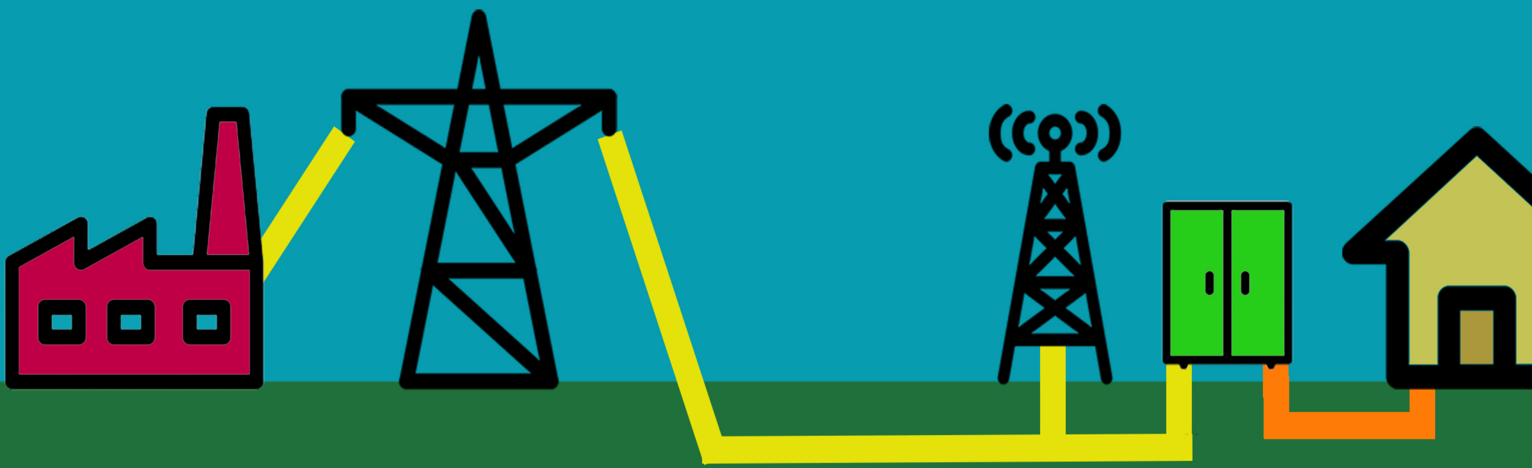
Policy recommendations to improve data quality in the electricity chain

Martijn Pronk

Master of Science Thesis



Delft University of Technology

Complex Systems Engineering and Management

Policy recommendations to improve data quality in the electricity chain

Master of Science Thesis

For the degree of Master of Science in Complex Systems Engineering and Management (CoSEM) at Delft University of Technology

Graduation committee:
Chair of the committee: Prof.dr.ir. M.F.W.H.A. (Marijn) Janssen
First supervisor: Dr.ir. B.(Bert) Enserink
Second supervisor: Mw.drs. J. (Jolien) Ubacht
Third supervisor: Dr. A.M.G.(Anneke) Zuiderwijk-Van Eijk
KPN supervisor: J. (Jikke) op de Weegh

Student: Martijn Pronk(4159314)

Graduation date: November 30, 2017

Faculty of Technology, Policy and Management (TPM) · Delft University of Technology

The work in this thesis was supported by KPN. Their cooperation is hereby gratefully acknowledged.

Abstract

Data quality is essential in the modern world. The higher the quality of the data, the more information, knowledge and thereby wisdom can be retrieved. Besides better wisdom high quality data has another advantage: it can save costs since errors can be avoided. Low quality data is costing a lot of money to companies and will lead to information with more noise. The energy sector is also an industry where data is of great importance. Data quality is of importance in the electricity chain since a lot of parties depend on it. The greatest challenge emerging from this energy transition is the balance between decentralized electricity production and constant electricity consumption. However the data in the electricity chain is not always of great quality leading to problems. The main research question is: **Which measures can be taken to improve data quality in the electricity data chain?**

First, a literature review was conducted to identify the dimensions of data quality. This was done to answer the sub-question: **How can data quality be expressed in a chain?** Four indicators were found to be most recurring for expressing data quality: accuracy, timeliness, consistency and completeness. These dimensions are popular in literature but can not always be easily applied to data quality problems in chains. Four additional dimensions for data quality were selected to measure data quality in a chain: reliability, comparability, accessibility and interoperability.

In order to answer the research question a case study was conducted to investigate data quality problems. The case study in the electricity chain in the Netherlands shows the complexity when data is being used by multiple organisations. The research was conducted at KPN since they are a big user of electricity in the Netherlands. KPN is therefore also dealing a lot with the complexities in the chain since KPN has to manage thousands of connections. In the energy sector data quality is key since the stability of the grid is dependent on data. Based upon the electricity usage data electricity is produced by producers. When electricity usage data is incorrect this has consequences for the total use. When too much or too little electricity is consumed this could lead to a surplus or a deficit on the electricity grid.

As a part of the case study the energy sector was analysed through interviews, desk research and available data on electricity connections. The electricity chain was analysed in order to get a clear understanding of processes and data. Multiple parties are working together

in the electricity chain to measure Metering company (Meetverantwoordelijke) (MV), transport Distribution System Operators (Netbeheerders) (DSO), Transmission System Operators (Netbeheerder) (TSO) and produce (Supplier (Leverancier) (LV)) the electricity. KPN is a big multi-site electricity user that is responsible for almost 0,8% of the Dutch electricity consumption. Most of the stakeholders in the chain have their own administration and it is a challenge to align all the data. It is essential to have a high data quality in order to make sure the grid is stable. Two important processes are the nomination process and the regular process of data for A1-connections. The second sub-question that has been answered is: **What is the current state of the electricity chain data and processes?** By doing this, important processes and data were identified that have an effect on the data quality in the chain. The analysis of processes and data lead up to different data quality problems.

Fourthly the different data problems are identified, also the causes that led up to these data quality problems. By doing this the third sub-question is answered: **What are data quality problems in the electricity chain?**. The different parties in the electricity chain are independently working on data quality. This research brings forward seven different types of data quality problems in the electricity chain. Problems regarding the calculation of electricity use, nomination of electricity on the grid, sharing of important data are a few examples. It is important to work together on the data quality in order to realise improvements. Electricity is not nominated correctly on the grid because incorrect electricity usage profiles are used. This causes an imbalance on the grid that is not necessary. Multiple parties have the correct data but this is not always broadly shared. After the identification of the data quality problems through a feedback session with experts more insight is created these problems. An analysis was executed of what the expected effect of the policy recommendations was. So also the fourth sub-question is answered: **How do data quality problems affect the electricity chain?** Incorrect data is generated in the electricity chain because the DSO does not feel responsible for calculating electricity usage data while it is essential for the other parties in the grid. The structure of the energy market plays an important role in the handling of data. From a historic perspective DSOs have had a powerful position where they own the data. This leads to discrepancies in the data that is spread. In the case study example not everybody is feeling responsible for the data quality while the data quality in this chain should be a shared responsibility. The data quality problems originate mainly from the different data quality dimensions that were in the chain category. Responsibilities regarding data quality need to be clearly divided between the partners in the chain. Responsibilities from the DSO should be moved to other organisations (LV or MV) that have a much bigger interest in working with correct data.

