the workshop participants is the impact of the PCT when the data is not opened up. The participants had not thought about this which made it interesting to discuss this scenario. If the data is not opened up, the current operating systems will continue to be operated as they were. The PCT would then remain in its current state. It has emerged from appendix C that the turnover and the total number of customers within the PCT has been falling considerably in recent years. The participants acknowledged that change is needed within the current design of the PCT system. If the BAG data also continues to develop, this may eventually become a major 'competitor' of the PCT and the PCT will lose market value even more. Failure in opening up the data can therefore entail major risks. The participants in the workshop, however, were not yet fully convinced by moving straight to an implementation plan. Other options for redesigning the PCT system should also be considered in order to make informed decisions about open data within the PCT system.

The above sub sections provides an insight into the most important results of the workshop on the new PCT system. Following this workshop, a survey was also sent to all participants. The results from this reinforce the results of the workshop and can be found in appendix E. The overall conclusion of all workshop participants was that a lot of added value can be created with open data in this new design. In the ODIM there is also a reflection phase on this point in which one can go back to phase one to map the current system again. This was not necessary in this case, and therefore will be moved on to phase four.

5.6 APPLYING ODIM ON PCT – PHASE IV EVALUATION

In the elaboration of this case, phase IV of the ODIM was chosen to provide a detailed elaboration of the validated system following from phase III. This new system is shown schematically in figure 5.10. The most important change resulting from the workshop has been incorporated in this. If the open data system is implemented, all sources that are now also open data will also be published as open data. This will be done using a teaser to attract more users to the PostNL DS APIs or other data services. By showing users that the PCT data is better than the BAG data, much value will be created with open data for PostNL DS. The system shown in figure 5.10 is for this case the final design that has been validated through the workshop in the previous phase.
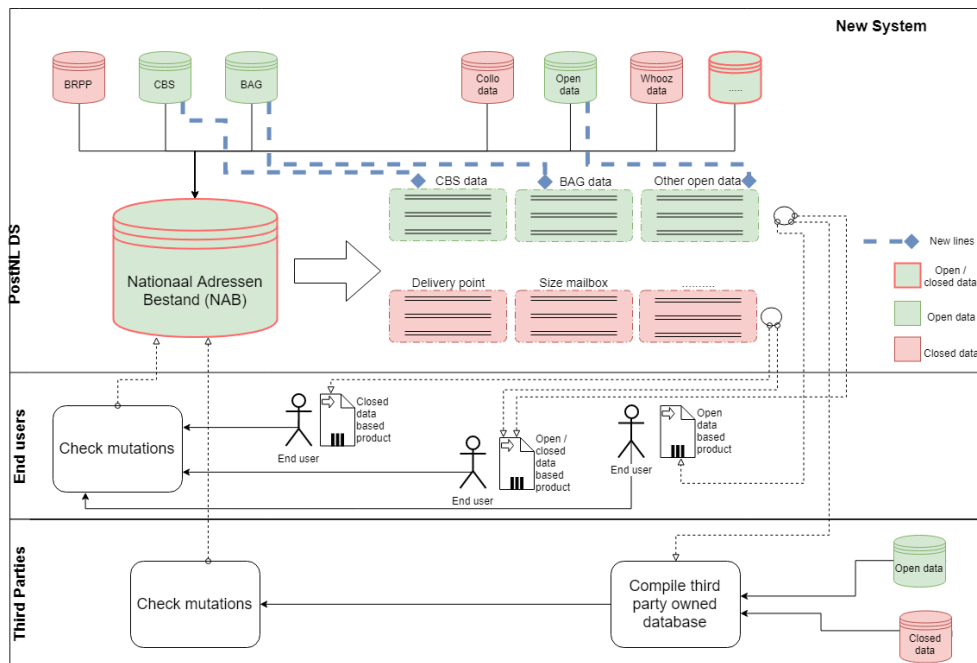

Figure 5.10: Proposed new final design of the Postcodetabel

5.6.1 Meeting the requirements

In the first phase of the open data implementation methodology, requirements have been drawn up with which the open data in the new system must comply. These are requirements that can also be seen as Key Performance Indicators to determine whether the new system works well. In this fourth phase of the ODIM it is useful to reflect on these requirements. The requirements are again shown in table 5.4 with now also indicated whether the new system complies with this. In this study, the new system remains a potential change and the new system has not yet been tested in reality. The statements in the second column of the table, whether the system meets the requirements, are therefore based on the results from the expert panel workshop and the visual representations of the new system (figure 5.10 and appendix D).

Requirements of the open data in the new system	Requirement met in new system of the PCT?	Explanation
The data infrastructure should be a single place for datasets from a variety of other infrastructures/places	Yes	All data sources, both internal and external, are stored in the new NAB file
The open data should allow the users to integrate the data with other datasets retrieved from other sources	Yes	In the new system it is possible for users to make changes to the data
Use easy to learn terms to describe the data in the dataset	Yes	Terms have been used that also occur in the BAG and others that must be clear to users
The data must be easy to find by using specific keywords	No	Must be included in implementation plan of new system
The dataset should facilitate filtering, sorting and ordering the search results	No	Must be included in implementation plan of new system
The dataset must be delivered in one easy to understand interface, this interface can not change over the years	Yes	These conditions have been drawn up within the preparation of the NAB
The dataset should provide a context in which the dataset has been created	Yes	The new system makes clear where the data comes from and what has been done with it
The dataset should provide examples for which it can be used	Yes	This is also included in the BAG description of the data and is also included in the new PCT
The conditions for publishing the data should be clear to the users	No	Must be included in implementation plan of new system
It must be easy to make changes in the dataset	Yes	There is a direct link for users to the NAB
It must be easy to add new attributes to the dataset	Yes	Extra columns have been added to the dataset within the NAB, so that new attributes must be easy to add
The dataset should provide interaction between the providers and the users	Yes	Both users and other third parties have access to the data and can propose adjustments
The dataset should assist in viewing who has viewed the data and for what reason	No	Not yet included in the current design
The system should provide tools for interactive communication between providers and users	Yes	Ideas for a platform on which users can submit changes have been prepared
The system should provide tools to make interaction possible between users	No	Now only the NAB for interaction between provider and user. Further elaboration is required
The system should provide insights in the real time quality of the data	Yes	Important condition before the data is opened in the NAB. Better design of IT systems needed for this
The system should provide information of the context for what the users used the data	No	Only from the provider for which the data can be used. Not yet from the users.

Table 5.4: Meeting the open data requirements in the new PCT system

5.7 APPLYING ODIM ON PCT – PHASE V COMMUNICATION

The design must be communicated in this final phase of the ODIM. Within this case, the final design has already been presented in the previous phase in figure 5.10, therefore it will not be discussed again here. In the communication phase of this new system there is actually an implementation plan for this system, something that was often referred to as a result of the expert panel workshop. This should contain specifications about the technical details of the datasets that are made public, instructions for use for end users of the datasets, an extensive plan on what kind of platform the data will be made available and how the feedback mechanism on the open data will work. Given the scope of this research, namely the focus on the potential value of open data within systems, this is not further addressed. When a company actually chooses to implement the new system, the afore-

mentioned aspects must be further elaborated. The ODIM drafted in the previous chapter does discuss how a company can choose a business model that is suitable for the data that has been opened up based on the value determination of the new system. Within this case, based on the work of Zeleti and Ojo (2018) and insights obtained from the interviews and the workshop, a freemium model is a suitable business model. This is further explained in appendix F.

5.8 SUMMARY CHAPTER 5

In this chapter an answer is formulated to the fourth sub question: *What value will opened up data create when applying the open data implementation methodology on a real case?* The complete open data implementation method has been implemented in a real life case in this chapter. After drawing up various selection criteria, it was decided in this study to carry out a case at the commercial company PostNL DataSolutions. The case relates to postal code and address data in the Netherlands. With this data a product is built, the PCT, which is sold to customers. As the value of this product has fallen enormously in recent years, this case is suitable for determining whether an open data strategy provides added value for the company. After mapping the current system in detail, a proposal was made for a new system design in which open data has a central role. This design has been validated through an expert panel workshop with PostNL DataSolutions employees. They were asked what they thought of this new system and what they thought of the value of open data measured on five benchmarks: organizational, cultural, economical, legal and technical. The most value for open data within this new system is, and therby forms an answer the to central sub question in this chapter, in the organizational, economical and cultural dimensions. This demonstrates that by applying the open data implementation methodology insight can be gained into which datasets are applicable to be made public within this case. The value of the open data can be achieved within different domains.

6 | EVALUATION OF THE OPEN DATA IMPLEMENTATION METHODOLOGY

This chapter describes the final phase in design science research, the research approach that is followed in this research. In this chapter the developed artefact, the open data implementation methodology (ODIM), is evaluated. This chapter examines whether the methodology developed in chapter 4 and applied to a single case study in chapter 5 can also be applied more broadly. With this, this chapter answers the sub question:

What are the limitations of the open data implementation methodology for application at commercial companies?

According to DSRM, the evaluation phase focuses on observing and measuring how well the designed artefact supports a solution for the described problem. The described problem in this study addresses the fact that commercial companies are not yet using the full demonstrated potential of open data. The ODIM is developed for this and this chapter discusses whether the method can be used at other commercial companies or parts of the method. Within the DSRM literature several evaluation methods are proposed including observation, analytics, experiments, testing, descriptive analysis, interviews and focus groups (Hevner & Chatterjee, 2010). The latter two methods have been chosen for this study, the focus group being held in the form of an expert panel workshop. This chapter is further subdivided into three different sections. Firstly, the approach to the evaluation is discussed, followed by the possibilities within ODIM that can be applied at other companies. Finally, the limitations of the ODIM are described and a proposal is made how the ODIM can be better implemented in the next case by applying a small adjustment.

6.1 APPROACH OF THE EVALUATION

As indicated in the introduction to this chapter, the open data implementation methodology is evaluated. To do this, it is first useful to repeat the purpose of the methodology. Because it has been shown that commercial companies currently do little with open data, and according to research there is a lot of potential for this, the ODIM aims to make it clear to companies which datasets can be made open within their business systems. In addition, the ODIM also serves to make it clear to these companies what value this open data can create. To test this, the ODIM was applied to a case at the largest postal delivery company in the Netherlands, PostNL. Datasets to be made public have been identified and their potential value has been tested

in an expert panel workshop. The newly developed system was presented to the participants in this workshop.

6.1.1 Shortcomings within this evaluation

As indicated in the introduction of this section, the expert panel workshop was given to the people who work with the PCT system. They were asked to evaluate the value of the open data within the new system. At the same time, this workshop was used to evaluate the open data implementation method. Combining the workshop for both elements in the research is not entirely valid, but it does fit within the scope and time frame in which this research was conducted. A more complete method of validating the open data implementation method is to first validate it externally with, for example, open data experts. They can propose adjustments to the method before it is applied to a case. In addition, this also contributes to the objectivity of the method's assessment. Objectivity will increase by not using the same people for the value assessment of the open data and the evaluation of the method. This point in the design science research approach in this research could be improved. This also helps to make it easier to use the method on another case.This way of researching, reflecting and zooming out from a specific case to a larger whole is possible, but it does have disadvantages as mentioned before. The researcher of this study is aware of this fact.

However, there are elements in the evaluation of the method that are useful. As there were several points to reflect on the method from both the expert panel workshop and from the interviews with the people at PostNL DS, this is discussed in the following sections. These points that are made here are therefore based on this single case study from which a lot can be learned for a subsequent study. These are positive points and improvement options that can already be determined after a single case.

6.2 POSSIBILITIES OF THE ODIM

Within the ODIM a distinction has been made between five different phases based on the engineering design process methodology by Dym et al. (2008). The results of the workshop, own observations and the interviews show that a number of steps within these phases were useful. The capabilities of the ODIM listed below have been identified as valuable for commercial companies to apply when they want to implement an open data strategy or explore the options of open data.

- The first phase within the ODIM is about the complete mapping of the entire current situation. The steps that have been taken have proved to be of great value throughout this method. The results of the workshop showed that the design was mapped out in great detail in the first phase. By subsequently asking questions about all types of data

sets that occur in this system, a good foundation can be laid for the design of the new system with open data. Because the workshop did not provide feedback on the fact that the current system was modeled incorrectly and that the new system was therefore considered reliable, the power of this step demonstrated within the open data implementation methodology. Mapping the current data system at a commercial company by means of process models is therefore valuable for open data implementation.

- Secondly, the requirements of open data in the new system to be developed also came back in the first phase of the ODIM. These requirements were mainly related to the requirements that open data must meet in the design of the new system. These requirements helped a lot in compiling this system. By making these requirements well communicated to the participants, it contributed to the positive response to the new system. A cautious conclusion from this is that these requirements are broadly applicable to other cases, possibly supplemented with specific requirements per company.

- The analysis based on the costs, revenues and observing peers (in this case the Dutch government) turned out to be a correct step within the methodology. Because the participants were confronted with the enormous decrease in the revenues of the product (the PCT) and with the fact that the Dutch government provides an almost the same type of product but then for free, the participants started thinking. This also contributed to the fact that the participants were open minded to the new system design in which open data had a central role. This can be explained by the fact that the BAG data acted as a kind of best practice.

- The next point that offers many possibilities to apply the open data implementation method to another commercial company is to present the development of a single design. Where Dym et al. (2008) write a lot about the usefulness of multiple designs, submit them to the client and then make a choice between them, it turned out not to be eccentric added value here. Because the participants had heard of open data but did not yet know the value for companies (appendix E), only one design helped in thinking about this system. The disadvantage could be that the researcher did not make the right choices in the beginning to design the system.

- The benchmarks established to test the value of open data within the new system are considered valuable. Having open data implementation viewed by stakeholders from a technical, economic, cultural, organizational and legal dimension proved to be a good approach. These benchmarks can also be adopted for the next application of the ODIM at another company.

6.3 LIMITATIONS OF THE ODIM

In addition to the fact that many elements that were proposed within the open data implementation method turned out to be valuable for wider application in other companies, a number of limitations have also been discovered. These limitations of the method are again based on the single case study elaborated in the previous chapter and in particular on the results of the workshop.

- First of all, the ODIM proposes two other ways to analyze which datasets may be suitable to open up within the current system of a company (phase one). These are analyzes based on the needs of the users and on the basis of the "ease of the release". An extensive investigation into the wishes of the customers, for which they currently use the data within their own systems for example, would be necessary for this. An analysis of the technical details of the datasets and on what kind of platform they can be made public is also needed to investigate this. These were not included in the case discussed in this study because the other two that were investigated proved to be sufficient.

- The benchmarks set up to test the value of the open data within the newly designed system proved valuable but also have limitations. Because the participants were forced, in kind of a way, to view the system from these five dimensions, there was no further consideration of other possibilities of possible value creation of the opened data. A good addition can be to let companies determine what the benchmarks are to estimate the value of the system. It is important to bear in mind that it should not be at the expense of the previously set requirements for open data from phase one in the ODIM.

- A frequently heard point of feedback from the workshop was about the implementation of the new system. A disadvantage of applying the open data implementation method is that the system remains conceptual. The workshop participants saw a lot of potential in the datasets that could be opened, but this could only be confirmed for them if this actually happened. To be able to reach this step, an implementation plan was often referred to. This is now mentioned within the ODIM in the final phase, but is not further elaborated in the case under consideration given the scope of this study. When the system is approved in phase three and a final design is made in phase four, an implementation plan for phase five can be considered. This can possibly also be presented earlier to the most important stakeholders within the current system to better guarantee success.

- The engineering design process method of Dym et al. (2008) contains a feedback cycle that is also reflected within the ODIM. From the stakeholders and based on own observation, it was concluded that this moment of feedback to earlier phases can be valuable. For example, companies can choose to completely stop using an open data implementation strategy or they can choose for a different dataset to open

up. This feedback cycle has not been worked out in sufficient detail in the case studied because there was no need for this from the problem owner. The assumption is therefore that these feedback cycles are valuable, but this should be confirmed in a subsequent case.

6.4 POSSIBLE ADJUSTMENT OF THE ODIM

In the previous two sections a number of possibilities and limitations are mentioned of the open data implementation method of how it is presented in chapter 4. These are the possibilities and limitations of the method if it is applied within another commercial company. It has been demonstrated that the ODIM serves well to identify which datasets within a current system of a commercial company are suitable to be opened up based on the proposed techniques. This was one of the reasons why the ODIM was designed with which it can be said that based on this case this is successful and is likely to be the case with other cases. The other reason for which the open data implementation method is designed is to identify the value of the opened data. The benchmarks found in the literature proved suitable for assessing this, but these were also a limitation. For this reason, as an adjustment to the ODIM, it is recommended that the creation of the benchmarks also be adapted to the needs of the company where the case is being conducted. This adjustment to the method is shown schematically in figure 6.1. The researcher is aware that there are also other options for adjusting the ODIM, but this is seen in this study as the most important adjustment when the ODIM is applied to another case.

6.5 SUMMARY CHAPTER 6

This chapter reflects on and evaluates the open data implementation method, the artefact of this research from design science research approach. This chapter forms the final phase in the design science approach and provides the answer to the question:

What are the limitations of the open data implementation methodology for application at commercial companies?

The limitations in the method are determined in response to the given workshop that was discussed in the previous chapter and in detail in appendix D. The most important limitation was that the benchmarks for value creation of open data within the newly developed system may be too limited. When the method is applied to another case, it is valuable to derive more benchmarks proposed by the company. The ODIM is seen as a suitable method for demonstrating which datasets are suitable to be opened up and what the new system with open data should look like.
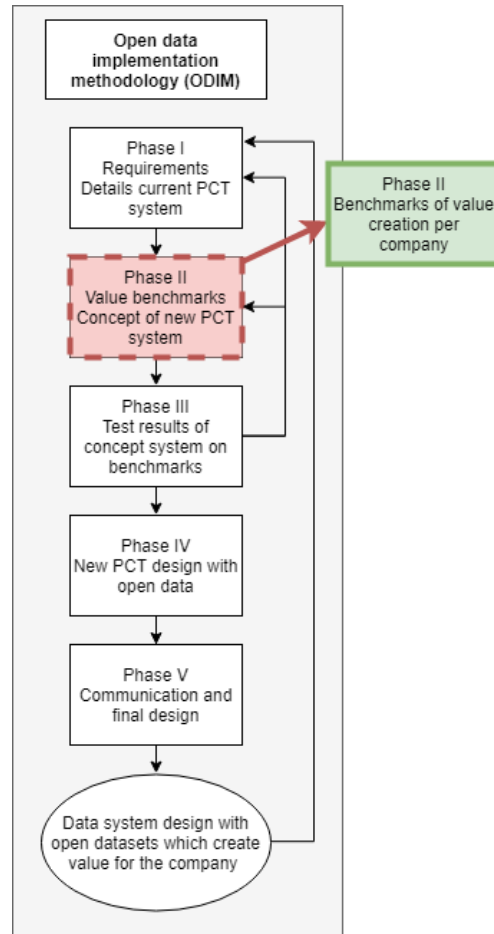
Figure 6.1: Adjustment of phase II of the open data implementation methodology after evaluation

7 | DISCUSSION

The results of this research are discussed and reflected on in this chapter. The aim of this research was to develop a method that can be used by commercial companies to see within their systems which datasets are suitable for being made public and what value this open data can create for a company. The methodology has been named "the Open Data Implementation Methodology" in this research and has been developed as an artefact within the design science research approach. This method was subsequently tested at the Data Solutions department of the commercial company PostNL, in the Netherlands. However, the design of this research and the methodology developed are not perfect. This is partly because the research is based on many assumptions and there are various limitations to the methodology. Regarding validation, there are also a number of limitations. Many of these assumptions are discussed in this chapter.

This chapter is made up of various parts. First of all there is a reflection on the way in which open data was viewed in this study. This is followed by a reflection on the research design used in this research, the design science approach. This is followed by a reflection on the methodology and the assumptions and the choices made herein for the composition of this method. The outcomes and elements of this chapter help to put this research into perspective in order to make the results usable and possibly support future research.

7.1 REFLECTION ON OPENING UP DATA FROM THE PERSPECTIVE OF THE RESEARCHER OF THIS STUDY

First, it is good to consider how the concept of open data was approached in this study. Due to a certain form of subjectivity, the possibilities of open data are viewed from a optimistic angle. On the other hand, this has been supported by many studies from both science and the business environment itself. By starting from such a starting point, the researcher has assumed in a certain way that when data is opened up at companies, these companies can get a lot of value out of this. In the introduction of this study, a number of risks have been mentioned about opening up data. Because of these risks, companies can choose not to open up their closed datasets and to set up their business model how they have done so far. Not all of these risks will be covered in this section, but the one highlighted in the introduction is also good to mention here. The European Union's privacy legislation, in

the form of the GDPR, makes companies (but also other organizations) a lot more careful when it comes to data. Not just about opening up the data, but in general, companies are a lot more vigilant about what happens to their data. This was also found by the researcher when conducting the case study at PostNL. In the many conversations that were held with the employees of this department, the GDPR regulations often came up. The biggest question within the PCT portfolio, and actually within one of the most important data that PostNL owns (address and postal code data), is whether this is not personal data. There appears to be a great deal of uncertainty about whether or not an address is a personal reference. The one argues that this is of course the case because we know who lives there, the other says it is not because it is about a building that is an object and not a person. By investigating whether parts of the PCT data can be opened up, this was a frequently heard concern.

In this study is it tried to cover both discussed points. The BAG data, the data from the Dutch government, was used as a best practice, which led to the people who had to assess the value of the new system looking at this system differently. Risks were estimated to be lower due to trust in the government, which also publishes address and postal code data, and if they can, why can't we? The point at which open data was viewed from a positive perspective was also attempted to cover in this study. By having people think critically about the risks and the possibilities of open data within the new system within the five dimensions set up in the workshop, the participants were not biased. This has helped to investigate the value of open data at commercial companies from an optimistic but critical angle.

7.2 REFLECTION ON THE USED APPROACH

In this research it was decided to use the design science approach. The choices for this can be deduced from the extended literature review in chapter 3. It has been shown here that research into open data has been done before, but certainly not often, using the design science approach. Another reason for using this approach is the artefact that is being developed in this research approach. The identified problem that companies do little with open data while there is great potential for it can be well investigated by offering these companies an possible helping tool in the form of a methodology, the artefact in this study. Using design science research was a suitable method for this, which led to a systematic approach to method development. By having the research follow this approach from the start, the research is being pushed in a certain direction. The question that sometimes came up was whether a different approach might not lead to other, maybe even better results. One of those approaches that sometimes emerged was action science research. Although there are many ways to tackle research into the value of open data, this approach is briefly explained in the next sub section to place this research in that perspective.

7.2.1 Suitability of the design science research approach

Looking at the available literature, there is much to be found about the similarities and differences between design science research and action science research (AR). To indicate in this sub section that it has been useful to use design science research and not action science research (which is mentioned here as an example to indicate that other ways of doing research on open data are also possible), a few of these differences are explained using Iivari and Venable (2009). Design science research is focused on designing innovative solutions in the form of an artefact as opposed to AR. In the conceptual context of this research, this was therefore a suitable research method. Nevertheless, AR can be used well as a supplement to the design science approach. Action science research is "aimed at exploring the reasoning and attitudes that underlie human action, and producing more effective learning" in individuals, organizations, and other social systems" (ARGYRIS & SCHÖN, 1989). Partly for this reason, AR would be a good addition for example to evaluate the ODIM, which has now been done within the design science approach, but by using the same case as in the demonstration phase there is much room for improvement.

7.3 LIMITATIONS OF THE METHODOLOGY

In this study, two important differences have to be distinguished. The structure of the research, and with it each separate chapter, follows the approach of design science research. In addition a method has been developed, the artefact within this approach, which is built on the basis of a standard design engineering process methodology by Dym et al. (2008). The approach in this study has already been discussed in the previous section, this section reflects on the choices made within the developed methodology. These are listed below.

- In the literature review in chapter 3 one of the conclusions is that research has mainly been done into the value of open data in a qualitative way. One of the studies found for this study that looks at open data from a quantitative side is research done by Marin de la Iglesia (2014). By comparing the revenues and costs of open data, quantitative statements could be made about the value of open data within systems of companies. Although this research has once again investigated open data in a qualitative way, the method will nevertheless serve as a tool for companies to be able to use it quantitatively. The method has proved being successful in identifying datasets that have the potential to be opened up. Because this research has been mapped out in detail, it is easier for companies and scientists to obtain quantitative results with this data. This research therefore contributes to the way in which open data research has been conducted.

- The above point also directly partly reflects on the choice within the requirement analysis method in phase two of the ODIM. The decision was made to only look at the functional requirements of open data because this offers a good first step towards an open data system. Non-functional requirements, such as costs, performance, security and privacy, can ensure that the implementation of a new system containing open data is an obstacle for the company. This can also result in the end users of the data not being satisfied. This part is discussed in the proposed methodology in the final phase of the communication of the design. Implementation techniques, business models and other aspects that are required for this must be further elaborated in this phase. These non-functional requirements may therefore have a place within the developed method, but must be further investigated in detail in order to be able to make well-founded statements.

- In order to test the value of open data within the newly developed system, it was decided in this study to use the identified factors from the literature review. These emerged as challenges within open data ecosystems, but could also be converted into dimensions within which value can be achieved (Attard et al., 2016). This turned out to be a good way because participants in the workshop were able to determine well what they considered valuable in terms of open data and where there were risks. However, these benchmarks can be compiled in many different ways. An idea that is also suggested in chapter 6 is to let companies determine this more actively. Every company is unique and will have its own benchmark when developing a new system. Within PostNL this worked well for this research, but for a subsequent research this can be done in consultation with the researcher and the company.

- A frequently heard comment during the execution of this research within PostNL was: "But how do we know for sure what the value is if we do not open up the data?" On the one hand, this was a legitimate comment because this conceptual design does not actually include opening up the data. That is the next phase that would require an implementation plan. However, this method contributes a lot to consciously creating in companies what the value of open data can be. In addition, it also provides a great deal of insight into its own data systems, so that it can be said that the ODIM contributes to the value creation of open data for companies, even without actually opening the data. This confirmation also came from the workshop where the participants were already actively thinking about which parts of the data they would open up and what value this could create.

- Using the methodology of Dym et al. (2008) to design the ODIM is an important direction in this research. Choosing a different method would probably have yielded different results. There is a lot of literature available about designing systems where in this research, in chapter 4, three different methods are explained and compared. Reflecting on this choice, it can be stated that the method of Dym and

Little has been a good choice, in particular because of the iterative nature of this process. Although this has not been discussed extensively in the case study, this feedback cycle is of great value for companies in terms of open data. By evaluating later whether users are actually using the data that has been opened up or evaluating it in a different way, this contributes to the value of the open data for users but also for the company that has opened up the data. By going back to phase one or phase two in the ODIM, and look critically at the current system and see whether another data set can be opened up or if an open data strategy for the company turns out not to be a suitable strategy can be useful.

- Dym et al. (2008) discuss the added value of multiple designs to make a choice between them. Despite this, it has been decided in this study to only work with one design and to have it assessed regarding the value of the open data in a case. This was also reflected in chapter 6 and it turned out to be the right choice in the context of the case in this study. Presenting people a design containing open data, of which they said they had no knowledge about, should be limited in complexity. Presenting only one design contributed to this.

- In the third phase of the methodology, the value of the open data in the new system was tested on the basis of an expert panel workshop. This value could perhaps also have been tested by means of simulation models or, for example, the execution of a pilot. In this case, however, the workshop proved to be of great added value because presenting a real life situation to experts about the system leads to many insights. The execution of a pilot did not fit into the time frame of this study and a simulation model is not immediately suitable for estimating a possibly rather subjective variable as value dimensions. Letting people who understand the current system determine the value of open data was therefore valuable.

7.4 GENERALIZATION OF THE FINDINGS FROM THE CASE STUDY

The results of this research are based on a systematic approach to the development of the method and on testing this method on a single case. Since this case was carried out at a company that trades within the logistics sector, the main question of this study, which will be answered in the next chapter, is also addressed in this way. Concerning the generalizability of this research on a subsequent case, it is therefore best to make a statement regarding commercial companies within this logistics sector. The value of open data within the case lies in particular in the economic and cultural dimension. The value of postal code and address data is large for the logistics sector. The results of this one case, to make the postal code and address data partly open, can therefore be valuable for this entire sector. In particular because the

Dutch government also has this data made open in the form of BAG data. This contributes to the efficient design of the market and challenges other parties within this market (such as perhaps more e-commerce companies) to also look at the possibilities for an open data strategy. However, since the methodology has only been implemented on a single case, it is difficult to be able to make well-validated statements about the general applicability of the method, also within other sectors. It is therefore advisable to test this and to talk with other open data experts to further check the validity of the method. Since the method should also be validated in a better or different way to be able to make informed statements about the generalizability of this study, the results of this study are based on the single case. This is further extensively explained in sub section 6.1.1.

7.5 AND NOW, ARE COMPANIES GOING TO OPEN UP THEIR DATA?

The central question in this study is how open data can contribute to value creation within commercial companies. An important first step in this is that companies must realize the potential of open data and then actually open up the data. Another question that can be asked in this study is whether this method will contribute to this. The first answer is yes. Because the ODIM has shown that it is suitable to look at the current system in a structured way and to see which datasets can be opened up and why. By looking at open data in such a way within the company, this stimulates companies in the open data development in the world. However, the identified value that comes from the ODIM will be company specific and difficult to make statements for all companies. The ODIM assists companies in the process of open data and helps in the selection process whether or not to open the data. This research contributes to the entire scientific field around open data. However, because the implementation method of the system is missing, the ODIM will not immediately cause changes at the company level. If an extra year had been used for this research, this could perhaps have been looked at, but in particular also the effect of ODIM within other types of companies and sectors. In this way, the value of open data is hopefully even more recognized.

7.6 SUMMARY CHAPTER 7

This chapter offers a reflection on many of the choices made in this study. This chapter places the entire research in a certain perspective and well-founded statements can be made regarding future research in the next chapter. For example, it has been shown that the design science approach is a good method of developing the ODIM, the artefact in this study. There are also many conclusions to be drawn about the value of open data within commer-

cial companies based on the single case study that has been done. However, this makes it difficult to make statements for other commercial companies, for which several cases need to be studied. Since the open data implementation method has only been tested on a single case, it is difficult to make firm statements about the broader applicability of the method to other cases. The results from the case study do, however, motivate the ODIM to continue to be applied to another case, given the positive outcome of the possibilities of opening up data and the value thereof. Given the scope and time frame of this research, the method has not yet been applied to multiple cases.

8 | CONCLUSION

This chapter provides a total overview of all findings in this study. This chapter forms the last part of this report and presents the conclusions on the research questions proposed. The conclusions are followed by recommendations for future research into the value of open data within commercial companies. The chapter concludes with the conclusions concerning the scientific and social relevance of this research.

Open data has been playing an increasingly prominent role in the entire data landscape since recent years. Because more and more research is being done into the possibilities of open data and the fact that more and more organizations are making their data available or making use of open data sources, the development of open data is enormous. Several studies have shown (Berends, Carrara & Radu, 2017; Manyika et al., 2013) that there is a large (economic) potential for open data. However, up to now it is often government organizations that make data openly available or that use open data. Companies are much less interested in this, partly because the value of opening up the data is not yet recognized. At present, open data is still seen by companies as an unnecessary glimpse into business data for the competition and a loss in total sales (Ahmadi Zeleti et al., 2016; Kitchin, 2014). Since the potential has been demonstrated but is not yet fully utilized, companies can be helped in this. To achieve this, the aim of this research is to develop a method that gives companies insight into which datasets can be opened up and what possible value this yields. This research therefore provides an answer to the following main research question:

How can open data contribute in value creation of digital products or services for commercial companies in the logistics sector?

To answer this question, four different sub questions have been prepared. By first answering these sub questions in the following sub sections, the answer to the main research question can be formulated.

8.1 ANSWERING THE RESEARCH SUB QUESTIONS

This section deals with all answers to the sub questions in this study. This research has been compiled on the basis of the design science approach whereby a separate sub question has been formulated for each phase in this approach. With this, each phase forms input for the subsequent phase. This will be further explained on the basis of the answers to the sub questions.

1. WHAT ARE IMPORTANT ELEMENTS FOR COMMERCIAL COMPANIES WHEN CONSIDERING OPENING UP DATA? The first sub question concerns the consideration that organizations make when they want to open up data. On the one hand the answer to this sub question provides insight into the current open data ecosystem and on the other hand shows the urgency of the identified problem. This sub question was answered on the basis of an extended literature review. First, the insights into the current open data ecosystem: This shows that research into the influences of open data takes place mainly within government organizations of economically developed countries. This finding confirms the problem identified in the introduction. Answering this sub question provided insight into the fact that a feedback cycle is required within the methodology to guarantee a correct evaluation of the potential open data system. In order to ultimately be able to determine the value of open data at companies with this method, many common challenges within earlier open data projects were investigated. These challenges are the elements that are important for organizations when they consider opening up data. This is the second part of the answer to the sub question.

Important elements for companies when they consider opening up data are in technical, economic, legal and social dimensions. When investigating the effects of opening up data for companies, these elements should be included in the testing phase of the new system. By including these elements in the consideration, well-founded statements can be made about the value of the new system. This also adds to the value of applying the method to other commercial companies.

2. HOW CAN A METHODOLOGY BE DESIGNED TO IDENTIFY THE VALUE OF OPEN DATA? The factors from the first sub question served as input for the second sub question. By analyzing how a methodology can be developed to determine the value of open data at companies, this sub question forms the answer to the goal of the solution, the second phase in design science approach. Factors from the previous sub question should be reflected in this methodology. In order to determine whether open data can be of value at a company, existing systems must be analyzed and adapted to a new system in which open data plays a central role. A lot of research has already been done into compiling these types of new systems. To answer this sub question, three different system design processes were compared. Because the iterative nature, the feedback cycles are considered important, a methodology has been chosen based on the engineering design process methodology developed by Dym et al. (2008). By adapting the steps in this method and making it suitable for open data projects, this newly developed methodology can be applied to a real life case.

3. WHAT METHODOLOGY CAN BE IMPLEMENTED TO LET COMPANIES CREATE VALUE WITH OPEN DATA? The third step in design science research is the development of an artefact. According to Offermann et al. (2010), one of the possibilities within design science research is to design a methodology as an artefact. With the framework prepared by Dym et al. (2008) from the second

sub question and the important factors for companies concerning open data from the first sub question, the methodology can be compiled. The artefact in this study is therefore the methodology that has been named "The Open Data Implementation Method" (ODIM). The purpose of the ODIM is that companies can apply this method and in this way find out which data can be made open within their current systems and what value the new system with open data provides for the company. The complete design of this method is shown in figure 8.1. In this figure the steps are shown on the left are according to the engineering design method of Dym et al. (2008) and the steps on the right are according to the ODIM developed in this study.

In the first phase of the method, it is important to first map the entire current system of the company. By making process models of the current system, performing analyzes based on the costs, revenues, users or observation of peers, compiling requirements for open data and involving stakeholders in the process at this stage, the current system can shown in detail. This phase provides a good insight into which data can be opened within the current system. The second phase consists of creating a conceptual design and setting benchmarks for value creation at companies. The factors identified in the first sub question were used to prepare these benchmarks. Research from Attard et al. (2016) and Susha, Zuiderwijk et al. (2015) has shown that these challenges can be converted into value dimensions. In this way, companies can assess the new system based on these five dimensions. These dimensions, and therefore points where open data can create value within new business systems, are: technical dimension, economic dimension, legal dimension, social dimension and organizational dimension. In the third phase, this new system must be tested on the basis of these benchmarks. When this system meets the requirements and according to the people who has tested the system value has been identified, it is possible to proceed to the fourth phase. Herein, adjustments that have emerged in the test phase are applied and a final design is presented. In the fifth and final phase, this design is communicated with, for example, an associated business model that the company can apply to their new system design.