Finally different policy recommendations were drawn up in order to improve the data quality in the electricity chain. The fifth and final research sub-question is: **How to improve data quality in the electricity chain?** This leads to the policy recommendations that could improve the data quality and processes within the entire chain. The majority of the policy recommendations are all aimed at redefining the responsibilities for the parties in the chain. Putting the responsibility closest to the party that has the biggest interest in good data quality is therefore essential. Different policy recommendations focussed on improving communication between the different parties in the electricity chain. Other policy recommendations were focussed on transferring responsibilities to the parties that have the biggest interest in good data quality, which currently is not always the case.

At the beginning of my research there was a lack of insight in data quality in the electricity chain. Currently there is a lack of uniform of use in literature about Data Quality Indicators for cases. My research contributes scientifically by a case study investigating of data quality in the electricity chain. My research contributes to society by painting a clear picture for the involved parties in the electricity chain. Right now there is a lack of insight in the electricity chain and the problems that are present regarding data quality. For the different users of the data in the electricity chain it has become clear what problems are that is cause by a low quality data.

Table of Contents

Acknowledgements

This thesis is written in fulfilment of the Master of Science in Systems Engineering, Policy Analysis & Management (SEPAM) at Delft University of Technology within the Faculty of Technology, Policy & Management (ICT department). The research is carried out during a 7 months (April – November 2017) internship at KPN in The Hague.

My first contact with the energy team at KPN was while I was in Chile. It was 30 degrees outside and I was sitting there suited up (while wearing shorts) looking for a stable internet-connection. A few months later I would start my master thesis for real. Thanks Daan, Jikke and Silvana for accepting me as a real team member at KPN. Thanks for the freedom you have given me to just make a plan myself and start different side-projects. It was great to have you all contributing in different ways to my thesis. I know for sure that the 3 of you will make KPN a greener company!

I would like to thank Bert, Jolien, Anneke and Martijn for being in my committee. Their assistance and guidance has been helping me to bring this graduation to a higher level. I know that I this has not been the smoothest process but I want to thank them from the bottom of my heart for the intense supervision in the final 2 months of the thesis. The proces of writing a thesis was not easy but their advice helped me to find the right direction. Their knowledge and academic experience on how to structure a research helped out a lot.

A special thanks to my family. Lots of times I was saying: "I should stop, leave the Netherlands and open a hostel on Honolulu." they nodded but did urge me to finish my thesis first. Thanks for believing in me when I had given up all hope.

Last but definitely not least: the taart-graduation-group. Laurens, Micha, Tom and Eline made me survive the whole graduation process. We have been through a lot but no matter the atmosphere, we always kept saying to ourselves: "Afstuderen is een feestje!". It is one big lie of course, but it was great to share experiences and eat cake every 2 weeks. Thanks a lot for the support!

Delft, University of Technology | Martijn Pronk
November 27, 2017

“Life is like a box of chocolates. You never know what you’re gonna get.”
— *Mrs. Gump*

Glossary

List of Acronyms

A1 Artikel 1 lid 2/3-aansluiting

C-AR Central Connectiondatabase (Centraal Aansluitingenregister)

DSO Distribution System Operators (Netbeheerders)

EAN European Article Number (Europees Artikel Nummer)

EDSN Energie Data Services Netherlands (Energie Data Services Nederland)

GV Bulk consumer of electricity (Grootverbruik)

GWh Gigawatt hour

KPN Koninklijke PTT Nederland N.V.

KV Small consumer of electricity (Kleinverbruik)

kWh Kilowatt hour

LV Supplier (Leverancier)

MV Metering company (Meetverantwoordelijke)

NIO Networking & IT Operations

PV Programme-responsible (Programmaverantwoordelijke)

SJV Standaard JaarVerbruik

TSO Transmission System Operators (Netbeheerder)

VEMW Association for Energy, Environment and Water (Vereniging voor Energie, Milieu en Water)

Chapter 1

Problem identification

1-1 Introduction

Currently the entire electricity market is undergoing a big transition known as ′The energy transition′. The energy transition is shifting the entire system from a centralized electricity production to a more decentralized production. Decentralized electricity production in particular consists of renewables like wind and solar energy. This trend can also be seen in the composition of the Dutch electricity mix, renewables are steadily growing [11]. Consequently the coming years the Dutch electricity infrastructure will be much more dependent on renewable sources.

The greatest challenge emerging from this energy transition is the balance between decentralized electricity production and constant electricity consumption [8]. Renewables tend to be less predictive and not able to provide a constant supply of electricity thereby jeopardising the stability of the grid [5]. A lack of sun or wind reduces the stability of the grid. This means that the electricity grid gets reliant on an external factor like the weather. Therefore, the electricity production from renewables is harder to forecast than electricity produced from fossil fuels [30]. Fossil fuels are able to provide a more stable supply of electricity. The forecasting of renewables is more accurate on the short term however not on the long term [44].

Correct data is needed to reduce imbalances

With more renewables being added the Dutch electricity system will be more volatile and less predictable. More renewables lead to potential bigger imbalances [12], [25]. Imbalance is caused by a mismatch between supply and demand on the electricity market. Within the Dutch energy-system TenneT is in charge of keeping the balance on the electricity grid. An imbalance between supply and demand should be kept to a minimum in order to guarantee the delivery of electricity and prevent blackouts [15]. Based on electricity usage data and electricity production data both demand and supply are monitored by Tennet. The costs for TenneT of keeping the grid in a balance was 52 million euro in 2016. TenneT however also has revenues from keeping the balance on the grid, which is 71 million euro [36]. ″The complexity of the electricity market and uncertainties in assessing electricity production from the likes of wind and solar power makes estimating the grid expenses payable a complex task.″ according to TenneT [36].

Data on the supply of electricity is closely facilitated by the two electricity markets (ENDEX and APEX) where electricity can be bought, days, hours or 15 minutes in advance [4]. A way to monitor the demand more closely is through smart meters. Smart meters provide more information that will lead to a better management of the Dutch electricity grid compared to when this detailed information is not available [14]. Parties in the electricity chain are able to see increases in electricity usage and able to act accordingly. Currently the electricity usage in the Netherlands is based upon estimates. Knowing how much electricity is used is an utopia to parties in the electricity chain. But that utopian future might be closer than you think: the EU and the Netherlands have a goal of having 80% of the connections equipped with smart meters in 2020 [16]. Smart meters are however not a holy grail, smart meters are not always correct and there are also missing values or outliers [13]. But with smart meters there will be more data available on a granular level. Having this data on a more local level makes it easier to manage the electricity grid with the growing share of renewables.

Data plays a vital role in order to combat the imbalances. TenneT is able to predict the imbalances better and take action when more data is available. However in order to act it is essential that the used data is of high quality. One measure to act against these imbalances is through the emergency power collection (noodstroompool). The emergency power collection has a function that it can restore the balance of the grid by changing supply and demand. The emergency power collection is a selection of heavy electricity users that can influence the grid by increasing or decreasing their electricity usage. When high quality data is provided it becomes easier for TenneT to forecast electricity use. With all the renewables that are added to the network it is essential that this goes quickly with high quality data.

What is data exactly and what is information? What are the steps from turning data into wisdom? Why is there a need for data quality? All these questions will be addressed in paragraph 1.2. In paragraph 1.3 the players in the Dutch energy sector are introduced. Furthermore the setting of the research and the dependencies that exist between the different actors are presented. The problem statement and the research objective will be further elaborated in 1.4. The final paragraph will conclude with the next steps for this research.