4. WHAT VALUE WILL OPENED UP DATA CREATE WHEN APPLYING THE OPEN DATA IMPLEMENTATION METHODOLOGY ON A REAL CASE? The methodology that has been compiled based on the previous three sub questions forms the artefact in this study. In the design science approach the next phase consists of the demonstration of this artefact. By doing this in the context of this research it is possible to identify where value is gained with open data in a new system developed at a commercial company. This demonstration of the artefact was done on the basis of a case study. The case study was conducted at the data department of the largest postal delivery company in the Netherlands, PostNL Data Solutions. The case relates to the address and postal code data in the Netherlands. For years this was sold as a unique product because PostNL is the legal owner of the postal code. Until a few years ago the Dutch government published an almost exactly matching dataset as open data on their website. Applying the open data implementation methodology was therefore an excellent opportunity for this case because
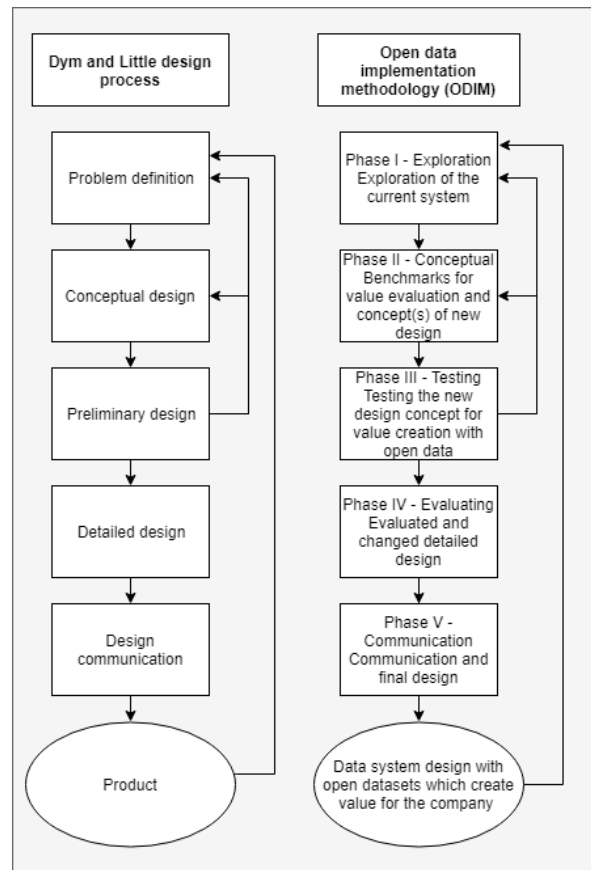
Figure 8.1: The Open Data Implementation Methodology. This picture can also be found in chapter 4. It is repeated here for easy reference.

there is experience and knowledge regarding open data and especially that the company is open to implementing a new strategy. After a extensive analysis of the current processes, a proposal was made for the design of a new system in which open data has a central role. This was then tested in the form of an expert panel workshop with PostNL employees. The central question here was what value the potentially open data could deliver in the new system. The conclusion was that the new system recognizes the value of open data in particular with regard to the economic and cultural dimensions. Open data can provide a lot of innovation on existing product structures and structural growth of the business. Attracting new customers (within the economic dimension) also emerged as an important value creation point. Open data would also be good for the branding and transparency of the company. Because users of the data can make real-time proposals for adjustments, this creates a lot of value. Leading the way in this trend of open data development is seen as valuable.

It can be concluded that applying the methodology to this case has been effective in identifying the value of open data within a newly developed system. The employees who gave feedback on the new system even felt that more data had to be made public than was proposed in the conceptual design of the system. So not only adressdata but every data that was also in the BAG database such as size of the building, building year and XY

coordinates of the building. This has been adjusted in the final design in the fourth phase of the ODIM.

However, it was concluded that there is a lack of an implementation method within the ODIM. Because it remains conceptual, it can remain difficult to estimate the actual potential value of the open data. There must also be taken account of many other elements such as the GDPR, security of the IT systems and user-friendliness of the data for the end users.

5. WHAT ARE THE LIMITATIONS OF OPEN DATA IMPLEMENTATION METHODOLOGY FOR APPLICATION AT COMMERCIAL COMPANIES? Now that the artefact from this study, the open data implementation methodology, has been tested in a real life case, the next and final phase of the design science approach can be researched. In this final phase, the artefact is evaluated and the limitations of the method become visible. The methodology was evaluated on the basis of an expert panel workshop. By having the participants in this workshop evaluate the new design that emerged from the method on the value for open data, the overall methodology could also be partially evaluated. The most important thing that emerged is that the benchmarks for value creation of open data at companies push the people who assess the system a certain, perhaps too positive, direction in assessing open data. An identified limitation of the method therefore had to do with the second phase and this phase has been adjusted in such a way that companies now have to compile their own benchmarks for the valuation of open data. Because valuation can differ per company, not only for open data within new systems, it is difficult to make generic statements about this. However, the benchmarks were set as globally as possible that it also provided a certain basis for determining the value of the open data system.

Another limitation of this method is that few statements can be made about the applicability at other companies. Because only one case has been worked out in detail in this study, the method must also be tested at other companies. Elements from the method, such as drawing up requirements and mapping the current system, prove valuable within the ODIM.

8.2   ANSWERING THE MAIN RESEARCH QUESTION

This research has used the design science approach to formulate an answer to the following main research question:

How can open data contribute in value creation of digital products or services for commercial companies in the logistics sector?

This research has focused on developing a methodology that companies can use to gain insight into their current systems and thereby determine which datasets have the potential to be opened up. In addition, the method serves to determine the value of this open data. Research following a case at PostNL Data Solutions, a commercial company in the Netherlands, has shown that

open data can create value in various areas. Firstly, a great deal of economic value can be gained by designing new types of products and services using open data. Secondly, a company also become a lot more transparent, which is good for your branding and thereby creates value for the company. Thirdly and lastly, it emerged within this case that the data became a lot more accurate and therefore better because of the feedback that end users can provide to the data sets that have been opened up.

The open data implementation methodology has demonstrated how value can be achieved with open data at a commercial company. However, it should be mentioned here that these results are based on a case within the logistics sector. In order to be able to make more statements about the value of open data within commercial companies, more research must be done and more cases must be researched by using the open data implementation method. One of the most important conclusions from this study is that the proposed methodology serves well to identify datasets that are suitable for open data for a company. The value of this open data for a company and for the end user is context dependent.

8.3 SCIENTIFIC CONTRIBUTION

On the one hand, this research has a practical effect in the form of a solution of the establishment for a new business system for PostNL Data Solutions. On the other hand, and perhaps the most important result in this research, is the scientific relevance of the subject. Because the results have a scientific relevance, they can also be used in future research and fill the identified knowledge gap.

This knowledge gap was briefly introduced in chapter 1 and extensively explained in the literature review in chapter 3. It emerged from this that research into opening up data is primarily focused on government organizations. It is often these types of organizations that make more use of this. This can be called special, since it appears that often non-public organizations are at the forefront of innovation. This often has to do with the regulations and the impact on users (of public organizations versus non-public organizations). It appears therefore that there is a gap in the current knowledge about this subject. Little is known about the effects of opening up data at commercial companies. As society becomes more open and transparent in more and more areas, for example through the growth in data availability and the growth of the internet, open data can have a major impact on companies.

This research therefore focused on filling this gap by conducting research into opening up data at commercial companies. It emerged from this that companies are often unaware of the potential value of open data and continue to think in traditional business systems in which the data remains closed. Products are then made that are often sold for a lot of money. The open data implementation methodology contributes to this by charting in an

insightful way which data has the potential to be made public and especially what value this creates for the company. This methodology needs to be further developed and can be of substantial value to the open data environment in the future. For this, however, more research needs to be done within other companies, but also for the evaluation of the method. This research serves as the correct basis in which important elements can be included in subsequent research. The benchmarks that have been drawn up in the first sub question contribute to making companies think about the different design of their systems. By also evaluating this method in other ways (see chapter 6), more well-founded statements can be made about the broad applicability of the methodology.

In conclusion, this research provides insight into how companies can open up data and what possible value this creates. For this a method is used according to previously proven design approaches. This turned out to be positive on the single case conducted in this study, making it a good addition to the literature on open data at companies. This research contributes to research for the entire open data ecosystem and offers many tools for future research.

8.4   SOCIAL CONTRIBUTION

This study investigates the possibilities for commercial companies to open up data. Giving companies insight into how and which datasets can be made public, also insights into the value of this opened up data is researched. This makes it an relevant topic in the societal field since this requires a different way of analyzing traditional business systems. To this day, data is of great value for a company because often only the financial values are considered. In conducting this research it also emerged that employees of commercial companies are not yet ready to open data within their company. A frequently heard response is that this will lead to revenue loss and unnecessary insights for the competition.

Applying the ODIM at PostNL, however, taught that by opening up certain data sets, current business systems can be set up more efficiently. This is partly because companies are going to look critically at their own data systems. In the case of the PostNL case, the open data strategy can have a major impact on the accuracy of address data and postal code data in the Netherlands. This helps the users of the data but also contributes to other systems within PostNL. For example, after opening up the data, the focus can be more on the users of the data instead of making marketable products from the data. In addition, this data can also have a extra societal value. Nowadays, where the e-commerce market is only growing and city centers are becoming fuller, efficient delivery of packages is desirable. Because the address data is more accurate, e-commerce companies will be able to benefit from this and better organize their services. Fewer packages will be returned which contributes to a more sustainable environment. But municipalities can also better organize their areas with this data that is being opened up. For

example, better statements can be made about where in the municipality the best new homes can be built.

To conclude, this research contributes to making companies think about the redesign of traditional business systems. Open data can play a central role in this, as demonstrated in the single case study at PostNL. When this is investigated at several commercial companies, but also at other types of organizations such as public organizations, open data can contribute to economic and cultural development within a country.

8.5 RECOMMENDATIONS FOR FUTURE RESEARCH

During the limited time that was available to conduct this research, a number of elements have been identified that may serve for future research. These are listed below.

1. The open data implementation methodology must be carried out at other companies to see if it is also suitable there. In this way, a better statement can be made about the value creation using open data by using this method.

2. A further deepening of the benchmarks of open data within companies must be carried out. Although this research has shown that value is business-specific, more research may lead to more widely applicable conclusions.

3. A pilot or entire implementation of the open data identified as valuable must be carried out within a company. End users of the data can be asked what value they actually derive from the data and the new system can be tested.

BIBLIOGRAPHY

Ahmadi Zeleti, F., Ojo, A. & Curry, E. (2016). Exploring the economic value of open government data. *Government Information Quarterly*, *33*(3), 535–551. doi:10.1016/J.GIQ.2016.01.008

Ali Hassan, M. I. & Twinomurinzi, H. (2018). A Systematic Literature Review of Open Government Data Research: Challenges, Opportunities and Gaps. In *2018 open innovations conference (oi)* (pp. 299–304). IEEE. doi:10.1109/OI.2018.8535794

Amui, L. B. L., Jabbour, C. J. C., de Sousa Jabbour, A. B. L. & Kannan, D. (2017). Sustainability as a dynamic organizational capability: a systematic review and a future agenda toward a sustainable transition. *Journal of Cleaner Production*, *142*, 308–322. doi:10.1016/J.JCLEPRO.2016.07.103

Andersen, A. B., Gür, N., Hose, K., Jakobsen, K. A. & Pedersen, T. B. (2015). Publishing Danish Agricultural Government Data as Semantic Web Data. (pp. 178–186). Springer, Cham. doi:10.1007/978-3-319-15615-6{_}13

ARGYRIS, C. & SCHÖN, D. A. (1989). Participatory Action Research and Action Science Compared. *American Behavioral Scientist*, *32*(5), 612–623. doi:10.1177/0002764289032005008

Armstrong, J. & Sage, A. (2000). *An introduction to systems engineering*. Retrieved from https://www.studystore.nl/p/9780471027669/an-introduction-to-systems-engineering

Attard, J., Orlandi, F. & Auer, S. (2016). Value Creation on Open Government Data. In *2016 49th hawaii international conference on system sciences (hicss)* (pp. 2605–2614). IEEE. doi:10.1109/HICSS.2016.326

Awazu, Y., Baloh, P., Desouza, K. C., Wecht, C. H., Kim, J. & Jha, S. (2009). Information-Communication Technologies Open up Innovation, Research-Technology Management. *52*(1), 51–58. doi:10.1080/08956308.2009.11657548

BAG bronhouders - Kadaster zakelijk. (2019). Retrieved from https://zakelijk.kadaster.nl/bag-bronhouders

Ben Ayed, A., Ben Halima, M. & Alimi, A. M. (2015). Big data analytics for logistics and transportation. In *2015 4th international conference on advanced logistics and transport (icalt)* (pp. 311–316). IEEE. doi:10.1109/ICAdLT.2015.7136630

Benbasat, I., Goldstein, D. K. & Mead, M. (1987). The Case Research Strategy in Studies of Information Systems. *MIS Quarterly*, *11*(3), 369. doi:10.2307/248684

Beno, M., Figl, K., Umbrich, J. & Polleres, A. (2017). Open Data Hopes and Fears: Determining the Barriers of Open Data. In *2017 conference for e-democracy and open government (cedem)* (pp. 69–81). IEEE. doi:10.1109/CeDEM.2017.22

Berends, J., Carrara, W. & Radu, C. (2017). *The Economic Benefits of Open Data*. Retrieved from http://www.europeandataportal.eu/

Boehm, B. W. (1988). A spiral model of software development and enhancement. *Computer*, *21*(5), 61–72. doi:10.1109/2.59

Bogner, A., Littig, B. & Menz, W. (2009). *Interviewing Experts*.

Borgman, J. (2017). InsideBIGDATA Guide to Use of Big Data on an Industrial Scale. Retrieved from https://insidebigdata.com/white-paper/guide-big-data-industrial-scale/

Canares, M. P. & P., M. (2014). Opening the local. In *Proceedings of the 8th international conference on theory and practice of electronic governance - icegov '14* (pp. 89–98). New York, New York, USA: ACM Press. doi:10.1145/2691195.2691214

Chan, C. M. (2013). From Open Data to Open Innovation Strategies: Creating E-Services Using Open Government Data. In *2013 46th hawaii international conference on system sciences* (pp. 1890–1899). IEEE. doi:10.1109/HICSS.2013.236

Charalabidis, Y., Zuiderwijk, A., Alexopoulos, C., Janssen, M., Lampoltshammer, T. & Ferro, E. (2018). *The World of Open Data*. Public Administration and Information Technology. Cham: Springer International Publishing. doi:10.1007/978-3-319-90850-2

Desouza, K. C. & Bhagwatwar, A. (2012). Citizen Apps to Solve Complex Urban Problems. *Journal of Urban Technology*, *19*(3), 107–136. doi:10.1080/10630732.2012.673056

Dijkman, R. M., Dumas, M. & Ouyang, C. (2008). Semantics and analysis of business process models in BPMN. *Information and Software Technology*, *50*(12), 1281–1294. doi:10.1016/j.infsof.2008.02.006

Dubé & Paré. (2003). Rigor in Information Systems Positivist Case Research: Current Practices, Trends, and Recommendations. *MIS Quarterly*, *27*(4), 597. doi:10.2307/30036550

Dyba, T., Dingsoyr, T. & Hanssen, G. K. (2007). Applying Systematic Reviews to Diverse Study Types: An Experience Report. In *First international symposium on empirical software engineering and measurement (esem 2007)* (pp. 225–234). IEEE. doi:10.1109/ESEM.2007.59

Dym, C. L., Little, P., Orwin, E. J. & Spjut, R. E. (2008). *Engineering design : a project-based introduction*. Wiley.

Eisenhardt, K. M. (1989). Building Theories from Case Study Research. *Academy of Management Review*, *14*(4), 532–550. doi:10.5465/amr.1989.4308385

European Union. (2016). General Data Protection Regulation. Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679

Foroohar, R. (2019). How much is your data worth? Retrieved from https://www.ft.com/content/3f2b0f0e-57cc-11e9-91f9-b6515a54c5b1

Frankowski, A., Van Der Steen, M. & Meijer, A. (2015). *Vorm geven aan inhoud Mark van Twist De Publieke Waarde(n) van Open Data*. Retrieved from https://www.open-overheid.nl/wp-content/uploads/2017/04/NSOB-15-17-DT_Publieke-Waarde-web-DEF.pdf

Gonzalez, R. (2010). A Framework for ICT-Supported Coordination in Crisis Response. Retrieved from https://repository.tudelft.nl/islandora/object/uuid:7528c7fd-9eac-43b7-95cf-2646aef4c132

Gurin, J. (2014). Open data now: the secret to hot startups, smart investing, savvy marketing, and fast innovation.

Hammell, R., Bates, C., Lewis, H., Perricos, C., Brett, L. & Branch, D. (2012). *Open data Driving growth, ingenuity and innovation Contents*.

Hevner, A. & Chatterjee, S. (2010). Design Science Research in Information Systems. (pp. 9–22). Springer, Boston, MA. doi:10.1007/978-1-4419-5653-8{_}2

Ho, L.-C. J., Liu, C.-S. & Ramanan, R. (1997). *Open-Market Stock Repurchase Announcements and Revaluation of Prior Accounting* (tech. rep. No. 3). Retrieved from https://www.jstor.org/stable/pdf/248481.pdf?refreqid=excelsior%3A190ff5e5f7ca0a6351009605fec99356

Huijboom, N. & van den Broek, T. (2011). Open data: an international comparison of strategies. Retrieved from www.epracticejournal.eu

Iivari, J. & Venable, J. (2009). Action research and design science research - Seemingly similar but decisively dissimilar. *ECIS 2009 Proceedings*. Retrieved from https://aisel.aisnet.org/ecis2009/73

Jaakola, A., Kekkonen, H., Lahti, T. & Manninen, A. (2015). Open data, open cities: Experiences from the Helsinki Metropolitan Area. Case Helsinki Region Infoshare www.hri.fi. *Statistical Journal of the IAOS*, *31*, 117–122. doi:10.3233/SJI-150873

Janssen, M., Charalabidis, Y. & Zuiderwijk, A. (2012). Benefits, Adoption Barriers and Myths of Open Data and Open Government. *Information Systems Management*, *29*(4), 258–268. doi:10.1080/10580530.2012.716740

Jetzek, T., Avital, M. & Bjørn-Andersen, N. (2014). Generating Sustainable Value from Open Data in a Sharing Society. (pp. 62–82). Springer, Berlin, Heidelberg. doi:10.1007/978-3-662-43459-8{_}5

Kalampokis, E., Tambouris, E. & Tarabanis, K. (2013). Linked Open Government Data Analytics. (pp. 99–110). Springer, Berlin, Heidelberg. doi:10.1007/978-3-642-40358-3{_}9

Kareken, J. H., Muench, T. & Wallace, N. (1973). *American Economic Association Optimal Open Market Strategy: The Use of Information Variables* (tech. rep. No. 1). Retrieved from https://about.jstor.org/terms

Kim, H. & Heo, U. (2017). Comparative Analysis of Economic Development in South Korea and Taiwan: Lessons for Other Developing Countries. *Asian Perspective*, *41*(1), 17–41. doi:10.1353/apr.2017.0002

Kitchin, R. (2014). *The data revolution: Big data, open data, data infrastructures and their consequences*. Sage.

Kotiadis, K. & Robinson, S. (2008). Conceptual modelling: knowledge acquisition and model abstraction. (Winter Simulation Conference), 951–958.

Lakomaa, E. & Kallberg, J. (2013). Open Data as a Foundation for Innovation: The Enabling Effect of Free Public Sector Information for Entrepreneurs. *IEEE Access*, *1*, 558–563. doi:10.1109/ACCESS.2013.2279164

Magalhaes, G., Roseira, C. & Manley, L. (2014). Business Models for Open Government Data. doi:10.1145/2691195.2691273

Maidique, M. A. & Zirger, B. J. (1984). A study of success and failure in product innovation: The case of the U.S. electronics industry. *IEEE Transactions on Engineering Management*, *EM-31*(4), 192–203. doi:10.1109/TEM.1984.6447537

Malamud, C., O'Reilly, T., Elin, G., Sifry, M., Holovaty, A., O'Neil, D., ...
Lessig, L. (2013). Principles of Open Government Data. *Open Government Working Group*. Retrieved from http://www.%20opengovdata.%20org/home/8principles

Manyika, J., Chui, M., Groves, P., Farrell, D., Van Kuiken, S. & Doshi, E. A. (2013). Open data: Unlocking innovation and performance with liquid information. *McKinsey Global Institute*, *21*.

Maramieri, J. (2014). *Journal of e-Learning and Knowledge Society.* Italian e-Learning Association. Retrieved from https://www.learntechlib.org/p/148271/

March, S. T. & Smith, G. F. (1995). Design and natural science research on information technology. *Decision Support Systems*, *15*(4), 251–266. doi:10.1016/0167-9236(94)00041-2

Marin de la Iglesia, J. L. (2014). Alternative estimation of "public procurement advertised in the Official Journal as % of GDP" official indicator using open government data. *Computers in Industry*, *65*(5), 905–912. doi:10.1016/J.COMPIND.2014.03.004

Ministerie van Economische Zaken. (2017). *Analyse toekomst postmarkt*. Retrieved from https://www.rijksoverheid.nl/documenten/rapporten/2017/07/10/analyse-toekomst-postmarkt

Ministerie van Infrastructuur en Milieu. (2014). Convenant inzake postcodes.

Offermann, P., Blom, S., Schönherr, M. & Bub, U. (2010). Artifact Types in Information Systems Design Science – A Literature Review. (pp. 77–92). Springer, Berlin, Heidelberg. doi:10.1007/978-3-642-13335-0{_}6

Opdenakker, R. (2006). Advantages and Disadvantages of Four Interview Techniques in Qualitative Research. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, *7*(4). doi:10.17169/FQS-7.4.175

Parmenter, D. (2015). *Key performance indicators : developing, implementing, and using winning KPIs*. Retrieved from https://books.google.nl/books?hl=nl&lr=&id=bKkxBwAAQBAJ&oi=fnd&pg=PA101&dq=key+performance+indicators&ots=cYZXfoi2bs&sig=5z_raXiZZlycF_uokgZKHdO6CYw#v=onepage&q=key%20performance%20indicators&f=false

Peffers, K., Tuunanen, T., Rothenberger, M. A. & Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems*, *24*(3), 45–77. doi:10.2753/MIS0742-1222240302

Radl, W., Skopek, J., Komendera, A., Jäger, S. & Mödritscher, F. (2013). And Data for All: On the Validity and Usefulness of Open Government Data. In *Proceedings of the 13th international conference on knowledge management and knowledge technologies - i-know '13* (pp. 1–4). New York, New York, USA: ACM Press. doi:10.1145/2494188.2494228

Royce, W. W. (1987). *Managing the development of large software systems: concepts and techniques*. IEEE Computer Society Press. Retrieved from http://www-scf.usc.edu/~csci201/lectures/Lecture11/royce1970.pdf

Safarov, I., Meijer, A. & Grimmelikhuijsen, S. (2017). Utilization of open government data: A systematic literature review of types, conditions, effects and users. *Information Polity*, *22*, 1–24. doi:10.3233/IP-160012

Schermer, B. & Wagemans, T. (2009). *Onze digitale schaduw: Een verkennend onderzoek naar het aantal databases waarin de gemiddelde Nederlander geregistreerd staat*. Retrieved from https://openaccess.leidenuniv.nl/bitstream/handle/1887/15888/rap_2009_onze_digitale_schaduw.pdf?sequence=1

Stellman, A. & Greene, J. (2005). *Applied software project management*. O'Reilly Media, Inc.

Susha, I., Grönlund, Å. & Janssen, M. (2015). Driving factors of service innovation using open government data: An exploratory study of entrepreneurs in two countries. *Information Polity*, *20*, 19–34. doi:10.3233/IP-150353

Susha, I., Zuiderwijk, A., Janssen, M. & Grönlund, Å. (2015). Benchmarks for Evaluating the Progress of Open Data Adoption. *Social Science Computer Review*, *33*(5), 613–630. doi:10.1177/0894439314560852

Thomson Reuters. (2019). Study finds organizations are not ready for GDPR compliance issues. Retrieved from https://legal.thomsonreuters.com/en/insights/articles/study-finds-organizations-not-ready-gdpr-compliance-issues

van Ooijen, C. (2014). *Het risico van de mobiele burger : Publieke... (2014) | www.narcis.nl*. Retrieved from https://www.narcis.nl/publication/RecordID/oai:tilburguniversity.edu:publications%2Ffebb9414-a8aa-47d0-b103-6ed88f8ae22f

Vert, S. & Vasiu, R. (2015). *Smart Cities and Communities View project Dynamic textures View project Integrating linked open data in mobile augmented reality applications-a case study*. Retrieved from www.temjournal.com

Volpi, V., Ingrosso, A., Pazzola, M., Opromolla, A. & Medaglia, C. M. (2014). Roma Crash Map: An Open Data Visualization Tool for the Municipalities of Rome. (pp. 284–295). Springer, Cham. doi:10.1007/978-3-319-07863-2{_}28

Wang, H.-J. & Lo, J. (2019). Factors Influencing the Adoption of Open Government Data at the Firm Level. *IEEE Transactions on Engineering Management*, 1–13. doi:10.1109/TEM.2019.2898107

Wee, B. V. & Banister, D. (2016). How to Write a Literature Review Paper? *Transport Reviews*, *36*(2), 278–288. doi:10.1080/01441647.2015.1065456

Yin, R. (2003). Case Study Research: Design and Methods. *Thoasand Oaks: SAGE publications*.

Zeleti, F. A. & Ojo, A. (2018). Capability Model for Open Data. In *Proceedings of the 11th international conference on theory and practice of electronic governance - icegov '18* (pp. 403–411). New York, New York, USA: ACM Press. doi:10.1145/3209415.3209492

Zuiderwijk, A. (2015). *Open data infrastructures: The design of an infrastructure to enhance the coordination of open data use* (Doctoral dissertation, TU Delft). Retrieved from https://repository.tudelft.nl/islandora/object/uuid:9b9e60bc-1edd-449a-84c6-7485d9bde012/

Zuiderwijk, A., Helbig, N., Gil-Garcia, J. & Janssen, M. (2014). Special Issue on Innovation through Open Data - A Review of the State-of-the-Art and an Emerging Research Agenda: Guest Editors Introduction. *Journal of theoretical and applied electronic commerce research*, *9*(2), 1–2. doi:10.4067/S0718-18762014000200001

Zuiderwijk, A. & Janssen, M. (2014a). Open data policies, their implementation and impact: A framework for comparison. *Government Information Quarterly*, *31*(1), 17–29. doi:10.1016/J.GIQ.2013.04.003

Zuiderwijk, A. & Janssen, M. (2014b). The negative effects of open government data - investigating the dark side of open data. In *Proceedings of the 15th annual international conference on digital government research - dg.o '14* (pp. 147–152). New York, New York, USA: ACM Press. doi:10.1145/2612733.2612761

Zuiderwijk, A., Janssen, M., Poulis, K. & Van De Kaa, G. (2015). Open Data for Competitive Advantage: Insights from Open Data Use by Companies. doi:10.1145/2757401.2757411

Zuiderwijk, A., Shinde, R. & Janssen, M. (2018). Investigating the attainment of open government data objectives: Is there a mismatch between objectives and results? *International Review of Administrative Sciences*, 002085231773911. doi:10.1177/0020852317739115

A | EXTENDED LITERATURE REVIEW ON OPEN DATA

Chapter 3 contains the results of an extensive literature study of open data. Specifically, this chapter provides insight into the considerations that companies have when they consider opening up data. Research has been carried out to substantiate these elements that come back in here. The most important papers are already listed in chapter 3 but are explained in more detail below. This appendix serves as a supplement and further deepening.

1 (Jaakola et al., 2015) This paper provides insight through conducting a case study on how an open data service can be set up. In addition, it provides insight into how opening up data can become a part of employees' everyday tasks. The most important conclusion drawn from this is that opening the data alone is not enough. It is also important to encourage the usage of the open data. Challenges that emerge here are of a legal, economical and technical level.

2 (Susha, Grönlund & Janssen, 2015) The main drivers for businesses to adopt open data are described in this research. The paper examines the trend that open data is becoming more accessible and that it has already been demonstrated that it can generate a lot of potential for companies. The most important drivers were identified by means of a survey. These have to do with innovation and improving skills and expertise. It also emerges that the adoption of open data can vary greatly per business.

3 (Magalhaes et al., 2014) More and more government organizations are undertaking projects to open up databases. This study looked at the commercial re-use of OGD and what the effects are. 500 American companies were investigated through a survey. The conclusions are that the re-use of OGD ensures greater transparency, increased public sector efficiency and open data innovation. The framework used to analyze this data is of high scientific value and can easily be reused in other cases.

4 (Kalampokis et al., 2013) The reason that governments are increasingly opening up their data is that science has shown that this entails many opportunities. However, recent studies also show that the potential of this is not yet fully utilized. This study shows that the combination of OGD and statistical data analysis leads to new insights and possibilities. Based on a case study in the UK, this research shows that opening the data alone is not enough, but that an analysis of this data provides much added value. This added value is mainly in the area of social value.

5 (Marin de la Iglesia, 2014) This research has been conducted in six different countries in Europe. Calculations using a previously proven indicator have shown that the value of products can increase when aggregation of public

data takes place. Certain processes can be executed faster and cheaper when open data is used. Benefits were achieved because the data was more accurate and users could handle it better and had better insight into the processes. A challenge is to reach a audience as large as possible in order to spread this concept widely.

6 (Zuiderwijk et al., 2015) According to government organizations, open data has a lot of potential. This research examines the possibilities for companies to create competitive advantage with the help of open data. Various theories, surveys and interviews have been applied to draw conclusions about this. The most important conclusion is that a company must have in-house capabilities and resources to create competitive advantage using open data. Whether this is actually achieved has not been investigated and is also very difficult according to the researchers.

7 (Lakomaa & Kallberg, 2013) This research shows that open data can be of great added value for entrpreneurs. A survey in Sweden has shown that entrepreneurs gain a lot of value to data that is made available by other companies or government organizations. This pleases the possibilities of executing their business plans and the possibilities of innovation. The conclusion is that open data can be seen as an enabler to innovation for entrepreneurs. 8 (Radl, Skopek, Komendera, Jäger & Mödritscher, 2013) Much of the literature provides insight into the possibilities of open data for companies. This paper deals extensively with the data quality of open data and then with the usefulness of the data. By conducting a correlation test, this study shows that open data is often reliable and can be very useful for companies. It can bring benefits to various stakeholders and commercial use can also be applied.

9 (Andersen et al., 2015) A case study in Denmark on agricultural government data examines the value of establishing various links in open data. The writers show that opening up the data alone is not enough, but with the creation of certain links, via a certain framework, a lot of new insights can be gained. Challenges are for users who cannot always make these links themselves, so that the full potential of open data is not exploited. A number of technical preconditions for open data have also been identified.

10 (Vert & Vasiu, 2015) Linked data is a concept that is clearly expressed in this paper. A case study has shown that open data does not offer enough for users' mobile applications, for example. But by linking the data, or allowing users to adjust the data in real time, the data becomes much more accurate. In this way the APIs that are built on this data are much better and the users are more satisfied with it.

11 (Maramieri, 2014) This study shows that organizations, mainly government organizations, often do not open their data due to legal reasons. Regulations must be adapted to this and the options must be better highlighted. Only when this happens can open data grow to its full potential. Companies can then use it on a larger scale and may themselves also have the insight to open up (parts of) their data.

12 (Desouza & Bhagwatwar, 2012) A case study was conducted to solve complex urban problems. By opening up data, the government offers citizens the opportunity to add data to this themselves. This makes data such as traffic jams, weather and water consumption more accurate which can result in a safer city and it will improve the management of public utilities.

13 (Volpi et al., 2014) Where a number of studies have used a mobile application to allow governments to use open data more by citizens, this study uses a visualization tool. The conclusions concerning the use of open data is that it ensures a great deal of cooperation between citizens in order to optimize the use of the data. The research shows that the greater the input of data, both the initial open data and the user data, the wider the deployment. This directly causes the challenges.

14 (Wang & Lo, 2019) Top management support, government support, competitive pressure, perceived benefits and perceived barriers are all factors that influence the adoption of open government data by firms. Both social and technical challenges influence the adoption of open data by firms. These factors can serve as a basis for investigating why companies do or do not use open data, but also why companies themselves can or cannot open up data.

15 (Canares & P., 2014) This paper demonstrates the imbalance between the supply side and the demand side around open data. A case study in the Phillipines shows that the government has opened up certain data, but that the users, the citizens, often could not work with it or did not even know it existed. A challenge from this is to solve the mismatch between supply and demand of open data. Transparency and accountability are important factors in this.

B | TRANSCRIPT OF INTERVIEWS

To get a good idea of the complete system of the Postcodetabel, several interviews were conducted. These interviews were conducted with PostNL Data Solutions employees. The people were chosen from the two most important departments within this company for this research: Marketing & Sales and Data & IT. Various questions had been thought up in advance, but in practice the interviews often became open conversations about the entire PCT system and potential changes in this system. In total, seven interviews were conducted with employees of the Marketing & Sales department and a total of nine interviews with employees of the Data & IT department. Each interview had the same structure that can be traced to the following four core topics: identification of the problem, stakeholders in the system, product specifications and possible scenarios. The following section shows a list of elements and questions that were used per topic. The summaries of all interviews are shown in the section after that. As all the employees interviewed were Dutch, the details of the interviews are also given in Dutch. The interviews took place at the end of March and the beginning of April 2019.

B.1 FOUR DIFFERENT KEY SUBJECTS

- **Identification of the problem**
  Heeft het de PCT buiten commerciële waarde ook nog andere waarde?
  Vanuit welke behoefte is de PCT ooit ontstaan?
  Wat maakt het product hedendaags nog uniek en hoe is dat ten opzichte van vroeger?

- **Stakeholders in the system**
  Wat voor soort klanten nemen de PCT af?
  Product bestaat al 25 jaar, hebben de klanten genoeg inzicht in waar zij het product voor gebruiken?
  Welke partijen leveren datasets als input voor de PCT?
  Wat is de toegevoegde waarde ten opzichte van de concurrentie?

- **Product specifications**
  Hoe is het product opgebouwd en hoe is dit veranderd in de afgelopen jaren?
  Wat mist er volgens u aan de huidige productstructuur?

- **Possible scenarios**
  Wat ziet u als mogelijk toekomst scenario voor de PCT?
  Wat zijn de belangrijkste variabelen binnen de PCT om rekening mee
  te houden wanneer een nieuw PCT systeem wordt ontworpen?

B.2 SUMMARIES OF THE INTERVIEWS

In this section the most important elements from the key subjects from the
previous section are shown as concisely as possible. This includes all relev-
ant answers from the employees of both departments within PostNL DS.

IDENTIFICATION OF THE PROBLEM   PCT wordt steeds meer genaderd door
de concurrentie, wellicht zijn sommige er al voorbij. PCT gaat nu ook aanlev-
eren in BAG format. Voordeel van de PCT is dat dit product garandeert dat
de postcode en adrescombinatie een afgiftepunt is. Dit is in tegenstelling tot
bijvoorbeeld de BAG waarbij een postcode ook een elektriciteitshuisje kan
zijn of een failliet bedrijf bijvoorbeeld. Het PostNL imago is een hele belan-
grijke factor in dit gebeuren en de reden waarom veel klanten het product
de PCT nog wel steeds aanschaffen. Echter daalt de omzet e het aantal
klanten omdat met name de BAG, de open data variant van de PCT van de
Nederlandse overheid, groeit. Er ontstaan twee waarheden van postcode en
adresdata ook bijvoorbeeld omdat gemeentes besluiten zelf postcodes aan te
gaan maken, dit kan nooit de bedoeling zijn. probleem zit hem erin dat BAG
eigenaar is van straatnaam en huisnummer en PostNL van de postcode, dit
veroorzaakt scheve datasets. Gegevens worden wel uitgewisseld maar lang
niet altijd even goed. Hier is ook nooit een goede strategie op bedacht hoe
we hier goed mee kunnen omgaan. Dit heeft onder meer als gevolg dat veel
klanten worden niet eens gebeld worden omdat er angst heerst dat de klant
er dan achter komt dat deze het product eigenlijk niet meer gebruikt en de
waarde verliest. Hoe lang is het product dan nog levend? Algemene tendens
binnen Data Solutions is wel het product te behouden omdat er het meeste
geld in verdiend wordt maar er heerst grote twijfel over de toekomst van de
productstrategie. Voornamelijk door de open data variant BAG.

STAKEHOLDERS IN THE SYSTEM   Afname van de postcodetabel gebeurt
onder andere door gemeentes, grote ecommerce partijen, verzekeraars en
banken. Gemeentes zijn echter verplicht sinds een aantal jaar om het BAG
te gebruiken. Veel zijn dan ook overgestapt op het gebruiken van twee syste-
men: de PCT en het BAG. Niet elke gemeente houdt zich aan deze verpicht-
ing waardoor veel nog de PCT alleen afnemen. Dit zal gedurende de jaren
ook nog wel meer af gaan nemen. Veel slapende klanten die het product niet
gebruiken maar wel betalen. Groot deel ook oude klanten die vastgeroeste
systemen hebben die niet makkelijk kunnen schakelen naar een ander soort
postcode input. Maar zodra deze gaan innoveren is de kans klein dat zij
weer de PCT afnemen in de huidige staat. Daarom wordt de PCT ook al

wel aangevuld met de BAG data maar veel partijen zien toch meer in dan alleen de afname van BAG data en niet meer betalen voor PCT data. In totaal zijn er ongeveer 1300 klanten die de postcodetabel nu afnemen. Elke klant heeft zijn eigen systemen hierop aangepast. PostNL was altijd, aanname tot 10 jaar geleden, de partij met de marktmacht en had/heeft het beste imago. PostNL is ook de wettelijke eigenaar van de postcode. Dit zal altijd zo blijven. PostNL is de duurste aanbieder van deze data en biedt hetzelfde als veel concurrenten. Reden dat veel klanten bij PostNL zitten is het goede imago en daarom de betrouwbaarheid. Eindgebruikers maken nu nog de afweging in prijs (goedkoper bij de BAG) en kwaliteit (de PCT is beter). Andere concurreten, zoals postcode.nl, zijn ook alleen maar aan het groeien dus verandering in de strategie is nodig. Postcodedata zal wel altijd interessant blijven omdat veel uitspraken gedaan kunnen worden over bepaalde gebieden.