1-2 Need for data quality

So why is data quality important? Most decisions in businesses are based on information that is derived from a lot of data. An analysis based on data is always dependent on the reliability of that data. More importantly, if data cannot be trusted fully, the analysis that has been done loses value and affects the quality of the decision [41]. Hence within organisations it is crucial that data of high quality is collected in order to make well-founded decisions. High data quality therefore is an essential prerequisite in order to get value out of the data [9]. When data is not high quality you can risk that the data and analysis will not be used [29]. Therefore using high quality data is important in business and research. Data quality problems have an impact on 3 levels: operational, tactical and strategical [32]. At a operational level this could be about the amount of errors that take place. Incorrect data will lead to extra costs for the company. It is important that high quality data is used in order to make carry out the operations. On a tactical level it is important that the correct decisions are taken based on the best data there is available. Strategy is something that takes more time and therefore it is harder to see immediate directs of incorrect data. On a longer term it is harder to base a strategic course when the data is incorrect. Correct data (high quality data) is favored by companies since it adds more value than when the data is from a low quality. A lot of business decisions are based on information that has been deducted from data. However, what happens when the data quality is uncertain? Can decisions still be made? What is the effect on the other parties involved?

A saying in project management is that it costs 1 dollar to prevent an error, 10 times as much to correct the error, and 100 dollar for prevention when something fails [33]. This 1-10-100 rule shows that putting focus on data quality from an early point could be rewarding. Higher data quality comes at a cost. Decisions based on incorrect data can be financially disastrous for an organisation. An estimate of IBM is that data quality problems is costing the US three trillion dollars on a yearly basis [20]. However higher quality data can lead to savings on current spending. Costs are being avoided by making sure data is correct initially.

In order to establish a consistent set of terms in this report a clear definition of data and information is needed. Figure 1-1 shows the DIKW-pyramid which depicts the phases of turning data into wisdom. The following definitions are from Ackoff [1]:

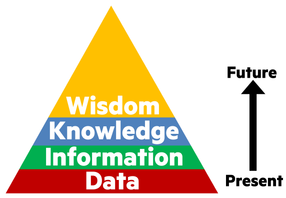


Figure 1-1: Evolution of data inspired by [1]

- '**Data** are symbols that represent the properties of objects and events.' These are mostly raw sources of information.
  For this example we will take a patient that arrived at the doctor. Data could be that the patient has a high blood pressure after it has been measured.
- '**Information** consists of processed data, the processing directed at increasing its usefulness.' Meaning can be derived from information.
  Information about this patient could be other measurements of his blood pressure. Based upon different measurements (data) can be observed that the blood pressure went up from January.
- '**Knowledge** is conveyed by instructions, answers to how-to questions.' The information is placed into a context.
  The knowledge of the doctor is that he knows that events can have an effect on the patient. After some questions the doctor finds out that the patient became a father in January.
- '**Wisdom** deals with values. It involves the exercise of judgment. Wisdom ... is the characteristic that

differentiates man from machines.' Wisdom is when applied actions can be taken. Wisdom therefore is most applicable for future actions. However this wisdom is obtained through data that describes actions in the past.
The doctor makes the connection between the patient's high blood pressure and becoming of a father. The doctor therefore recommends the patient to relax and start working less.

The quality of the data has an effect on the outcomes of a data analysis. When the data (measurements of blood pressure) is incorrect different knowledge and wisdom can be obtained. Maybe there is incorrect data and things are overlooked and different patterns are seen. Better information could be generated when the data on which it is based is improved. Data also plays an important role in the energy sector, a lot of data is exchanged between different parties in the electricity chain.

1-3 The Dutch energy sector

The energy sector is an industry where data is of great importance. Data is used by large companies to identify their electricity connections. The electricity connections in this case are between the utility grid and the consumer. The master data (stamgegevens) is data that is about the connection. The master data are standardized to a certain degree through the Informationcode [2]. This informationcode has been created to facilitate sending of data between partners in the electricity chain. Data found in the master data could be address, zip-codes, consumption of electricity, connection value etc.

In reality not everybody in this sector is agreeing on the master data (stamgegevens) because it is observed that data conflicts exist. What is the ZIP-code of the connection at Nachtegaalstraat 8 in Almelo: 7605BR or 7605BJ? Who needs to pay for the main location in Rijssen: organisation Y or organisation Z? What is the street number of the location at the Valduif: 7 or 66? How much electricity is used at the Dorpsweg 4 in Reeuwijk, 3.000 or 3.000.000 kWh? All these questions will be answered differently by parties in the electricity chain. The different parties have different data stored in their administration about supposedly the same object. Different examples of data quality related problems in the electricity chain will be identified. In this research it is not about who is right, but how these differences in answers are created, what the impact of these differences is and how they could be avoided.

The energy infrastructure is important for the entire Netherlands. In the Netherlands there are around eight million electricity connections [2015] that consume 119.000.000.000 Kilowatt hour (kWh) of electricity on a yearly basis [10]. The majority of the connections to the electricity grid are to households, but there are also industrial connections used for multiple purposes.

The Dutch Electricity market has been liberalized almost two decades ago. Figure 1-2, depicts the current electricity chain. As it can be observed, many parties are involved and are responsible for the production, measuring and delivery of electricity to consumers. Hence, all these parties have to cooperate together in the energy sector. In the chain depicted in figure 1-2 there are a lot of different parties that supply the electricity to the customer. The Supplier (Leverancier) (LV) is responsible for supplying electricity to the grid to be used by the end consumers. In the Netherlands there is a free market where 39 main LVs are supplying the grid of electricity. Another party involved is the TSO. In the Netherlands, TenneT is the only Transmission System Operators (Netbeheerder) (TSO) in the Netherlands transmitting electricity at a high voltage (110.000 Volt and higher). The low voltage network is operated by the eight Distribution System Operators (Netbeheerders) (DSO)s. These DSOs take care of the regional distribution of electricity (50.000 Volt and lower) and make sure only 230 Volt reaches the end consumers. Transmission (TenneT) and distribution (DSOs) are, in contrary to supply, monopolistic markets. There is no competition in this market because of the high sunk costs and investments [27]. Hence, economically it is best to connect houses to only one DSO. TenneT has a public utility function and is therefore fully owned by the government. The Metering company (Meetverantwoordelijke) (MV) measures the electricity that has been used at the connection. Both the LV and MV are operating on open markets where competition is high. Most LVs are also Programme-responsible (Programmaverantwoordelijke) (PV). The PV has the important task of providing information to TenneT. PVs provide information on how much electricity will be used by the connections that are serviced by the LV. This data is essential for TenneT in order to be aware of the electricity grid′s operation and needs.

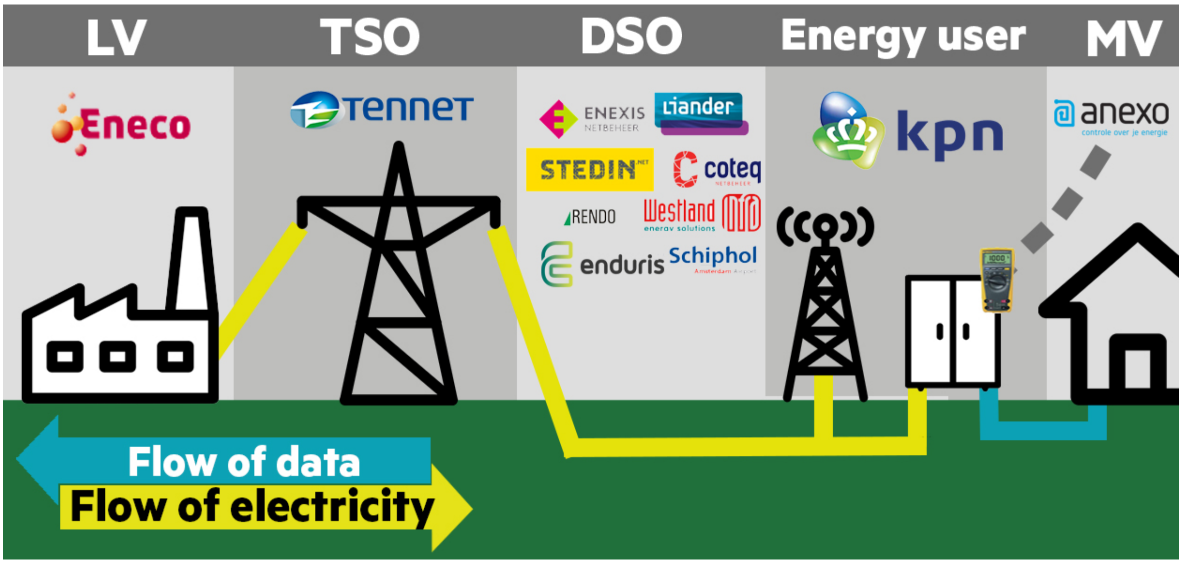


Figure 1-2: All parties involved in the energychain

Balancing the grid costs money since different parties, that are part of the emergency power collection, need to be paid for their service of lowering or increasing their electricity consumption to restore the balance. A reduction in imbalances therefore is financially beneficial. Through improving data better information and knowledge can be obtained about imbalances. These imbalances can be made smaller or even prevented with this information thereby saving money in unnecessary imbalances. For the users of electricity, DSOs and other parties in the chain it is in the best interest to lower the costs of the infrastructure while improving the reliability of the grid.

Data in chains

When data can be combined its value can increase even more for all the organisations involved. Since the energy sector is described by the electricity chain, it is necessary for data to be communicated and transmitted in the chain. Data in chains is the communication between different parties and the transmission of data from one party to the other. However, managing data in chains is not always an easy task. Even at an organisational level, the department which creates the data is not always the one that will use it or benefit from it. Therefore, the goals on data quality and context are not always aligned which leads to less improvements on data. The same problem arises between different organisations.

Each party is handling the data in a different way and they may need the data for different reasons. This makes the transmission of the data problematic. However, within a chain there also lies a great responsibility. All parties rely on each others data. Organisations in chains should therefore realise that a mistake from their side can have negative consequences down the chain. Parties create their own administration of the activities that they are performing. Therefore data exchange can be problematic and have a big impact on the different parties in the chain. The problem is that the department which creates the data is not always the one that will use it or benefit from it. Therefore, the goals on data quality and context are not always aligned which leads to less improvements on data. Within a chain some mistakes are more important than others since they have a bigger impact on the operation [7].

Something that complicates investment is when multiple parties are benefiting from good data. Who should pay for it?

Policy window: a new law

So there is a need to investigate the impact of (low) data quality in the chain, but there is another important reason why action should be taken. In december 2015 a new electricity and gas bill was rejected by the Dutch senate [22]. A new law ”Wet Voortgang Energietransitie” (VET)[37] is currently under construction. The next law (VET) will be made by the new government that will take office by the end of Oktober 2017. This new law will replace the law that has been rejected by the Dutch senate. In policy literature a situation like this is illustrated with the streams model. The streams model describes the moment in time where streams of problems, solutions and policy all meet and are all open: a policy window [23]. This so-called policy window is an opportunity to add legislation that could improve the data quality in the entire chain. Now a law is changed and the situation will be frozen again for some years. So this is an opportunity for the different parties to include certain pieces of legislation that are aimed at improving data quality. At the end of this year it is likely that consultations at big users of electricity and interest groups will be happening. This research will therefore come up with policy recommendations that can improve the data quality in the electricity chain.

1-4 Problem statement and Research Objective

The electricity chain is a vital infrastructure in the Netherlands. With more renewable energy and the transition in which more data is being collect it is is of the utmost importance to have high quality data that can be relied upon. The functioning of the electricity grid will rely more and more on data. For data to be trusted and accepted by all the different parties it can be challenging. In the energy sector the dependencies are very high when it comes to data. In the chain multiple organisations use, produce and analyse the data. So the parties in the electricity chain are working with the information that has been given to them by other parties. It is essential that the trust is maintained among parties and that the data quality is high. Improving the data could lead to a better utilization of the electricity grid at lower costs. But before this can be done first the data quality problems need to be identified.

As has been mentioned before there is a policy window where policy can be changed. Therefore the objective of this thesis is **to identify policy recommendations that can be made to improve data quality in the energy data chain**. A chain is only as strong as its weakest link. A single organisation that is not dealing correctly with data can be undermining the trust in the entire electricity chain. When incorrect data is discovered it should be analysed where this came from. But how can data quality be analysed? Policy recommendations to improve the data quality can only be drafted when a clear overview exists of the problems in the chain. The different parties in the chain can start a further discussion based upon the findings in order to improve the processes. This way, policy recommendations will be made towards all the parties involved, which will lead to an improvement of the respective internal and external data exchange processes.

1-5 Scientific contribution

This research contributes with a case study investigating of data quality in the electricity chain. Various papers have been written on how data quality can be improved within a closed environment. Something that is not always taken into account is when the data is exchanged in a chain and multiple parties rely on the data. Just typing in data quality on sciencedirect already provides 2,327,133 results. There is a lot of literature on data quality already, so there is also a need for a structure. For the different parties in the electricity chain it will become clear which data quality problems currently occur. Previous research in this dept on the Dutch electricity chain has not been conducted yet to my knowledge. By developing multiple policy recommendations the parties in the electricity chain know how certain problems can be solved. The developing of policy recommendations for the dutch energy sector is something that will contribute to the already existing literature. The majority of the findings from the thesis are applicable to the Dutch electricity market however multiple outcomes are rather general and can be applied in other situations.

1-6 Next steps

In the next chapter the entire Research design of the thesis will be further explained. The research question will be explained together with the five sub-questions. The different chapter leading towards (policy) recommendation will be drawn out in chapter two.

Chapter 2

Research design

In this chapter the research will be further explained. Section 2.1 will be further elaborating on the research question for this thesis. The research approach and the different phases of the research together with the research design will also be analysed in detail in section 2.1. The research flow diagram will be presented in 2.2 together with the next steps ahead for this research.

2-1 Research Question and phases of the study

The main research question of this master thesis project is: **Which policy recommendations can improve data quality in the electricity data chain?** The main research question is divided into sub-questions that will be helping to answer it. The research subquestions all tackle a different aspect of the research question. The sub-questions are embedded in the research design.

A case study will be performed to see how the Dutch electricity chain functions. But before the case study will be started a review of literature will be done. In this thesis data quality that concerns multiple organisations will be the main focus. The description of the research below will be divided into these five phases. First literature research has been done on how data quality can be expressed. Secondly it was made sure that a proper analysis was executed on how the current system is exchanging data and carrying out certain processes. This has contributed to the identification of how data quality problems occur in the third phase. This third phase was to further analyse the system and find the causes that led up to data quality problems. The fourth phase was about the receiving more feedback on the different problems in the electricity chain. Experts in the energy field have been consulted about the data quality problems that have been observed in the energy sector. A fifth phase is to analyse the effect of the policy recommendations that have been drafted. Here below a more elaborate description of the different phases will be given:

Figure 2-1 depicts the different research phases that needed to be employed as they were presented in the previous lines together with the research questions that try to address.