PRODUCT SPECIFICATIONS In de beantwoording van deze sectie zijn alle specificaties rondom het hele PCT systeem in kaart gebracht. Deze worden hier niet verder behandeld omdat dit in hoofdstuk vijf al in de procesmodellen is weergegeven. De interviews hebben een aanzienlijke bijdrage gehad in de samenstelling van deze modellen. Product is echt heel erg goed. PostNL is de eigenaar van de postcode dus wij weten dat we het altijd bij het juiste eind hebben. BAG data kunnen bijvoorbeeld fouten in zitten door het verkeerd overnemen van de juiste postcode bij een adres. Wel wordt erkend dat het product eigenlijk al heel erg lang niet meer uniek is en er veranderingen nodig zijn om het een interessant product te laten blijven. Er zijn nu ook te weinig inzichten waar de huidige klanten de postcodetabel bijvoorbeeld voor gebruiken. Dit baseren we nog op contracten die in het verleden zijn afgesloten en klanten waarvan we ervan uitgaan dat ze nog op dezelfde manier het product gebruiken. Product is erop gebouwd dat mensen vroeger veel moesten betalen voor de opslag van data dat is nu niet meer het geval. Nu veel afhankelijker van de PCT, voor haar producten is dit altijd de basis. Variatie binnen het product kan bijvoorbeeld gezien worden in verschillen op reeks niveau of huisnummer niveau, met of zonder leestekens, met of zonder BAG variant.

Quote: " Voor mijn gevoel ontbreekt er op dit moment niks aan het product. Het product in de huidige staat is al 25 jaar op deze manier opgebouwd en weinig aan veranderd. De echte cashcow van het bedrijf maar omzet is flink aan het dalen."

Een belangrijke conclusie is dat de huidige productstructuur is niet aanpasbaar is. Hier moet een nieuw systeem voor worden bedacht. Algemene tendens is dat de PCT op dit moment beter is dan alle andere producten in de markt met name door de betrouwbaarheid van de data.

POSSIBLE SCENARIOS Stop met reeks producten en biedt alleen de PCT aan op huisnummer niveau. Vul dit eventueel aan met BAG data aangezien BAG data van grote waarde is en kan zijn in de toekomst. Op deze manier groeien BAG en PCT steeds meer naar elkaar toe en is er eventueel een mo-

gelijkheid tot de creatie van een centraal adres- en postcodebestand. PostNL kan er ook voor kiezen alleen de BAG data te gebruiken en dit aan te vullen met bedrijfsunieke data zoals de grootte van de brievenbus. Hiermee moet goed opgepast worden met de nieuwe privacywetgeving. Veranderingen binnen bedrijven zijn moeilijk, en zeker binnen zo een grote organisatie als PostNL. Wanneer systemen aangepast moeten worden is het goed om te beginnen bij de basis en vanaf daar opnieuw op te bouwen. Hou hierbij altijd het algehele doel van PostNL in je achterhoofd, en niet alleen de doelen van DS. Postcode is de basis van heel PostNL. En als data solutions zijn wij in dit scenario degene die alle data hebben zodat we naar waarheid kunnen communiceren en verkopen. Bijvoorbeeld bij sorteren en adres check wordt een andere dataset gebruikt dan hier bij data solutions, dat moet anders kunnen. Probleem is nu ook dat er veel discussie intern is over de waarde van postcode data. Wanneer we over gaan op een open data variant is deze discussie deels weg en kunnen we focussen op andere aspecten. Het gaat nu vooral over op korte termijn snel geld verdienen. Belangrijk is om te luisteren naar wat de eindgebruikers nou daadwerkelijk willen. BAG kwaliteit is niet goed genoeg maar dit is een kwestie van tijd

# C | DETAILED TECHNICAL ANALYSIS OF THE POSTCODETABEL

In this appendix all products within the PCT system are displayed in table C.1. This contributes to the complete picture of the PCT system and how the complex relationships work. In addition, this appendix contains various figures and graphs that support statements in the main text. This is mainly about the decline in sales of PCT products and the number of customers. This appendix is therefore partly intended as confirmation that a change in the current system is necessary and that a different strategy must be pursued.

Name of the product		
Postcodetabel	Postcode	PCT
op huisnummerniveau wekelijks	tabel op reeksniveau voor VSVS	huisnr 10-99 bestandslevering HB2
Mutaties postcodetabel op HUISNUMMERniveau	Nieuwbouw Adreslevering tbv Omnidata(Ivo Temmink)	PCT huisnr 10-99 Initieel HE2
Vernummeringen Postcodetabel	Nieuwbouw Adreslevering tbv Vetinair Centrum	PCT huisnr 10-99 mutaties HM2
Postcodetabel op reeksniveau	Nieuwbouw Adreslevering tbv Vomar	PCT huisnr 10-99 onderhoudsabonnement HO2
Postcodetabel op huisnummerniveau	Nieuwbouw Adreslevering tbv Wehkamp N.V.	PCT huisnr 1-9 bestandslevering HB1
Mutaties Postcodetabel op REEKSniveau	APB Adreslevering tbv Omnidata(Laurens J. Ruster)	PCT huisnr 1-9 Initieel HE1
Week tabel op reeksniveau	Nieuwbouwadressen voor VSVS	PCT huisnr 1-9 mutaties HM1
Postcodetabel op reeksniveau wekelijks	APB Adreslevering tbv KPN-Telecom	PCT huisnr 1-9 onderhoudsabonnement HO1
Woonplaatsen ervaringstabel 2	Postbodewijken voor VSVS	PCT huisnummerniveau
Woonplaatsen ervaringstabel 2w	APB Adreslevering tbv Omnidata(DMC)	PCT huisnummerniveau bestandslevering
mutaties postcodetabel huisnummerniv + diakrieten	APB adreslevering temp.adres.gbrdmdat	PCT huisnummerniveau dupl lev bij mutatieabonn HP2
Mutaties postcodetabel op reeksniveau Neckermann	Mutaties postcodetabel op reeksniveau Wehkamp	PCT huisnummerniveau duplicaatlevering
mutaties postcodetabel reeksniveau bagcompact	APB levering voor BRIDGIS	PCT huisnummerniveau duplicaatlevering HP1
mutaties postcodetabel huisnummerniv + diak week	Interfacebestand reeksmutaties voor de PCT-browser	PCT huisnummerniveau initieel
mutaties postcodetabel reeksniveau met diakrieten	Interfacebestand huisnummermutaties voor de PCT-browser	PCT huisnummerniveau mutaties
postcodetabel op huisnummerniv met diakrieten	Interfacebestand vernummeringen voor de PCT-browser	PCT netnummerbestand 10-99 NT2
postcodetabel op reeksniveau voor bagcompact	APB Adreslevering tbv Omnidata(Ren¿ Meijers)	PCT netnummerbestand 1-9 NT1
postcodetabel op reeksniveau met diakrieten	APB Aantallen levering op Postcode voor de Gouden Gids	PCT netnummertabel
PCMUTR met buurtcodes en XY maandelijks	Adres levering tbv AddressDoctor GmbH	PCT reeks 10-99 bestandslevering RB2
PCTR met buurtcodes en XY maandelijks	PCT reeks 10-99 Initieel RE2	PCT reeks 10-99 mutaties diakriet RM4
Verkort huisnummertabel +xy +geo	PCT reeks 10-99 onderhoudsabonnement RO2	PCT reeks 10-99 mutaties wijken en buurten WM2
Verkort huisnummertabel +xy +geo maand	PCT reeks 1-9 bestandslevering RB1	PCT reeks 1-9 bstd diakriet RB3
Postcodetabel op reeksniveau AdresXpress	PCT reeks 1-9 Initieel RE1	PCT reeks 1-9 initieel diakriet RE3
APB Aantallen levering op Postcode	PCT reeks 1-9 onderhoudsabonnement RO1	PCT reeks 1-9 initieel wijken en buurten WE1
gbr levering 2	PCT reeksniveau	PCT reeks 1-9 mutaties diakriet RM3
APB Aantallen ongeadreseerd(netto)	PCT reeksniveau bestandslevering	PCT reeks 1-9 mutaties wijken en buurten WM1
APB Aantallen levering op woonplaatswijkcode	PCT reeksniveau dupl lev bij mutatieabonn RP2	PCT Special
Deellevering postcodetabel 1000-1999	PCT reeksniveau duplicaatlevering	producten
Deellevering postcodetabel 2000-2999	PCT reeksniveau duplicaatlevering RP1	PCT Special Aantallen Ongeadresseerd
Deellevering postcodetabel 3000-3999	PCT reeksniveau initieel	PCT reeksniveau mutaties Neckermann
Deellevering postcodetabel 4000-4999	PCT reeksniveau mutaties	PCT reeksniveau mutaties Wehkamp
Deellevering postcodetabel 5000-5999	PCT reeksniveau 10-99 mutaties RM2	PCT Special APB Aantallenleveringen APBP
Deellevering postcodetabel 6000-6999	PCT reeksniveau 1-9 mutaties RM1	PCT Special APB Aantallenleveringen APBW
Deellevering postcodetabel 7000-7999	PCT Special	PCT Special APB Adresleveringen ADRES RENE
Deellevering postcodetabel 8000-8999	huisnummerniveau	PCT Special APB Adresleveringen ADRES15
Deellevering postcodetabel 9000-9999	PCT vernummeringenbestand 10-99 VN2	PCT Special APB Adresleveringen ADRES2
Postbodewijken	PCT vernummeringenbestand 1-9 VN1	PCT Special APB Adresleveringen ADRES4
Gbr_andes_cbs	PCT vernummeringentabel	PCT Special APB NieuwbouwAdresleveringen 100
GBRMXSTREAM2	PCT Woonplaatsen Ervaringstabel	PCT Special APB NieuwbouwAdresleveringen 5
Interfacebestand voor Nieuwe Nieuwbouwadressen	PCT Woonplaatservaringbestand 10-99 WP2	PCT Special APB NieuwbouwAdresleveringen 6
Deellevering postcodetabel 1000-1999 Huisnummer	PCT Woonplaatservaringbestand 1-9 WP1	PCT Special Gouden Gids
Deellevering postcodetabel 2000-2999 Huisnummer	PCT woonplaatservaringstabel	PCT Special Nieuwbouwadressen
Deellevering postcodetabel 3000-3999 Huisnummer	Postcodetabel op HUISNUMMERniveau	PCT Special Postbodewijk
Deellevering postcodetabel 4000-4999 Huisnummer	Postcodetabel op REEKSniveau	PCT Special Postcode tabel op reeksniveau - Jaar
Deellevering postcodetabel 5000-5999 Huisnummer	PCT huisnr 10-99 deellevering initieel HD2	PCT Special PostcodeTabel op reeksniveau - PCTR
Deellevering postcodetabel 6000-6999 Huisnummer	PCT huisnr 1-9 deellevering initieel HD1	PCT Special PostcodeTabel vernummeringen
Deellevering postcodetabel 7000-7999 Huisnummer	PCT huisnummerniveau deellevering	PCT vernummeringentabel Neckermann
Deellevering postcodetabel 8000-8999 Huisnummer	PCT reeks 10-99 deellevering initieel RD2	PCT
Deellevering postcodetabel 9000-9999 Huisnummer	PCT reeks 1-9 deellevering initieel RD1	BAG reeks 1-9 initieel PBRE1
GBRMXSTREAM3	PCT reeksniveau deellevering	PCT BAG reeks 1-9 mutaties PBRM1
AdresXpress huisnummer	PCT APB aantallen en perceelcodes	PCT huisnr 10-99 bstd diakriet HB4
Aantal ongeadresseerd per PC6 met bestelloop voor CI analyse	PCT BAG reeks 10-99 bstd PBRB2	PCT huisnr 10-99 bstd diakriet XY HB8
levering tbv APB-plus	PCT BAG reeks 10-99 initieel PBRE2	PCT huisnr 10-99 initieel diakriet HE4
Adres levering tbv datarotonde , alle velden in één tabel	PCT BAG reeks 10-99 mutaties PBRM2	PCT huisnr 10-99 initieel diakriet XY HE8
PCT huisnr 10-99 initieel verkort geografisch HE6	PCT huisnr 1-9 initieel diakriet HE3	PCT reeks 10-99 initieel wijken en buurten WE2
PCT huisnr 10-99 mut. diakriet XY HM8	PCT huisnr 1-9 initieel diakriet XY HE9	PCT XY
PCT huisnr 10-99 mutaties diakriet HM4	PCT huisnr 1-9 initieel verkort geografisch HE5	
PCT huisnr 1-9 bstd diakriet HB3	PCT huisnr 1-9 mut. diakriet XY HM9	
PCT huisnr 1-9 bstd diakriet XY HB9	PCT huisnr 1-9 mutaties diakriet HM3	

Table C.1: All different kind of products within the PCT system

```
figs/"Total revenue and number of customers sequence products over the last four year
```
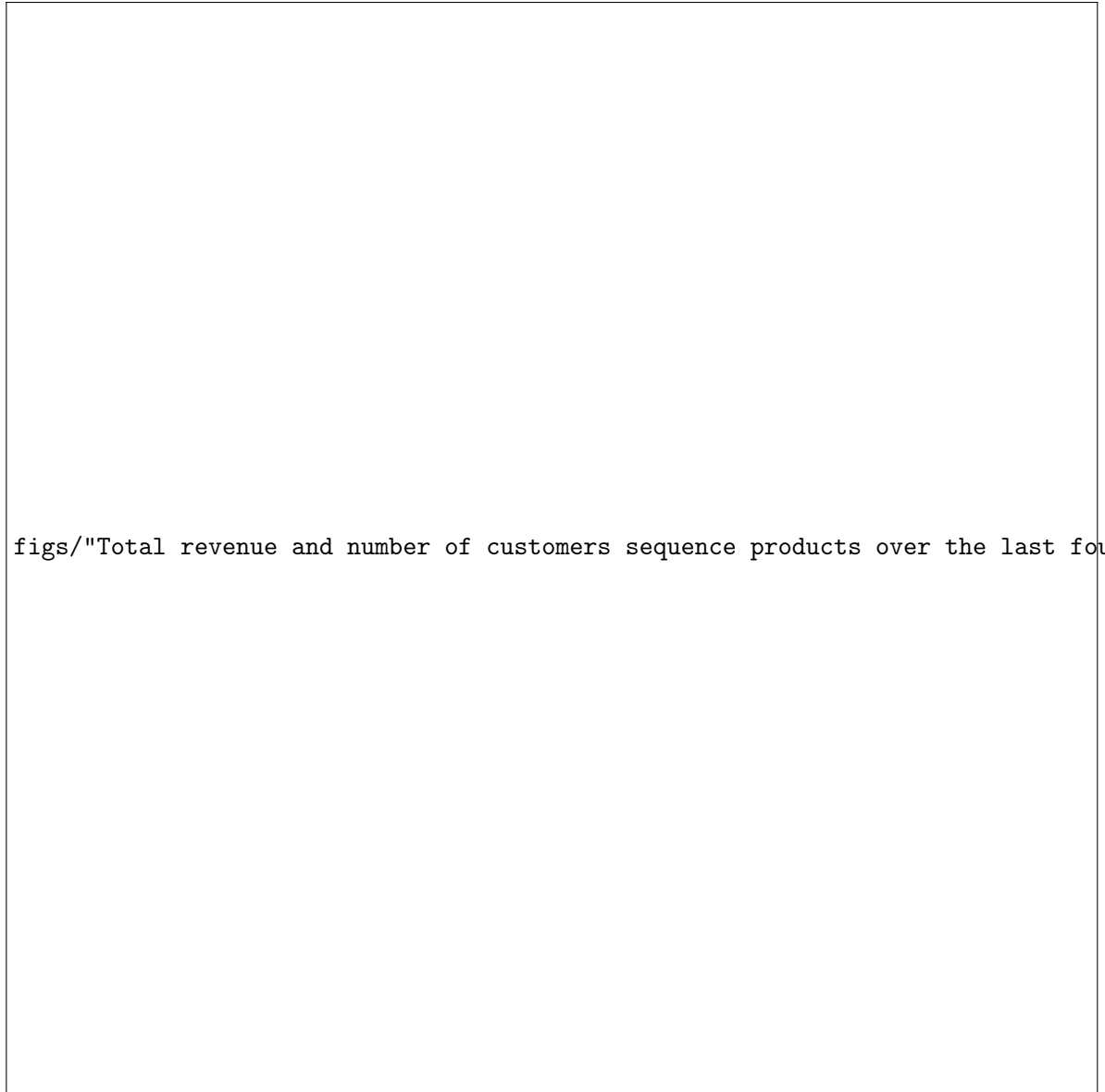
Figure C.1: Total revenue and number of customers sequence products over the last four years (CONFIDENTIAL)

figs/"Total revenue and number of customers special products over the last four years
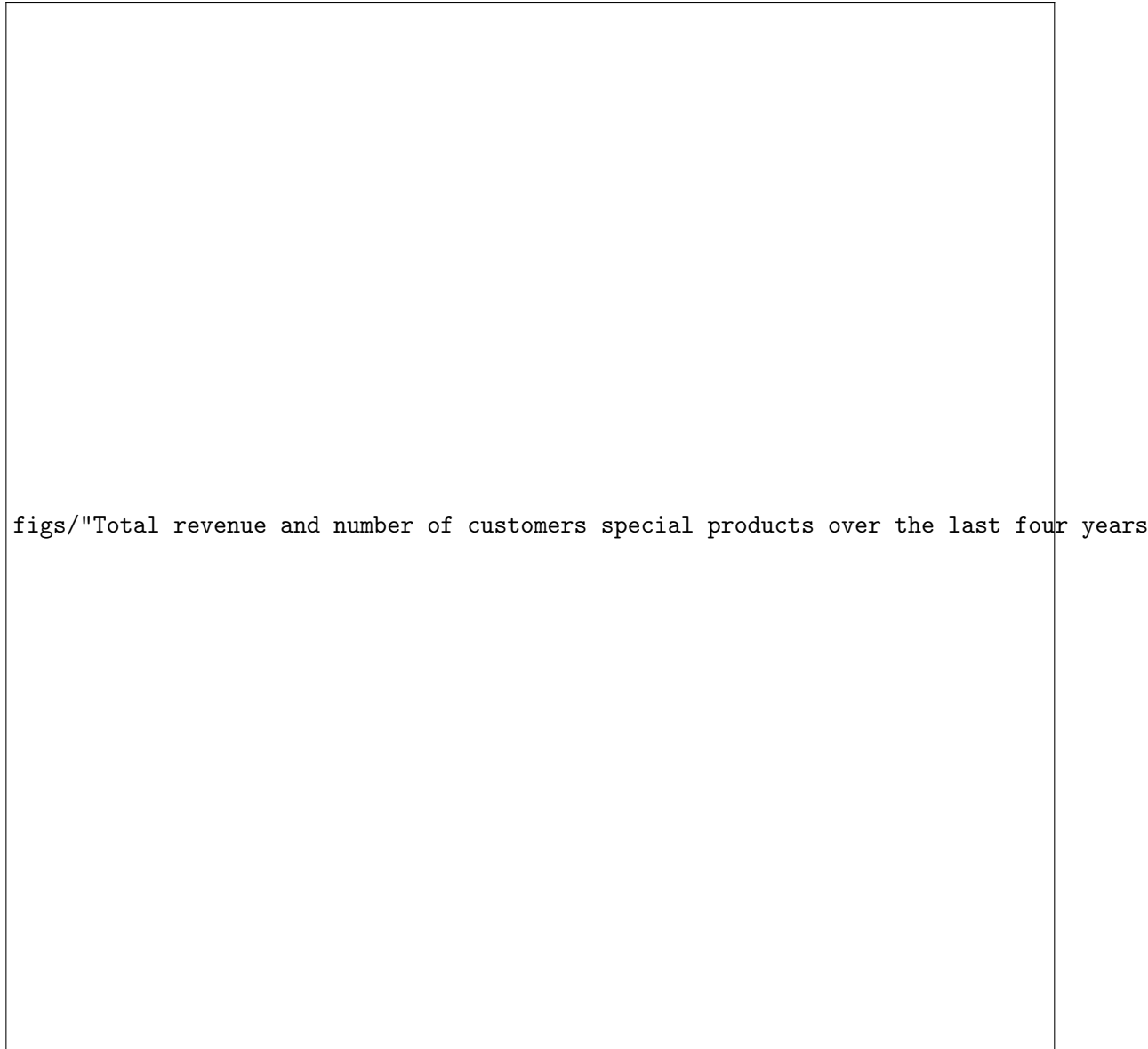
Figure C.2: Total revenue and number of customers special products over the last four years (CONFIDENTIAL)

figs/"Total revenue and number of customers housnumber products over the last four ye
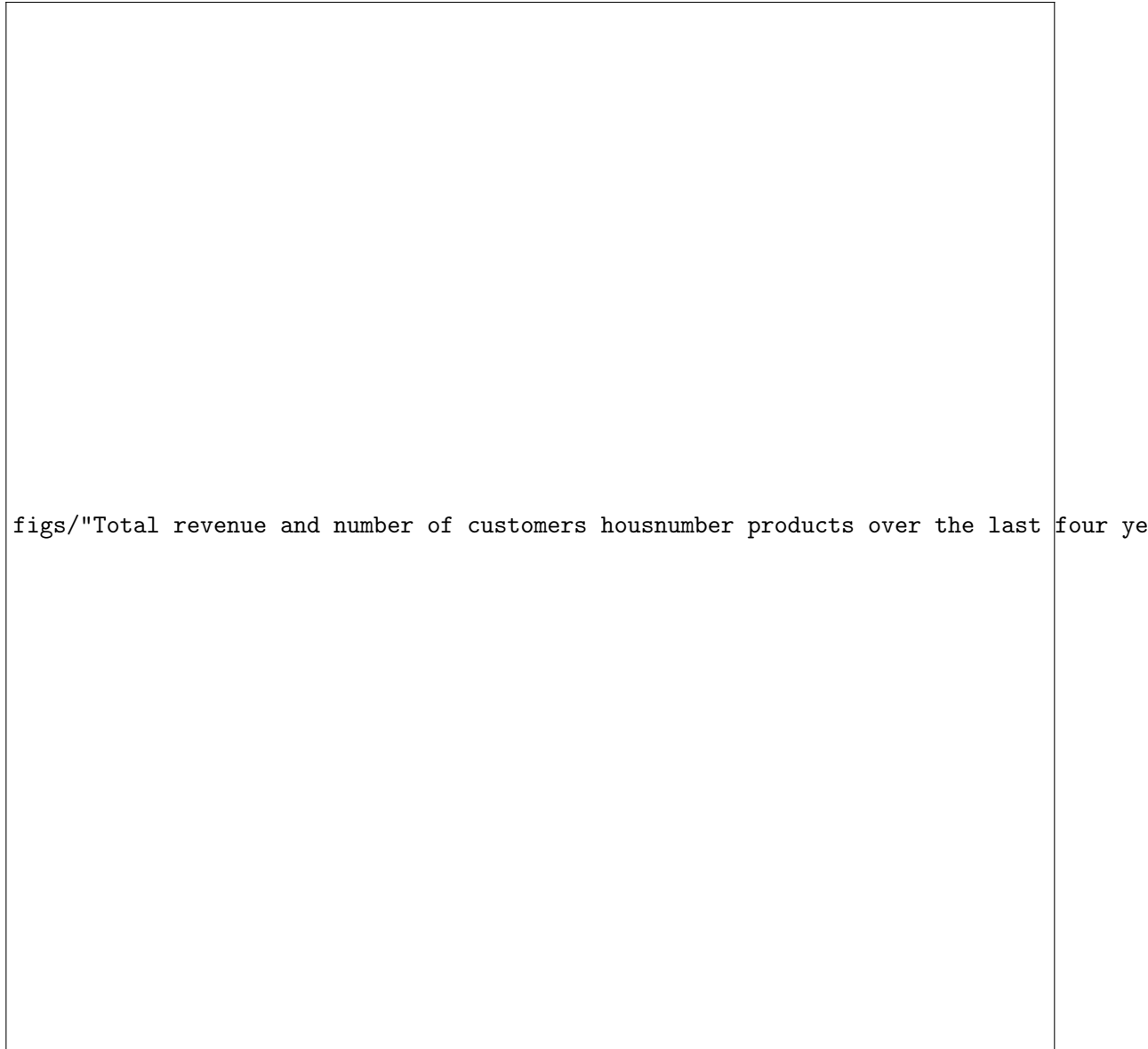
Figure C.3: Total revenue and number of customers housenumber products over the last four years (CONFIDENTIAL)

D | EXPERT PANEL WORKSHOP

In order to test the new system in which open data has a central role, it was decided in this study to organize an expert panel workshop. This contributes on the one hand to testing the potential value of the opened data and on the other hand to testing the operation of the ODIM. The applicability of the ODIM is further evaluated in chaper 6. This appendix explains as briefly as possible how the workshop was set up and what the results were of the various experts who participated in the workshop. In this way the value of the new system and the open data is tested.

D.1 SELECTION OF THE PARTICIPANTS

In view of the possibilities within this study, it was decided to invite the experts who are most involved with the system. The chosen participants, seven in total, knew how the current PCT system works and what the product does. This contributed to the further course of the workshop because it enabled the participants to better imagine the new system. There are two important departments within PostNL Data Solutions that deal with the Postcodetabel: Marketing & Sales and Data & IT. One department simply builds the product and the other sells it. For the workshop, participants from both departments were chosen for the widest possible vision of the system. This prevents a certain way of tunnel vision. In the workshop the participants are divided into two groups with people from a certain department in each group.

D.2 PRACTICAL SETUP

Table D.1 shows the overall planning of the workshop. First of all, a brainstorm session took place based on a general introduction to open data. When the participants knew what open data is and what the possibilities are, the new PCT system could be presented to them. The participants then spent the last part of the workshop on the value assessment of the open data within the new potential PCT system. This assesses the value of the open data and the functioning of the ODIM. The latter because it is tested whether the new system has been established correctly.

Number	Activity	Goal
1	Introduction (12:30 - 12:45)	Explain the purpose of the research and explain the concept of open data. Making sure that the participants not only thought about finances but also about other things
2	Brainstorm value of open data (12:45 - 13:00)	High level brainstorming within which dimensions open data can have value according to the participants. This makes the participants think actively about the positive things about open data, but also about the risk
3	Present findings (13:00 - 13:10)	Learn from the results of the other group
4	Break (13:10 - 13:15)	
5	Introduction open data in the PCT system (13:15 - 13:30)	Presenting the new PCT system with a central role for open data
6	Discussion (13:30 - 13:55)	Detailed brainstorm about the possible value of open data within the PCT within the same five dimensions as in the high level brainstorm. Presenting to the other group
7	Closing (13:55 - 14:00)	

Table D.1: Agenda of the expert panel workshop

D.3 RESULTS OF THE EXPERT PANEL WORKSHOP

The most important and useful results for this study are the results on the data opened up within the new PCT system. The brainstorm, which the workshop started with, was of value to the participants but not to discuss further here. The results per group are explained in more detail in the following sub sections. Each group is asked the three same questions when discussing the PCT system.

D.3.1 Results group 1

This section presents the results of the brainstorming session and the discussion of group 1. The ideas they generated during the workshop are also graphically represented in figure D.1.

People in this group fulfill the following positions within PostNL DS:

1. Director IT, Management team - Data & IT

2. Data manager - Data & IT

3. Product owner - Commerce

4. Solution consultant - Data & IT

WHICH DATA SHOULD YOU OPEN AND WHY   All data that exactly matches the BAG data would be made public. This is a confirmation of the new PCT system presented. We want to make this available free of charge or perhaps a small contribution for questions about opening up the data. This must be worked out in the implementation strategy.

MOST IMPORTANT REQUIREMENT FOR THE NEW SYSTEM   First of all you have to look at the postal code covenant what you can or cannot do with the data. These are agreements made with the government (legal dimension). You must then have a very good self-service portal, look at your pricing model, security and abuse (organizational dimension). In summary, these risks fall under a detailed business strategy if an open data strategy is applied. So there is a need for a business implementation plan.

WHICH PART CONTAINS THE MOST VALUE   We would like to add a reliability feature to the data that has been opened up. For example, if address data is made public, we want to add a feature to it whether you could deliver mail here or not. If the user wants to know why this quality mark has been given, the user must purchase an additional product. Users then see the value of the data, but for deeper insights, they need to purchase an extra product. We therefore see the most value in the economic field because we create a new kind of dependence on the user. By making new products on this open data, we can grow further as a company (economic dimension). In addition, we see a lot of value in the branding of the company. By publishing data that has always been paid for, users will have confidence in the quality (culture dimension). It is important that you show that you make the same data public as the BAG, but better. And if users want more insight, we can offer them, and BAG cannot, for example. Your brand name that you have built up in recent years is very important in this.

D.3.2   Results group 2

This section presents the results of the brainstorming session and the discussion of group 2. The ideas they generated during the workshop are also graphically represented in figure .

People in this group fulfill the following positions within PostNL DS:

1. Data architect - Data & IT

2. Solution consultant - Data & IT

3. Sales manager - Commerce

WHICH DATA SHOULD YOU OPEN AND WHY   We already use different open data sources. Our idea is to collect all these sources and publish them under one format. This therefore also includes BAG data that we already use in our systems to supplement the BRPP data. Group 1 confirms that this was

also their idea what they want to open up. Both groups agree on this point. However, group 2 would make even more data public than just what can be found within the BAG.

MOST IMPORTANT REQUIREMENT FOR THE NEW SYSTEM Group 2 acknowledges adding the attribute to the opened data. This will then really count as a teaser for new users. This is therefore an important condition for the data that has been opened up.

WHICH PART CONTAINS THE MOST VALUE One of the important things you create with open data within the PCT is transparency (culture dimension). By recognizing that the product is decreasing in value and be transparent about this and then publishing the data, you will get a brand upgrade. You are actively concerned with the honesty towards the users of the products of the company. Another major advantage is also in the focus on certain areas that were previously underexposed. You encourage your company to respond to innovation from the outside world, but also to the proper maintenance of your IT environment (technical and organizational dimension).

Waardecreatie voor commerciële bedrijven met behulp van open data
Workshop PostNL DataSolutions
Data has a better idea

Planning

Planning	
12:30 - 12:45	Introductie
12:45 - 13:00	Brainstorm waarde van open data
13:00 - 13:10	Presenteren bevindingen
13:10 - 13:15	Pauze
13:15 - 13:30	Inleiding open data binnen PCT
13:30 - 13:55	Discussie
13:55 - 14:00	Afsluiting

Introductie

- Stage bij PostNL DS in het kader van masterscriptie voor de opleiding Technische Bestuurskunde aan de TU Delft

- Uitdaging in combinatie wensen PostNL en de voorwaarden voor een scriptie

- Inmiddels aantal maanden verder

Afbakening project

- Opdracht geformuleerd rondom PCT

- Aanpassen van de strategie die nu voor de PCT producten wordt gebruikt

- Is een open data strategie hiervoor geschikt?

Doelstelling

- Het ontwerpen van een **methodologie** die inzichtelijk moet maken welke **data** kan worden **opengesteld** en welke **waarde** dit kan creëren voor **commerciële bedrijven**.

- Testen binnen het PCT systeem

Doel van deze workshop

- Informatie over open data en algemene bevindingen van waardecreatie

- Reflecteren op het potentiële nieuwe systeem van PCT

Hiermee validatie op de casus zodat ik later de methode kan valideren

---

Definitie open data

"Open data is data die vrij gebruikt kan worden, hergebruikt kan worden en opnieuw verspreid kan worden door iedereen"

Open data eisen

- *Beschikbaarheid en Toegankelijkheid*
  - Voor iedereen beschikbaar als download bij voorkeur via het internet

- *Hergebruik en herverspreiding*
  - Toestaan van hergebruik en samenvoegen met andere data

- *Universele deelname*
  - Geen discriminatie t.o.v. bepaalde organisaties of individuen

---

3 mythen over open data

1. Open data is slecht voor de business
   - Als je data openbaar zet levert het geen geld meer op, en alles waar de klant niet voor betaald heeft geen waarde

2. Alle data moet open worden gezet
   - Wanneer voor een strategie betreft open data wordt gekozen moet alle data gelijk openbaar worden gezet

3. Open data zorgt alleen voor meer werk
   - Om datasets openbaar te zetten moeten ze worden aangeleverd in een specifiek format wat alleen maar zorgt voor meer werk

---

Succesvol voorbeeld

- **Geolytix**: Startup uit de UK gespecialiseerd in geografische informatie
- Levert **gratis** open data met gegevens uit Londen over onder andere:
  - Al het openbaar vervoer
  - Locatie en specificaties van alle supermarkten
  - Postcode gebieden
  - "Points of interests" (ziekenhuizen, bioscopen, sporthallen, etc)

- Leveren producten aan klanten in de vorm van **betaalde** API's
  - Waar het beste een nieuwe vestiging openen
  - Waarom werkt een strategie in een bepaalde wijk wel en in een ander niet?

- Resultaat: enorme groei in omzet en klanten in de afgelopen paar jaar. Waarde nu geschat op 3,3 miljoen euro
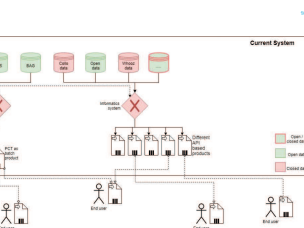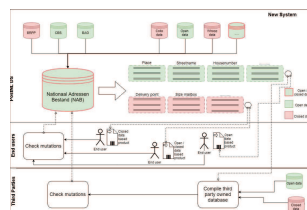

Planning

Planning	
~~12:30 - 12:45~~	~~Introductie~~
12:45 - 13:00	Brainstorm waarde van open data
13:00 - 13:10	Presenteren bevindingen
13:10 - 13:15	Pauze
13:15 - 13:30	Inleiding open data binnen PCT
13:30 - 13:55	Discussie
13:55 - 14:00	Afsluiting

Waarde bepaling van open data

Technisch — Data makkelijk te gebruiken
Wetgeving — Licentie voor gesloten data
Cultuur — Vergroten van de transparantie
Economisch — Stimuleren van innovatie
Organisatie — Vergroten van de efficiëntie

Brainstorm waarde van open data

Bespreek met elkaar en plaats in 1 van de 5 vlakken:

- Voordelen / waarde van open data

- Risico's van open data

Groepsindeling

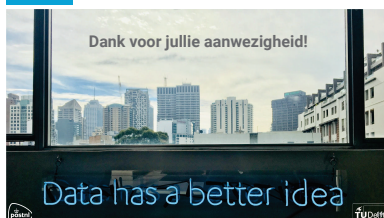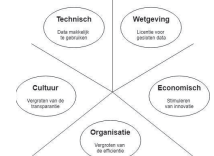| Groep 1	Groep 2
Eric van Duin MT directeur IT	Fred van Velzen architect
Freek Werner proces manager data (data manager)	Job Jansen solution consultant
Marion de Rink - Walter product manager	Ellen Beekenkamp sales manager
John Vernes solution consultant	Ron Bartels

Pauze

Planning	
~~12:30 - 12:45~~	~~Introductie~~
~~12:45 - 13:00~~	~~Brainstorm waarde van open data~~
~~13:00 - 13:10~~	~~Presenteren bevindingen~~
13:10 - 13:15	Pauze
13:15 - 13:30	Inleiding open data binnen PCT
13:30 - 13:55	Discussie
13:55 - 14:00	Afsluiting

Planning

Planning	
~~12:30 - 12:45~~	~~Introductie~~
~~12:45 - 13:00~~	~~Brainstorm waarde van open data~~
~~13:00 - 13:10~~	~~Presenteren bevindingen~~
~~13:10 - 13:15~~	~~Pauze~~
13:15 - 13:30	Inleiding open data binnen PCT
13:30 - 13:55	Discussie
13:55 - 14:00	Afsluiting

Inleiding open data binnen PCT

- **Omzet** van PCT laatste jaren sterk **dalend**

- **Rationalisatie** gedaan op PCT portfolio

- Nieuwe (lange termijn) **strategie** gewenst voor het nieuwe portfolio

---

BAG data en PCT data

- Opkomst van BAG data:
  - Open data variant van adresgegevens en gebouw specificaties

- Veel overeenkomsten met PCT
  - straatnaam, huisnummer, plaats, verblijfsobject, XY coordinaten etc.

- Creatie van twee "waarheden" betreft adresgegevens in Nederland

- Nieuwe strategie → open data

Current System

New System

Verschillen tussen systemen

- Terugkoppeling mogelijk voor gebruikers naar database

- In plaats van batch/API producten nu open/gesloten data gebaseerde API's (de open data variant is nog de huidige PCT batch)

- Ook niet gebruikers van de producten kunnen aanpassingen aandragen

- 1 centrale plek waar data wordt bewaard, "de Datafabriek"

Potentiële waarde van het nieuwe systeem

- Kwaliteit van adresgegevens gaat omhoog door real time aanpassingen aan data
- Kostenbesparing
- Creëren van nieuwe business (data as a service)
- Grote stijging van het aantal klanten
- Positieve branding voor PostNL, koploper op gebied van data
- Verhogen van de transparantie leidt tot betere klantinzichten
- Aantrekken van nieuw talent binnen de organisatie

Planning

Planning	
~~12:30 - 12:45~~	~~Introductie~~
~~12:45 - 13:00~~	~~Brainstorm waarde van open data~~
~~13:00 - 13:10~~	~~Presenteren bevindingen~~
~~13:10 - 13:15~~	~~Pauze~~
~~13:15 - 13:30~~	~~Inleiding open data binnen PCT~~
13:30 - 13:55	Discussie
13:55 - 14:00	Afsluiting

Discussie

- Wat zouden jullie openbaar zetten en waarom?

- Wat zijn voor jullie belangrijke vereisten aan dit nieuwe systeem?