Finally, the research method that was employed for each research question is also presented.

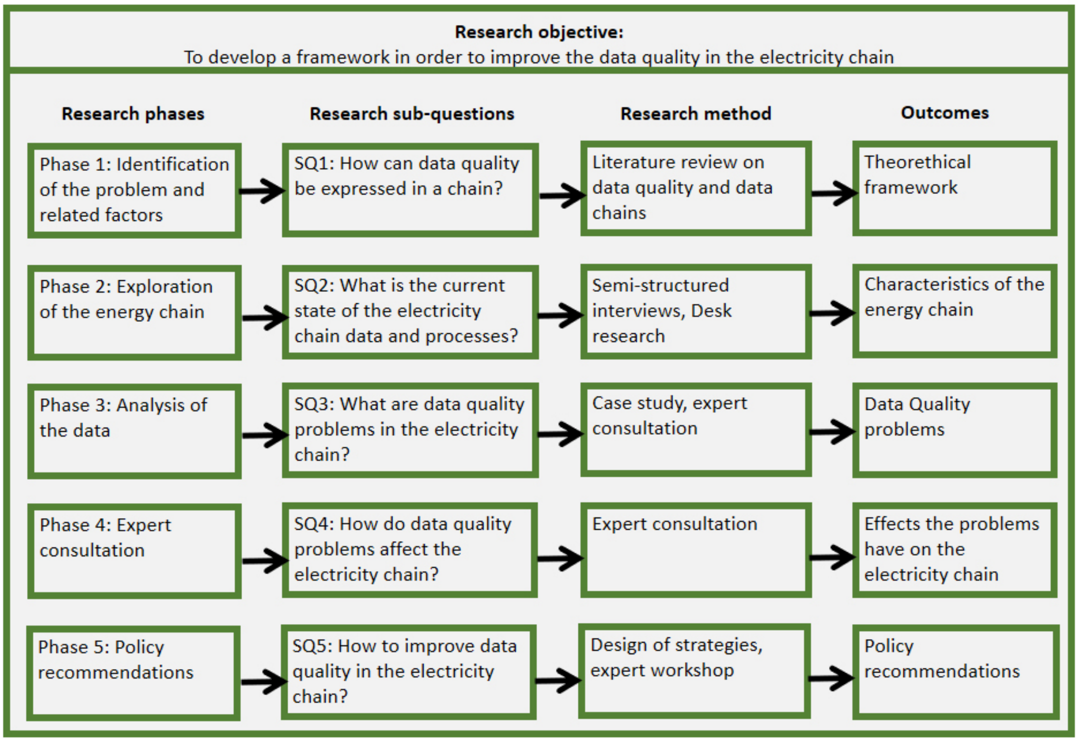


Figure 2-1: Overview of the research design

2-1-1 Phase 1: Identification of the problem and related factors

To answer the research question mentioned above, different steps have been executed. First, a literature review on Data Quality and Data in chains will be done, to become acquainted with the existing techniques and policies. All this is done in order to answer the first sub-question: **How can data quality be expressed in a chain?** This will be done with the help of a theoretical framework to define (the concept of) 'data quality'. This framework will later on be used in order to analyse the case.

Literature on Data Quality

Once literature has been investigated, important data quality dimensions will be identified. These data quality dimensions will be used to analyse the data quality problems in the energy data chain. Data quality and the governance in the chains is important to improve data quality. With this in mind we will look into the existing literature on how data is being exchanged.

Setup of literature review

The following setup of the literature reveiw has been inspired by Kitchenham [24].

- **Selection of studies** The literature for the review has been mainly found on Scopus, ScienceDirect and Google Scholar. Different keywords have been used in order to find the literature: Data quality, Data quality indicators, Data quality in chains, Data quality indicators in chains, Data Quality, Data Provenance, Data Chains. To see if the literature was valuable for the research the abstract was read and if considered of interest read and analysed. Next to searching the databases snowballing was used as an additional strategy to find relevant literature. Most papers that have been studied have been peer reviewed. The relevance of the literature has been decided after the abstract has been read.

- **Study quality assessment** I have used grey literature as well as important academic papers from Scopus, ScienceDirect and Google Scholar.

- **Data extraction and monitoring progress** Batini did a literature research on the most important literature when it comes to data quality indicators/dimensions. Therefore I am not afraid that I have missed out on very important papers.

- **Data synthesis** At the end of chapter 3 the data quality dimensions that have been found in the literature are brought together.

2-1-2 Phase 2: Exploration of the case

The case study will explore data handling at KPN, the largest telecom provider in The Netherlands. KPN is a large user of electricity and electricity data. An identification of the current processes is vital to understand the system. Besides interviews with different partners in the chain also desk research at KPN has been conducted to obtain a better insight into the system. All this is done to answer the second sub-question: **What is the current state of the electricity chain data and processes?**

Case study

Koninklijke PTT Nederland N.V. (KPN) is a big telecom provider in the Netherlands. KPN's network is essential for the connectivity of the companies, datacenters, millions of devices and homes in the Netherlands to the internet. Operating a fixed network and a mobile network means that a lot of electricity is needed to make this happen. The electricity use of KPN is one of the biggest in the Netherlands. On a yearly basis 761 Gigawatt hour (GWh) (2016) [26] are consumed which amounts to 0.7% of the total Dutch electricity consumption. KPN has approximately 20.000 connections spread over the Netherlands which makes them a so called multi-site user. Multi-site users have multiple electricity connections spread over the Netherlands. Examples of other multi-sites are: telecom providers, government organizations and utilities. Those organisations differ vastly from the households due to their higher electricity needs.

Nevertheless, more insight in the electricity chain is needed for KPN. The presented work is focusing on the case study of KPNs position for simplification of information. The manual administration of the electricity connections of KPN is an intensive task. The energy team therefore wants to simplify and automate this process. But before things can be automated a better insight in the data and the processes need to be obtained. According to KPN there are different uncertainties in the system which need to be identified. Once the high quality data is used, knowledge can be derived from that to be used for the future.

Yin has distinguished a clear design for Case studies. Since for this research also a case study is used multiple elements from Yin′s work will be used to clarify the use of the case study [43]. There are different ways to obtain evidence, for this research mainly documents and semi-structured interviews will be used. First documents from KPN and the energy sector are analysed. These documents will provide an initial understanding of the energy sector. The semistructured interviews will be used multiple times to obtain more information, to verify findings and to review the work.

Semi-structured interviews

More insight into the electricity chain is needed from different perspectives. Experts throughout the chain have been consulted. This has been done by speaking to different parties in the electricity chain. First the causes of incorrect data have to be identified, before the mistakes can be reduced through different recommendations [6]. For this research it was important to talk to the different parties in the chain: Distribution System Operators (Netbeheerders) (DSO)s, Supplier, Multi-site users and the Measuring organisation. A detailed summary of the different interviews can be found in Appendix A.

2-1-3 Phase 3: Analysis of the data

All parties have given insight in how the different processes run. After all information from the different parties is gathered it is a task to identify the problems that occur when data are transferred from one part to another and have to be used by various parties in the electricity chain. The information that has been provided by the interviewees has helped to answer the third sub question: **What are data quality problems in the electricity chain?** Also lessons and knowledge can be created from the analysis of the data. Through the errors that arise from the analysis of data exchange, processing and analysis thereof, it will be possible to identify the mistakes and malfunctions in the electricity chain. The causes of the mistakes can be identified in this phase.