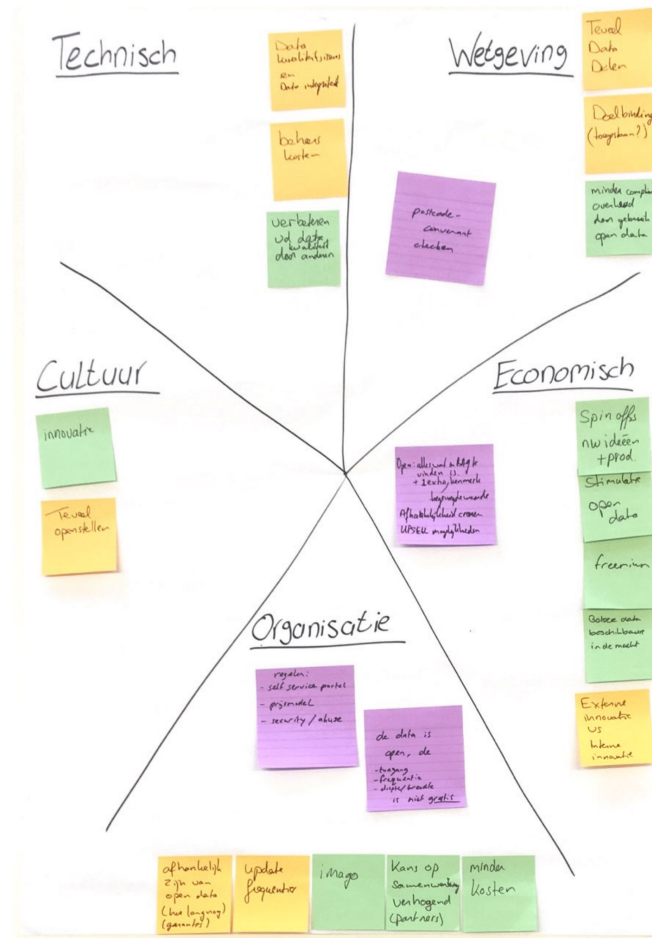- Waar zit de meeste waarde in en waarom? (denk aan de 5 gebieden van waardecreatie)

Afsluiting

Vragen / opmerkingen?

Dank voor jullie aanwezigheid!
Data has a better idea

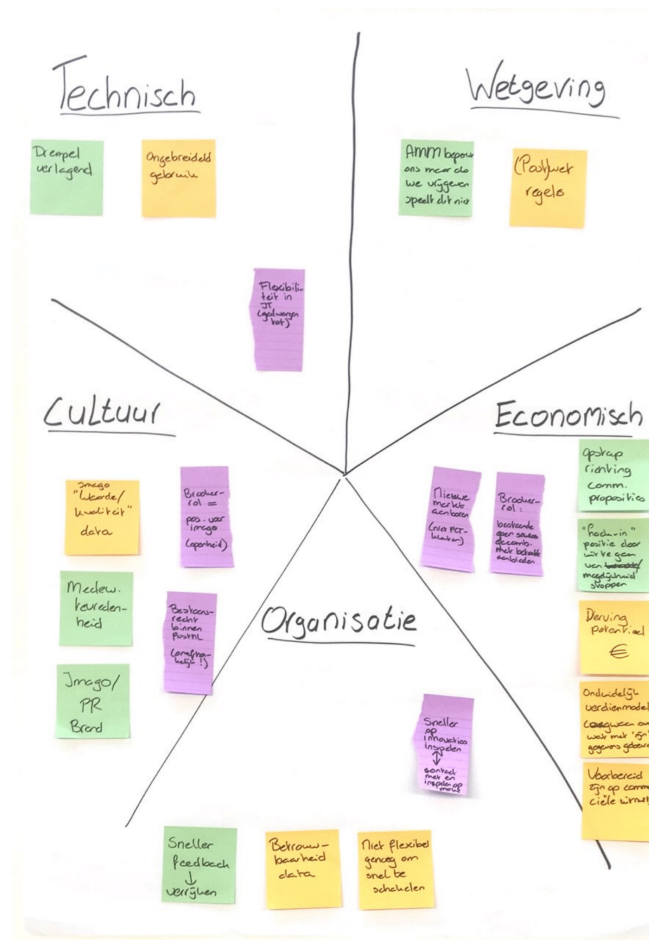**Figure D.1:** Graphic representation of the results of the workshop of group 1

Figure D.2: Graphic representation of the results of the workshop of group 2

E | OPEN DATA SURVEY

Following the workshop on open data within the PCT system, a survey was also conducted among all participants of the workshop. First of all, the questions of the survey are presented in the next section. The results are shown schematically in the section after that. As all participants in the workshop were Dutch speaking, this appendix is also shown in Dutch. The most important results, conclusions and differences between the groups are discussed in chapter 5.

E.1 SURVEY SETUP

Open data strategie PCT voor PostNL Data Solutions

Beste,

Naar aanleiding van de workshop over open data die je hebt gevolgd op 05-07-2019 stuur ik deze vragenlijst. Naast dat we de bevindingen plenair hebben besproken ben ik ook benieuwd naar jouw persoonlijke visie omtrent dit onderwerp. Het invullen duurt slechts een paar minuten en zou erg helpen in het vervolg van het onderzoek. De resultaten zijn anoniem en bij het verzenden van dit formulier stem je automatisch ermee in dat de resultaten kunnen worden meegenomen in de masterscriptie over open data. Deze zal voor iedereen vrij te verkrijgen zijn op de website van de TU Delft.

Nogmaals heel erg bedankt voor je medewerking!

Vraag 1 Onder welke afdeling valt de functie die je binnen DS vervult?

- Commercie
- Data & IT

Vraag 2 Voorafgaand aan de workshop:

- Had ik wel eens gehoord van open data
- Wist ik wat open data was
- Kende ik bedrijven die data openstelde of gebruik maakte van open data
- Wist ik wat voor waarde open data kon opleveren voor een bedrijf

Schaal: volledig mee oneens - oneens - neutraal - eens - volledig mee eens

Vraag 3 Na de workshop:

- Heb ik een beter beeld gekregen van wat open data is
- Vind ik dat open data meerwaarde kan opleveren voor commerciële bedrijven
- Ben ik geïnteresseerd in mij verder verdiepen in de waarde van open data

Schaal: volledig mee oneens - oneens - neutraal - eens - volledig mee eens

Vraag 4 Ik erken de waarde van het openstellen van data voor commerciële bedrijven op het gebied van:

- Technologie
- Wetgeving
- Cultuur
- Economisch
- Organisatie

Schaal: helemaal geen meerwaarde - geen meerwaarde - neutraal - wel meerwaarde - heel veel meerwaarde - geen mening

Vraag 5 Door bepaalde datasets binnen de PCT open te stellen zie ik in dat:

- De kwaliteit van deze gegevens omhoog gaat
- Er kostenbesparing kan optreden
- Er veel mogelijkheden zijn voor nieuwe business
- Het aantal klanten kan stijgen
- Dit positieve branding is voor PostNL
- Dit leidt tot betere klantinzichten door transparanter te zijn
- Dit leidt tot het aantrekken van nieuwe talentvolle werknemers

Schaal: volledig mee oneens - oneens - neutraal - eens - volledig mee eens - geen mening

Vraag 6 Heb je verder nog vragen, opmerkingen of tips?

E.2 RESULTS

Six out of the seven participants in the workshop responded to the survey. The results are listed below.

Onder welke afdeling valt de functie die je binnen DS vervult?
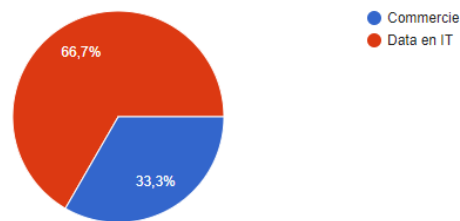
6 antwoorden

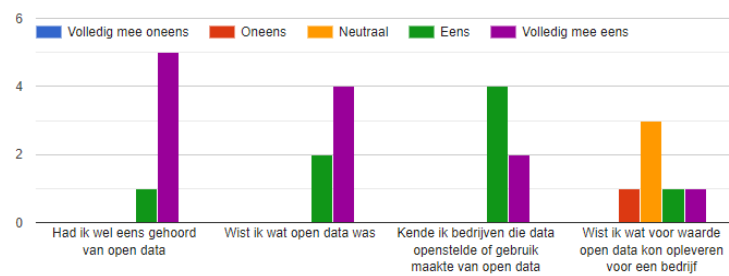**Figure E.1:** Results survey question 1

Voorafgaand aan de workshop:

Figure E.2: Results survey question 2

Na de workshop:

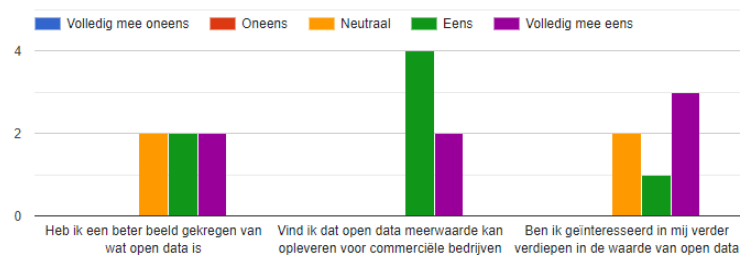**Figure E.3:** Results survey question 3

Ik erken de waarde van het openstellen van data voor commerciële bedrijven op het gebied van:
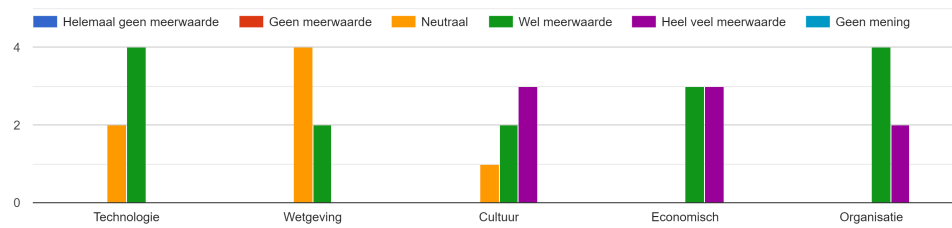

Figure E.4: Results survey question 4

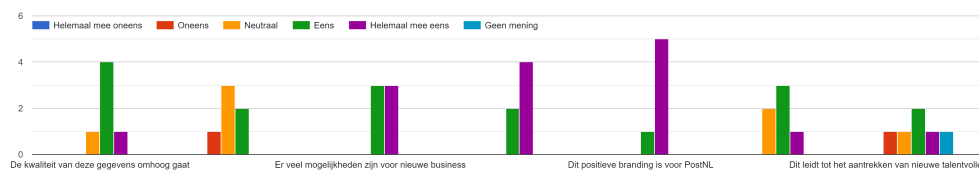Door bepaalde datasets binnen de PCT open te stellen zie ik in dat:

Figure E.5: Results survey question 5

Figure E.6: Results survey question 6

F | OPEN DATA BUSINESS MODELS EXPLAINED

One of the shortcomings in this study is a concrete implementation plan. Because the new system of the PCT has only been tested for potential value, and given the scope of this research, this has not been included in this study. However, the interviews and in particular the expert panel workshop showed that there is a great need for this. An implementation plan ensures that it will become much clearer for the employees who work with the data that has been opened up how the system will take shape. Since the case, and the rest of the research, is about open data at commercial companies, concrete business models are desirable in an implementation plan. The research of Zeleti and Ojo (2018) extensively examines the possible business models that can be used for open data projects.

The case of the PCT showed that a type of freemium business model can be a desired model. In particular, giving a teaser to the potential users of the end product can lead to other new users. The most important elements of value assessment according to Zeleti and Ojo (2018) are shown in figure F.1.
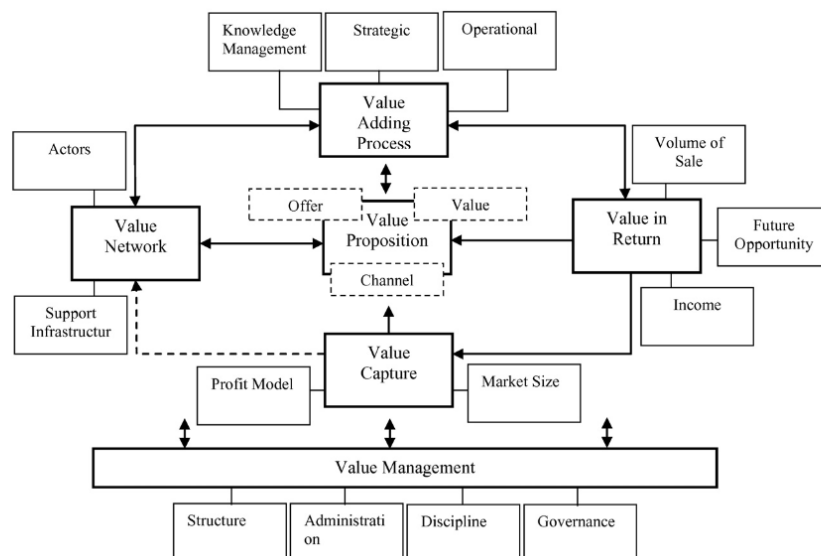
Figure F.1: The 6-V business model conceptual framework adapted from Zeleti and Ojo (2018)

G | SCIENTIFIC ARTICLE

See next page for the scientific article associated with this research.

Creating value through opening up data by commercial companies: Evaluation of the Open Data Implementation Methodology

Michael van der Wielen[a]

[a]*Submitted in partial fulfillment of the requirements for the degree of Master of Science in Complex Systems Engineering and Management*

ABSTRACT

Research has already shown that there is a lot of potential for open data. However, until now it has been mainly government organizations that use open datasets or governments that open up the data. To utilize the full potential, companies must open up more data. To show the value of the open data for companies, the open data implementation method was developed in the study. By using this method, companies gain insight into which datasets within their current systems are suitable for being opened up. Testing this method at a commercial company has shown that open data can add a lot of value in economic, social and organizational areas. More research is needed in applying ODIM to other companies and the risks of opening up data.

1. Introduction

Open data initiatives have received increasing attention in the last few years. Open data is becoming an increasingly well-known subject in science, government, companies and citizens. The reason for this broad interest in recent years is twofold. First of all, the growth in the availability of data has increased enormously [17]. More and more organizations have access to data, which means that the interest in the possibilities with this data only increases [3]. The second reason for the growing interest is the demonstrated potential of open data projects. Several studies have shown that there is a lot of (economic) potential when more organizations use open data or when more organizations make their data openly available [14, 2]. In recent years, more and more successful projects have been found in which open data had a crucial role.

The potential of the open data that has been demonstrated can be used when more open data is available. And it is precisely this availability of open data that poses a major problem in the current design of the data landscape [9]. This can largely be explained by the fact that at the moment mainly government organizations open their data or make use of open data sources when executing projects. These organizations have various reasons for this, such as increasing transparency within society. To be able to use the full potential of open data, companies also have to contribute to this. This can be done by opening up data sources them self or by using open data. This research focuses on the first, opening up current closed data sets at commercial companies. There are various reasons why companies do not open their data. Companies often think that open data is an unnecessary insight into conventional data and that opening up data leads to large losses in terms of market position and annual turnover [7].

This shows that there is currently too little insight for companies into the potential value of open data. When companies start to see this value, there is a chance that the full potential of open data will be exploited. With this, current data-driven ecosystems can be arranged differently and processes and projects can be organized more efficiently [11].

In a world where more and more data is available and great potential has been demonstrated with regard to open data, companies need to invest more in this. This research focuses on the value creation of open data for companies. This value is currently not being realized, so this requires tools at companies to stimulate the development of the open data environment.

The following chapter 2 will describe the methodology with which the main research question formulated in this paper will be answered. This question is: *How can open data contribute in value creation or digital products or services for commercial companies?*. Before this can be described in the next chapter, it is useful to provide a number of definitions for the remainder of this paper. Data is considered open data when the data is: complete, primary, timely, accessible , machine process-able, made online in persistent archives and as accurate as possible. Besides this the data access must be non-discriminatory, the data formats must be non proprietary data and the license must be unrestricted and bear no usage costs [4, 13]. When writing about open data in the remainder of this paper, data means what meets all the above characteristics. In addition, this research focuses on the value of open data at commercial companies. These are companies with a profit motive and act in the interest of the customers they serve. When writing about companies or businesses in this paper, only commercial companies are meant.

After the methodology is described in the next chapter, chapter 3 follows an analysis of the current literature on open data. This describes the most important benchmarks of open data for companies. Chapter 4 describes the development of the open data implementation methodology that can be applied at companies to gain insight into open data within their current business systems. This methodology is tested in chapter 5 at a commercial company in the Netherlands, PostNL Data Solutions. Here, among other things, it becomes clear through an expert panel workshop what value a company can create with open data. The conclusions and future recommendations of this entire study follow in chapter

7.

2. Research Approach

In earlier research, the development of an open data implementation method investigated the value of open data for companies [20]. In this research, the focus was on developing and testing this method and the results could be traced to the implication of analyzing current business systems and how open data plays a central role in this. This research builds on this by going deeper into the benchmarks used to determine the value of potentially open data at commercial companies. The Design Science Research Methodology (DSRM) was used for this research. In the available literature much has been written about different ways in which this methodology can be applied in scientific research. Design science is concerned with "producing and applying knowledge of tasks or situations in order to create effective artefacts "[16]. The artefact that was developed in the previous study into the value of open data at companies is the aforementioned open data implementation methodology (this study will now be referred to as ODIM). DSRM consists of six different design process steps: 1) identify the problem, 2) define objectives of a solution, 3) design and development, 4) demonstration, 5) evaluation and 6) communication [8, 18]. By using an expert panel workshop in the demonstration phase of the aforementioned design science approach, insights are gained into the value of open data at commercial companies. This research is an enrichment of the available literature on open data and the earlier research by deepening the benchmarks of value creation. These benchmarks, so-called value creation dimensions, are explained in the next chapter and tested in chapter 5.

3. Background Literature

An extended literature study into open data has confirmed what has already emerged in the introduction of this paper. Currently, open data projects are mainly implemented by government organizations or it is governments that make data available openly. In addition to the aforementioned reason for more transparency within society, this also has to do with the fact that open data research is currently taking place mainly in countries with highly developed economies [12, 15]. The most important part of this paper of the extended literature review is the analysis of the effects of open data. By researching these effects in the past, well-founded statements can be made about the potential value that open data can offer. By conducting research into both the effect of open data at government organizations and at companies, a framework can be established for value evaluation within the open data domain. This framework is used in the following chapter within the open data implementation methodology. By putting the effects of earlier research in a framework, newly designed systems can be tested at companies, including open data.

A number of examples that emerge in the literature concerning the effects of open data are: for example, new markets can be opened and innovative IT solutions are made possible [6]. Also the generation of social, environmental and economic value when using open data is stimulated [10]. Users are sometimes not aware of the possibilities that developers create with the open data by developing various APIs. In addition to the possibilities that the APIs can create for users, the needs also differ per organization. For example research has demonstrated that in open data research it is difficult to draw conclusions that can be generalized [19].