Data analysis of the electricity chain

Besides interviews the organisations have been asked to provide all the information they have about KPN. Through KPN a lot of data can be accessed and the data analysis can reveal a lot about what data is transferred, the quality of these data and how these data are treated. A data analysis on multiple fields will be conducted to see if there are any differences. Input for the data analysis will be the interviews with the experts in the field. After talking to the experts a guided analysis of the data can be executed. The data analysis will show the data quality problems that exist in the chain. Together with the interviews it provides a clear insight in how the electricity chain works.

Beheading the outcomes of the analysis it will show that clear responsibilities in the chain are established when it comes to the physical network; but when it comes to the data in the the electricity chain the responsibilities are demarcated less clearly.

2-1-4 Phase 4: Expert consultation

The findings on data quality, exchange and handling that have been discovered have been presented and discussed with the experts. Their input was needed for the fifth sub question: **How do data quality problems affect the electricity chain?** Once these problems and their consequences are identified, the next phase will be handling the identification of the points to improve. The experts are working in the energy field. Their feedback therefore is important to increase the robustness of the recommendations that will be drafted up later in phase 5. This verification procedure is a check to see if the behaviours and findings match the feeling of the participants. In september 2017 all people that have been interviewed before have been invited to discuss the findings and recommendations. In this meeting a discussion is started with all the parties that are involved in the electricity chain. The findings have been presented together with ways to keep the data current and more valuable for all the parties.

2-1-5 Phase 5: Development of the recommendations for policy makers

When all information is collected it is possible to draft up policy options. These recommendations could potentially be used in the new law concerning electricity (VET). These recommendations blend in well with the fifth research sub-question: The fifth sub question is: **How to improve data quality in the electricity chain?** Based upon the input throughout the report recommendations can be developed and are described in the final chapter. These different recommendations will be created using the input from literature and the experts in the field. The policy options will influence the processes that cause the current low quality data. The policy options are loosely discussed with experts in the field and extensively discussed within KPN. For this research the data quality and the recommendations will be of applied value to the entire chain.

2-2 Overview

Here below a flow diagram can be seen that graphically represents the thesis. Through different sub-questions this thesis will come up with policy recommendations in the end. The green boxes in the figure below are representing the answers to the different research sub-questions.

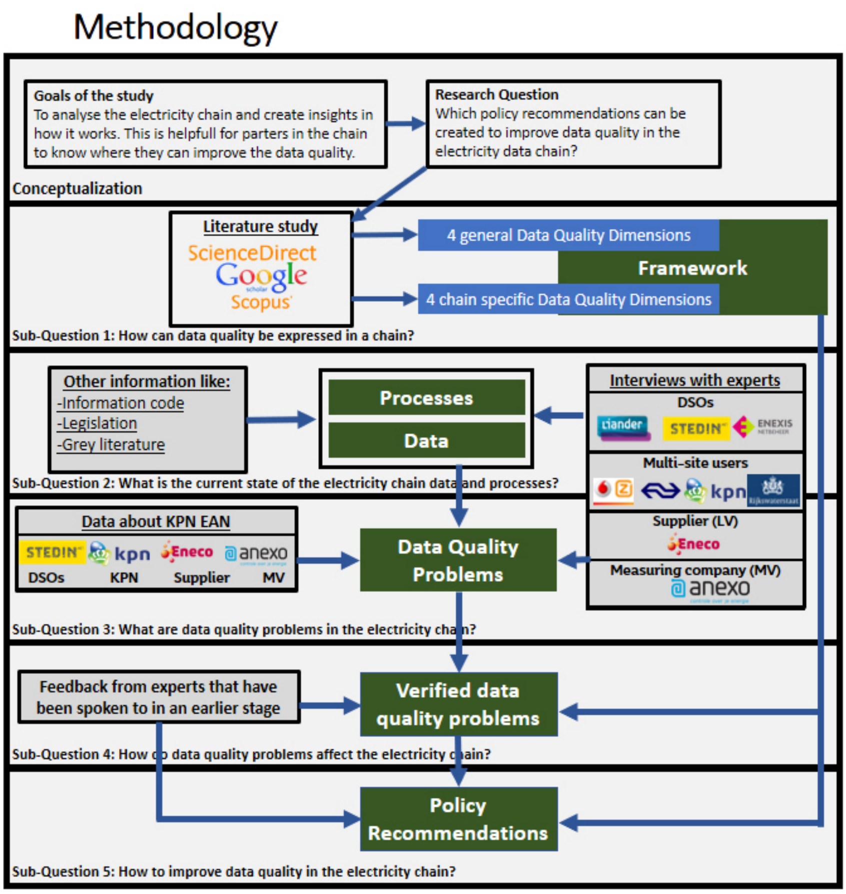


Figure 2-2: Flowdiagram of the research

In the next chapter the literature will be explored in order to define what data quality is. Based on the outcomes of the literature review a theoretical framework will be constructed.

Chapter 3

A theoretical framework for analyzing data quality in the energy chain

What is the definition of high quality data and how could the data quality be determined? In this chapter a theoretical exploration of the concept data quality will be described. Dimensions for data quality will be derived from the literature. Besides data quality for just the companies it is also very important to take into account the effects of companies that are part of a chain. First existing literature will be identified together with the important findings regarding data quality in 3.1. In 3.2 four different Data Quality Dimensions (DQDs) will be selected that are used most in literature. An explanation of the different concepts together with definitions of the DQDs will be given. In 3.3 four other DQDs will be handled, these DQDs are suited better on capturing data quality in a chain. In 3.4 a theoretical framework will be presented based upon the 8 DQDs from 3.2 and 3.3. By executing all these steps an answer will be given to the first research sub-question: **How can data quality be expressed in a chain?**. The theoretical framework, that is the answer to this question, will be used further on in this research. .

3-1 How can data quality be defined?

In large supply chains the only way to measure the performance is through data [18]. Trust in the data is needed and this is also based on the mutual trust that has been placed in the partners of the electricity chain. Once trust in the data or the partner is lacking it is impossible to use the data for an analysis. What information should companies share with others? Is the information crucial, and does it need to be shared? Data quality therefore is a concept that is hard to describe since multiple parties have different perspectives on good data. A focus on the user of the data and the needs of other partners in the chain is crucial in order to maximize the value for the different users. ″One man′s trash is another man′s treasure.″ In this context it means that certain data that are useful for one party might not be useful for another. Organizations in general however tend to act in their own interest and focus on the data that is essential to them. This first section will cover literature on data quality. After that other literature will be covered that also looks into the needs for data quality of parties in a chain. Different perspectives on how to describe data quality will be presented.

In 2001, Dough Laney introduced three Vs: Volume (amount of data), Velocity (speed of new data) and Variety (different forms of data) [28]. However over the years other V′s have been added to that: Veracity (uncertainty of data), Value (value that can be obtained through analysis) and Validity (credibility of data). The most important factor however can in this research can be demarcated as Veracity. When the data is not trusted or correct then it will be hard to create value out of it. In chains most data is used by multiple organisations. So the data that is created by organisation can add value to the processes of partners in the chain.

Information Resource Management is a way to systematically improve the Data Quality in a given system [3]. Most literature never acknowledges the difficulty that comes with a shared responsibility for data. However how can Data Quality Dimensions be applied to a chain? And why would you want to improve the data quality in a chain?There are different reasons to focus on improving data. One of them is to be a reliable/proactive partner. When you can maintain on high quality you are able to strengthen ties with your partners in the chain. To improve processes, you first need to be sure the data is correct. The better management is done on data, the better information can be obtained from the data. Data turns to information and Knowledge by proper analyses. The better the data the more value data can generate. However within a chain there is an additional complexity.