Because these needs can differ per organization, a framework must be developed for this to be able to compare open data implementation methods and to measure the value of the open data. Research has led to the following five key dimensions for value creation with open data within enterprise systems [1, 19]:

- Technical dimension: value creation within this dimension is often in the dataset itself. Because open data must comply with a certain format (machine processable and non-proprietary), it must make it easier for users to use the data. In addition, the data ambiguity is also reduced because the format in which the data is delivered makes the data set easier to understand.

- Policy / Legal dimension: to make datasets not open but closed, a lot has to be arranged. Licenses must be requested and kept up to date. Certain laws and regulations can also ensure that valuable data sets are not used at all and therefore do not create value.

- Economic / Financial dimension: this dimension can form a barrier if no money is made available within the company to identify which dataset can be made public. But in this dimension, the benefits, and therefore the value that can be created, lie mainly in the innovations that are possible. By opening up data, companies can build new data products with this. In addition, it also offers the company value by selling products that work with the data that is opened up. By opening up the data, it becomes clear to users what the dataset consists of. However, this does not mean that the user can actually do something with it. (Economic) value can be achieved here by companies and other organizations.

- Organizational dimension: the value creation within this dimension focuses primarily on increasing efficiency within the organization. By opening certain datasets, users can detect errors and adjust them themselves or make a proposal to adjust them. By partially outsourcing this task to the users, the data set will be more accurate and therefore provide more value for the users.

- Social / Cultural dimension: this dimension regards the feeling of the public towards open data in your organization. An important factor is the transparency of the company. It could reduce corruption by enabling greater transparency. It also promotes participation of

citizens and other organizations by providing more insight into the data. The accessibility and openness of the company is improved by being more transparent, which can provide added value for the company.

4. Open data implementation methodology

The important five dimensions of value creation for open data have been included in the development of the artefact. In the design science approach this is the next phase of the research and serves to ultimately find a solution for the identified problem. Where the problem in this study refers to the lack of commercial companies that see the value of open data, the artefact (the open data implementation methodology - ODIM) helps companies to see this value. Secondly, this method helps to analyze and critically view current data systems. By performing this analysis it becomes clear how the processes work in the current system and how that can be adjusted to make it suitable for open data.

As the implementation of open data within current systems is accompanied by the creation of a new design, use of a system design approach is desirable in this research. To logically follow the steps described in the ODIM, it was decided to use the system engineering design methodology to structure the processes in the ODIM [5]. This approach has been chosen in particular because of the iterative nature of the method. This also had to occur in the design of the ODIM. This is mainly due to the mid-term evaluation of the system that is required for companies. For example, when companies do not recognize the value of an open data set, or want to decide not to apply an open data strategy, they can go back to an earlier phase in the method. The methodolgy is ultimately about giving companies insight into which data can be made public within their current systems and which value creates this for the company. The ODIM consists of the following five phases, which are also shown schematically in figure 1:

- **Phase I - Exploration**
  The first phase of the ODIM consists of detailed mapping of all current systems within the company. By compiling process models, good insight can be obtained into the data that is suitable for opening up. This in combination with an extensive analysis of, for example, the current costs and revenues of certain data sets, provides a great insight into the implementation of open data for companies. This in combination with the requirements for open data and an analysis for which users in the current system use the data yields results for phase two.

- **Phase II - Conceptual**
  In phase two the design with the open data can be compiled. This also includes the important value creation dimensions that were identified in the previous chapter, in the literature background. By linking these value dimensions to the new design with open data, the value of this data can be tested.

- **Phase III - Testing**
  The third phase is the actual testing of the system. By allowing users and / or developers to determine the value of the open data in the system, well-founded conclusions can be drawn about the entire open data implementation process.

- **Phase IV - Evaluating**
  After it has become clear from the previous phase what the most important elements are within the concept design, the final design can be compiled in phase four. Any proposed changes are incorporated in this to create even more value with the data that is potentially opened up.

- **Phase V - Communication**
  In phase five, the final design is communicated to the person developing it. This includes a detailed communication plan and possibly a corresponding business model to maximize the value of the open data [21].

This conceptual design of the open data implementation methodology can now be applied to a real life case to test whether the method can identify value for open data. And if that is the case, within which dimensions.

5. Case study analysis

The artefact that was designed in the previous chapter in the design science approach, the open data implementation methodology, is applied in this chapter to the case of the Postcodetabel at the commercial company PostNL Data Solutions in the Netherlands. By applying this method, current systems at PostNL are analyzed and assessed whether an open data strategy would provide value to this company. Below is first an explanation of the case, after which the effect of each phase within the ODIM on the PCT case at PostNL is described. These are the most valuable results of this paper.

Due to the growth in the parcels market and the decrease in the mail market, the interest of (e-commerce) companies regarding postal code and address data is also changing. Even more than before, it is important for these companies to have the correct data. This ensures less return of parcels, which contributes to a better environment and a reduction in costs. PostNL has set up its own department to assist companies in this: PostNL Data Solutions (PostNL DS). This department makes digital products for customers. The establishment of this department within PostNL also has to do with the major changes that have taken place in the "digital world". Whereas previously households and companies were sent a red book with all the correct address and postal code data, everyone now has much easier access via the internet. This red book was replaced from the rise of the internet (around 25 to 30 years ago) by (to this day) the largest product (in terms of sales and turnover) within this department: the Postcodetabel (PCT). The PCT is sold to companies with the aim that these companies always have the correct postal code and address data of their current and new customers. By providing a
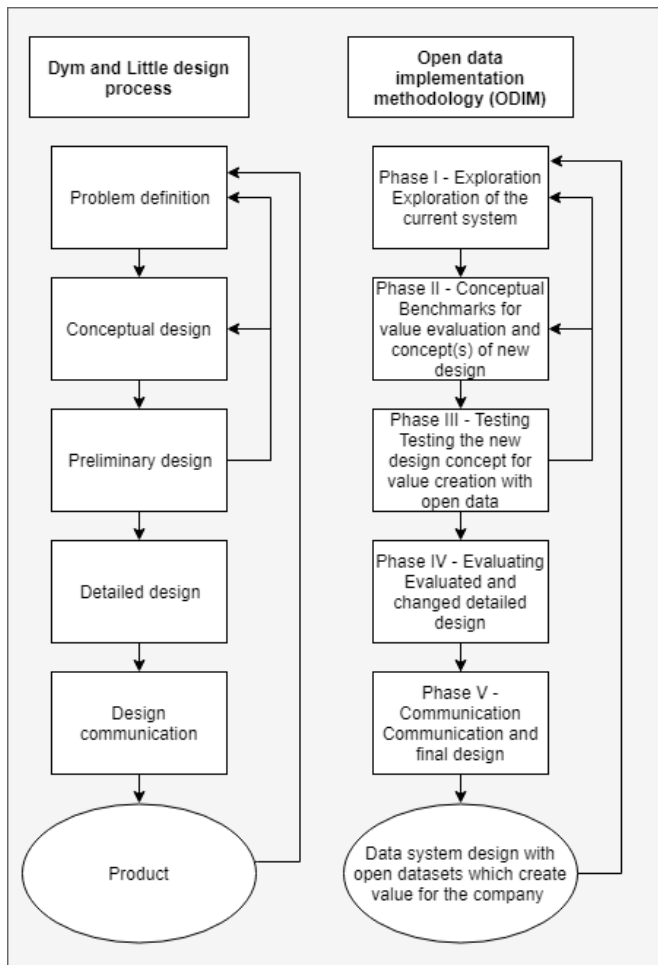
Figure 1: Open data implementation methodology

Figure 2: Results survey after conducted workshop in the case study

dataset in which all names of the cities are correctly linked to other properties such as the postal code, street name and house number, value is created with this product for customers.

Because the value of this product has fallen sharply in recent years and the number of users who have purchased the product, it is a good case to apply the ODIM to. Below is a summary of the results of each phase.

- **Results Phase I**
  The current system shows that PostNL DS currently uses many closed and open data sources within their systems. All these data sources are merged into a central database. Products are then made from this and sold to end users. Within this system, to the end users, there is no feedback mechanism and only closed data is delivered.

- **Results Phase II**
  A conceptual design is made from a new system containing open data and a feedback cycle. Users can respond via an interface if there are errors in the dataset. In this way, an attempt is made to add value to this new system.
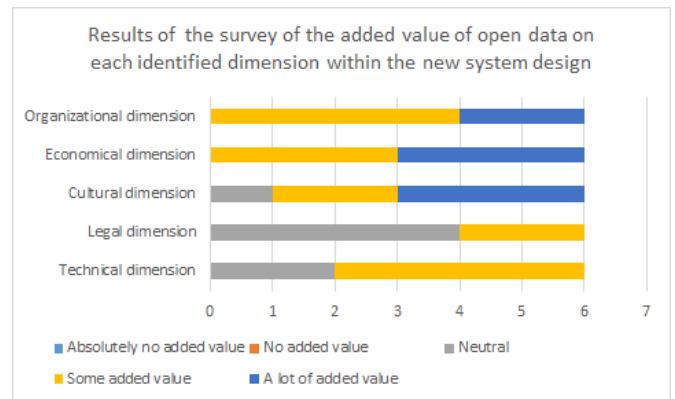
- **Results Phase III**
  In the third phase, the new system was tested for the value achieved with the open data. Results from this were positive. From an expert panel workshop, held with company employees, the feedback came that more data could possibly be opened up to add even more value to the system. The participants found the most value within the economical, organizational and social dimension. By adding open data, this stimulates innovation, leads to more transparency of the company and better organized IT systems. In this case, it has been demonstrated that implementing an open data strategy can provide much value for the company. The confirmation of the value creation of the new system in the different dimensions is shown in figure 2

- **Results Phase IV**
  After processing the adjustments that came out of the workshop, the final design has been put together.

- **Results Phase V**
  Within the communication phase there is an implementation plan that is outside the scope of this research. Further research into this is needed.

Based on this case, it can be concluded that open data adds value to existing systems of the company. Employees acknowledge seeing great value from the open data. However, the benchmarks drawn from the literature were too restrictive. People may have looked too positively at the implementation of open data due to the positive approach of the researcher. The ODIM was later adjusted in the evaluation phase by allowing the companies to determine the value of the open data for themselves by adding company specific benchmarks.

6. Evaluation

The results of this research, as described in the previous chapter, have shown how data can be opened up at a commercial company. The application of the open data implementation method has demonstrated that this method offers the

possibility of properly mapping the potential value of open data. However, there is an important part that is not included well enough in this study.

By conducting an expert panel workshop with the employees of the company where the data is potentially opened up, an attempt has been made to assess the value of open data within the new system. With this workshop, an attempt was also made in this study to evaluate the open data implementation method in its entirety. Because the system was assessed, statements can be made in a certain way about the operation of the method as a whole. This is a shortcoming in this study because both the value of the open data and the working of the method have now been tested with the same people. In order to be able to make better, more informed statements about the broad applicability of the ODIM, it must be evaluated in other ways. Given the scope and time frame of this research, this has not yet been done, but this would be a good addition to the whole study. Below are three examples of how this evaluation of the method could be done. If this is carried out in a subsequent study, this could make a substantial contribution to the future research into opening up data at commercial companies and perhaps the more efficient organization of many current business systems.

- A first possibility to better evaluate the open data implementation methodology is by presenting the method to open data experts. In the current state, the method is only assessed by the researchers themselves and employees of a commercial company. Important new insights can be obtained by discussing the method with an open data expert before applying it to another case. Open data experts may add new benchmarks for the value assessment of the opened data or a different view or the feedback cycle that is now incorporated in the method.

- A second method for evaluating the ODIM is to perform a simulation. By modeling a current business system with closed data, the value of that system can be objectively assessed. By subsequently modeling the new system, including the open data, a comparison can be drawn between the KPIs that have been drawn up. When it appears that the new system performs better than the old system, this is a confirmation of the way in which the ODIM is set up.

- A third and final way to conduct the evaluation is to do a pilot. The complete current structure of the method is based solely on a potentially new system with open data. Data has never actually been opened up. By designing the new system in such a way that parts of the data can be opened up in the beginning, the value of the opened data can be measured. In this way, the feedback cycle in the ODIM is also looked at more critically. If it appears that certain parts of the data are not suitable, it could be decided at an early stage of the process to take a different approach.

These are all examples of a different way of evaluating

the open data implementation methodology. Further research is needed to make the ODIM applicable within many other types of commercial companies. Starting the evaluation is a good first step in this.

7. Conclusions

Various studies have shown that open data has a lot of potential, both at government organizations and at companies. This value needs to be recognized by the companies for which a methodology has been developed in this study. This method helps companies to analyze existing processes and make it open data ready. The ODIM applied in this study to a PostNL case has proven to be able to create value by identifying why value can be achieved with open data within the company. Because the ODIM was applied to one case only in this study, it is difficult to draw conclusions about the value of the ODIM in other sectors. The most important conclusion is that parts from the ODIM, such as the requirements of open data and the process models of current systems, are useful for also applying to other cases. This is a way to find out in a structured way how a company creates value with open data. Further research into value creation using open data at companies is necessary to be able to make more statements about this. The ODIM must also be carried out at several companies and the risks of open data, such as the GDPR legislation, which has not been widely investigated in this study, must be taken into account. This research provides a lot of insight for both science and society what open data can create for value and how companies can identify this.

References

[1] Attard, J., Orlandi, F., Auer, S., 2016. Value Creation on Open Government Data, in: 2016 49th Hawaii International Conference on System Sciences (HICSS), IEEE. pp. 2605–2614. URL: http://ieeexplore.ieee.org/document/7427509/, doi:10.1109/HICSS.2016.326.

[2] Berends, J., Carrara, W., Radu, C., 2017. The Economic Benefits of Open Data. Technical Report. URL: http://www.europeandataportal.eu/.

[3] Borgman, J., 2017. InsideBIGDATA Guide to Use of Big Data on an Industrial Scale URL: https://insidebigdata.com/white-paper/guide-big-data-industrial-scale/.

[4] Charalabidis, Y., Zuiderwijk, A., Alexopoulos, C., Janssen, M., Lampoltshammer, T., Ferro, E., 2018. The World of Open Data. volume 28 of *Public Administration and Information Technology*. Springer International Publishing, Cham. URL: http://link.springer.com/10.1007/978-3-319-90850-2, doi:10.1007/978-3-319-90850-2.

[5] Dym, C.L., Little, P., Orwin, E.J., Spjut, R.E., 2008. Engineering design : a project-based introduction. Wiley.

[6] Frankowski, A., Van Der Steen, M., Meijer, A., 2015. Vorm geven aan inhoud Mark van Twist De Publieke Waarde(n) van Open Data. Technical Report. URL: https://www.open-overheid.nl/wp-content/uploads/2017/04/NSOB-15-17-DT_Publieke-Waarde-web-DEF.pdf.

[7] Hammell, R., Bates, C., Lewis, H., Perricos, C., Brett, L., Branch, D., 2012. Open data Driving growth, ingenuity and innovation Contents. Technical Report.

[8] Hevner, A., Chatterjee, S., 2010. Design Science Research in Information Systems, Springer, Boston, MA, pp. 9–22. URL: http://link.springer.com/10.1007/978-1-4419-5653-8_2, doi:10.1007/978-1-4419-5653-8{_}2.

[9] Janssen, M., Charalabidis, Y., Zuiderwijk, A., 2012. Benefits, Adoption Barriers and Myths of Open Data and Open Government. In-

formation Systems Management 29, 258–268. URL: http://www.tandfonline.com/doi/abs/10.1080/10580530.2012.716740, doi:10.1080/10580530.2012.716740.

[10] Jetzek, T., Avital, M., Bjørn-Andersen, N., 2014. Generating Sustainable Value from Open Data in a Sharing Society, Springer, Berlin, Heidelberg, pp. 62–82. URL: http://link.springer.com/10.1007/978-3-662-43459-8_5, doi:10.1007/978-3-662-43459-8{_}5.

[11] Kitchin, R., 2014. The data revolution: Big data, open data, data infrastructures and their consequences. Sage.

[12] Lakomaa, E., Kallberg, J., 2013. Open Data as a Foundation for Innovation: The Enabling Effect of Free Public Sector Information for Entrepreneurs. IEEE Access 1, 558–563. URL: http://ieeexplore.ieee.org/document/6584732/, doi:10.1109/ACCESS.2013.2279164.

[13] Malamud, C., OâĂŹReilly, T., Elin, G., Sifry, M., Holovaty, A., OâĂŹNeil, D., Migurski, M., Allen, S., Tauberer, J., Lessig, L., 2013. Principles of Open Government Data. Open Government Working Group URL: http://www.opengovdata.org/home/8principles.

[14] Manyika, J., Chui, M., Groves, P., Farrell, D., Van Kuiken, S., Doshi, E.A., 2013. Open data: Unlocking innovation and performance with liquid information. McKinsey Global Institute 21.

[15] Maramieri, J., 2014. Journal of e-Learning and Knowledge Society.. volume 10. Italian e-Learning Association. URL: https://www.learntechlib.org/p/148271/.

[16] March, S.T., Smith, G.F., 1995. Design and natural science research on information technology. Decision Support Systems 15, 251–266. URL: https://www.sciencedirect.com/science/article/pii/0167923694000412, doi:10.1016/0167-9236(94)00041-2.

[17] van Ooijen, C., 2014. Het risico van de mobiele burger : Publieke... (2014) | www.narcis.nl. Technical Report. URL: https://www.narcis.nl/publication/RecordID/oai:tilburguniversity.edu:publications%2Ffebb9414-a8aa-47d0-b103-6ed88f8ae22f.

[18] Peffers, K., Tuunanen, T., Rothenberger, M.A., Chatterjee, S., 2007. A Design Science Research Methodology for Information Systems Research. Journal of Management Information Systems 24, 45–77. URL: https://www.tandfonline.com/doi/full/10.2753/MIS0742-1222240302, doi:10.2753/MIS0742-1222240302.

[19] Susha, I., Zuiderwijk, A., Janssen, M., Grönlund, Ã., 2015. Benchmarks for Evaluating the Progress of Open Data Adoption. Social Science Computer Review 33, 613–630. URL: http://journals.sagepub.com/doi/10.1177/0894439314560852, doi:10.1177/0894439314560852.

[20] van der Wielen, M., 2019. Creating value through opening up data at commercial companies: A case study at PostNL Data Solutions. Ph.D. thesis.

[21] Zeleti, F.A., Ojo, A., 2018. Capability Model for Open Data, in: Proceedings of the 11th International Conference on Theory and Practice of Electronic Governance - ICEGOV '18, ACM Press, New York, New York, USA. pp. 403–411. URL: http://dl.acm.org/citation.cfm?doid=3209415.3209492, doi:10.1145/3209415.3209492.