Another general term in literature that is used frequently to classify data: Fitness for use. It is a term to describe if data is usable for the designated process [39], [34] [38]. It is up to the user to decide if the data is fit for use. Data clearly needs to have a high quality in order to be fit for use. High quality is needed since poor data quality can alter conclusions that are drawn based upon data. Quality control therefore is an essential tool that should in order to prevent this from happening [42]. Again the focus of these papers is on a small environment within a system. The ′fitness for use′ varies a lot depending who you ask about it. In order to make sure data quality is maintained or improved quality control is also very subjective. What is for instance an acceptable error rate when it comes to electricity usage?

Another well established approach for improving data quality is: Total Data Quality Management. This means that the data is continuously defined, measured, analysed and improved, thereby improving the data quality. [38]. Additional data cleaning campaigns are needed in order to improve the data quality. Also in this paper not a lot of effort is put in recognizing the complexity when multiple stakeholders are involved in the improvement process.

Horton does take into account the different organisations that come into play when data is exchanged in a chain. Horton defined multiple roles when it comes to the processing and usage of information (Figure 3-1) [19]. The four different roles are important and show the difference in interests. The different roles also do not have to be within the same company. When applying the framework to the electricity chain KPN is seen as a data user, while the DSO is the information supplier and another party, the EDSN, manages the data. It is important to know where the responsibilities lie within in the system and what the different roles are.

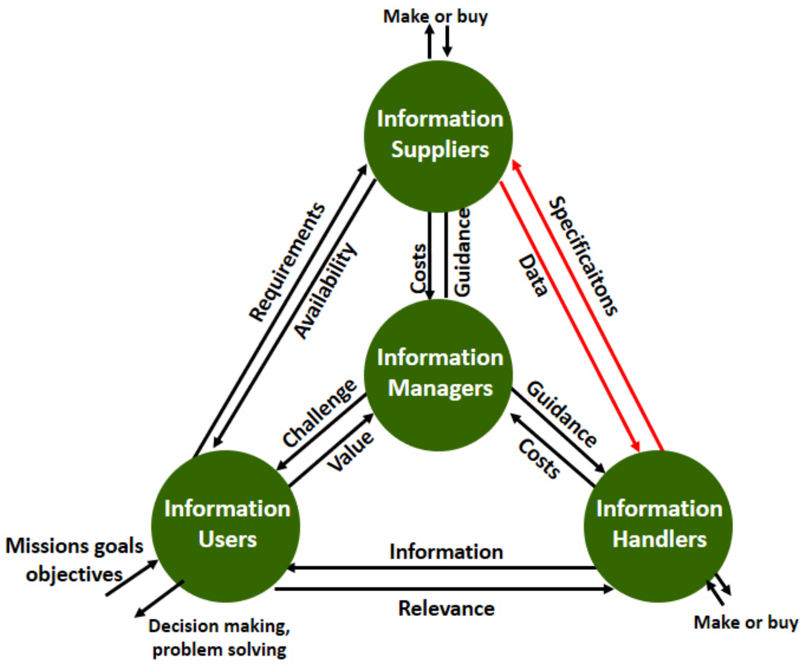


Figure 3-1: Overview of different roles by Horton [19]

3-2 Data Quality dimensions

For assessing data quality various dimensions can be found in the literature [32], [7]. A dimension is ″a set of data quality attributes that represent a single aspect or construct of data quality″ [40]. Good data quality dimensions can limit the uncertainty and can improve the analysis of the process [41]. So using the correct data quality dimensions is important. There are a lot of dimensions and indicators to data quality according to professionals [40] and research [31]. Accuracy, timeliness, consistency and completeness are four data quality dimension that are used a lot in important literature [34]. These four dimensions are considered to be the most used in literature and therefore will be explored into more detail.

3-2-1 Accuracy

″Accuracy of data refers to how closely the data correctly captures what it was designed to capture″ [21] Accuracy is needed to see if the correct value has been recorded. When a street name is matched to a wrong connection this can be considered inaccurate. When wrong payment data is connected to a customer this can be considered also an inaccuracy mistake. Some mistakes are more important than others since they have a bigger impact [7]. Accuracy therefore is variable in the impact it can have on an analysis.

3-2-2 Completeness

″Completeness of data refers to the extent to which the data collected matches the data set that was developed to describe a specific entity.″ [21]. Data records that have all fields filled in are of course the dream of every organisation. Unfortunately that is not always the case. Not all data is complete and therefore parties need to handle the non-completeness.

3-2-3 Timeliness

″Timeliness is the extent to which age of the data is appropriate for the task at hand ″[40]. When data is old it still can be considered good for the task. Or based upon the characteristic it can be outdated and therefore not current [7]. Within the electricity chain physical infrastructures are there for long periods of time. These infrastructural data elements within a chain are not really losing value after years. Electricity usage data however need to be much more recent in order to have value. The electricity usage value is most valuable when it is very recent as the planning of the power to be generated is based on them. In 20 years time the added value will be close to zero. It is important that the data is correct. However not all data needs the same timeliness in order to keep the data current and thereby contributing to the data quality.

3-2-4 Consistency

″Consistency is the process of checking data and if they are consistent. When Jaffalaan 5 is the street name and Delft is the city then the zip-code should be 2628BX. When this is not the case then the data is not consistent. Data can be inconsistent when different parties are using different procedures. There are a lot of parties involved and not everybody is sticking to the same principles. Standardisation of the data entries is preferred since it will then be easier to turn data into reliable information.

3-3 Complexity of data quality in a chain

A dominant way to measure data quality is through the different dimensions. Data quality is measured separately by organisations in a chain based upon the four dimensions mentioned in 3-2. However for chains a broader approach should be kept in order to unleash the full potential. Data quality can be good according to one party but might not be sufficient for another party. Therefore there is a need for more dimensions to be taken into account.

3-3-1 Reliability

″Reliability of data refers to the extent to which data is collected consistently over time and by different organisations either manually or electronically″[21]. When data is operated in a chain there is a big dependence between the parties. Reliability is an important dimension since a lack of it will not contribute to better data.

3-3-2 Comparability

″Comparability of data refers to the extent to which data is consistent between organisations over time allowing comparisons to be made.″ [21]. In the electricity chain there are different organisations having information about the same physical connection. It is therefore important that data can be compared between the parties in order to make sure there are not any differences.

3-3-3 Accessibility

″Accessibility of data refers to how easily it can be accessed, the awareness of data users of what data is being collected and knowing where it is located.″ [21]. Within the electricity chain there are multiple types of organisations involved. Access to the data and a spread knowledge of what information is available is therefore very important.

3-3-4 Credibility

″Credibility is used to evaluate non-numeric data. It refers to the objective and subjective components of the believability of a source or message.″ [9]. When multiple organisations are working with the data it is important that the data that is exchanged is credible.

3-4 Theoretical framework for further analysis

Here below in figure 3-2 the eight data quality dimensions are put together. The four common Data Quality Dimensions are seperated from the four Data Quality Dimensions that describes dq in chains. So there are two ′thermometers′ to measure whether there are any data quality problems in the chain. These eight dimensions will be used in chapter five to see how they relate to the identified data quality problems.

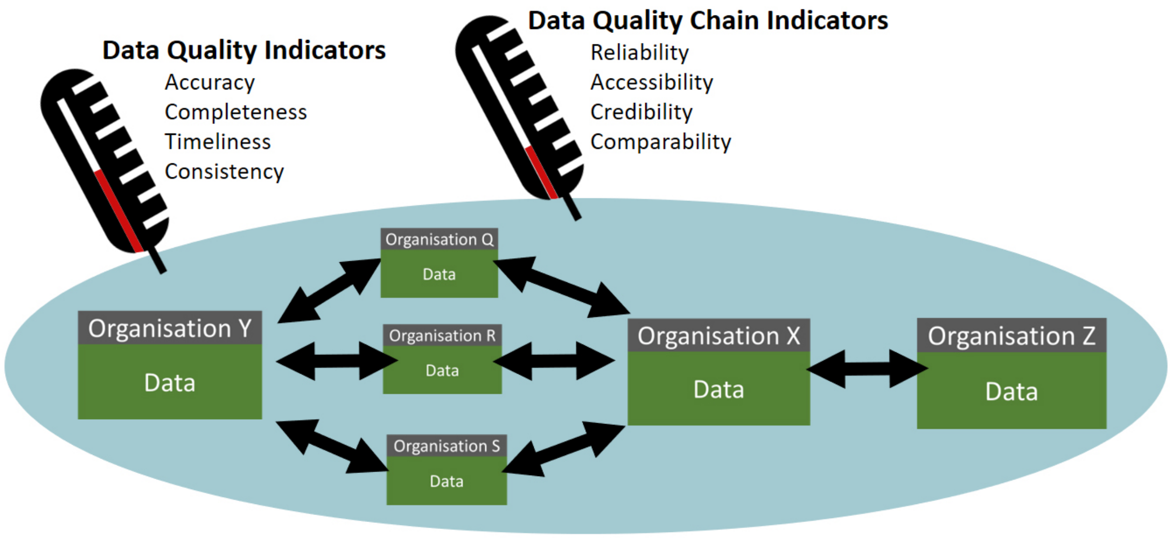


Figure 3-2: Framework of how the data quality in a chain is measured

3-5 Conclusion

This chapter had the purpose of answering the first research sub-question one: **How can data quality be expressed in a chain?** Expressing data quality within one single organisation is a tough job. Expressing data quality for multiple organisations add additional complexity. In chains this extra complexity consists out of multiple interests, responsibilities and interdependencies between organisations. The value of the data increases when it is of higher quality since more information can be deduced from it. This is vital for the costs, innovation and stability of the electricity infrastructure. There are multiple ways to describe data quality. For this research the first four data quality dimensions are: accuracy, completeness, timeliness and consistency. However what has been seen in literature is that a lot of literature does not take into account the multiple organisations that can be involved. Therefore four other Data Quality Dimensions (credibility, accessibility, reliability and comparability) have been added to the framework. These four extra DQDs that make data fit for use should be connecting better to the data quality problems in the chain. Once the problems are identified in chapter four they will be analysed together with the eight DQD that have been found in this chapter.

Chapter 4

Case study description: the Dutch electricity chain

There is a need for high quality data in the electricity market. However before data can be improved first the processes and data should be understood. In this chapter insight will be provided on the current state of the electricity chain. In paragraph 4.2 a detailed introduction to the electricity chain/market will be given. The information that is provided in this chapter has come from desk research and interviews with different players in the electricity chain. The interviews have also provided insight in which data is owned by which party. An overview of all the data in the chain is given in section 4.3. In order to better understand the market two important processes are highlighted in paragraph 4.4 All these steps have been executed in order to answer the second research sub-question: **What is the current state of the electricity chain data and processes?** in paragraph 4.5.

4-1 Case study protocol

Earlier in paragraph 1-5 the policy window was described that will occur in the coming period. A case study was needed where a lot of electricity connections were at the party. KPN has a lot of connections spread over the Netherlands. They need to work with all the different parties and therefore makes the KPN case very suitable for a case study. KPN wants to take this opportunity to provide new input to the law. Interest groups like Association for Energy, Environment and Water (Vereniging voor Energie, Milieu en Water) (VEMW)(where KPN is a member of) have the possibility to respond to the legislation before it goes to the Second Chamber.

The KPN energy team is positioned within the Networking & IT Operations (NIO) department at KPN and tasked with everything that is related to energy. They manage all connections, provide forecasts and try to create the best possible information out of all the data. The energy team interacts with many different parties. Some of these parties are MV, DSOs and LV (Eneco) and it is the team's responsibility to ensure that electricity is taken care of and paid for, thereby ensuring connectivity. The data provided to the energy team is obtained from the different parties that KPN needs to interact with, within the electricity chain. The energy team sees in their everyday work that the data quality is not very high. KPN has experienced, in their interactions with the different parties, that they have their own version of the truth. Within the energy team there is a feeling that the other parties are not using all the data or data is simply incorrect. An effect of this incorrect data is that KPN needs to correct their estimates of the electricity use. A consequence of the incorrect data is that big settlements need to be paid and huge amount of work is done on getting the numbers right. Therefore KPN wants to know how reliable the data is in the electricity chain. Figure 1-2 illustrated that the stream of data runs the opposite direction than the electricity. The data in their electricity chain heavily influences their estimates of the amount of electricity used (which is a KPI).

In this paragraph an explanation is given on why a case study was essential for this research. In order to analyse a system in general it is important to have a certain degree of access. Having an organisation as a case study that is interacting with other organisations was important. Preferably the organisation was also handling a lot of data in order to be exposed to lots of different problems. KPN was fulfilling the criteria of being an involved party that actively participated and had multiple data exchanges. Real life examples were needed in order to stress the importance of data quality management. Without doing a case study it would have been impossible to say anything about the Dutch electricity chain.

For this research it was important to talk to the different parties in the chain: Distribution System Operators (Netbeheerders) (DSO)s, Supplier, Multi-site users and the Measuring organisation. Multi-sites are a special category of users of electricity. Multi-sites have multiple connections to the grid spread over the entire country. Multisites also have different usage profiles of electricity compared to households. Their locations are not easily accessible and require a lot of work to keep track of. Speaking to them led to a broad view of the market since they have a lot of different connections. For the interviews also three other multi-site users (Rijkswaterstaat, NS and VodafoneZiggo) besides KPN have been consulted.

After the different multi-sites have been interviewed the DSOs followed. For this thesis the three main DSOs (Stedin, Enexis and Liander) have been interviewed. These 3 DSOs have dedicated account managers for KPN and also together have around 90

I have spoken to the interviewees shown here below. All of them are related to the connections of KPN.

Table 4-1: Interviewees in the electricity chain

Company	Category	Role interviewee	Date of interview
NS	Multi-site user	Energy advisor	12-06-2017
Vodafone/Ziggo	Multi-site user	Energy connection manager	21-06-2017
Rijkswaterstaat	Multi-site user	Energy expert	07-06-2017
Anexo	Measuring company	Energy expert	01-06-2017
Eneco	Supplier	Strategic Partner Manager	10-08-2017
Stedin	DSO	Accountmanager	25-07-2017
Liander	DSO	Accountmanager	22-08-2017
Enexis	DSO	Accountmanager	18-07-2017

The questions below were discussed with the participants. More information about the interviews can be found in 2-1-2 where more information is given about the semi-structured interviews.

Topics discussed during the interviews

- More information about the interviewee and the organisation.
- What are the different activities that the interviewee performs?
- What are the responsibilities of your organisation?
- How is the collaboration with the other parties?
- How does the interviewee use the C-AR?
- How does an A1 connection work at their organisation?
- How important is data quality for the interviewee and the organisation?
- Are there measures to check if the data is correct?
- What are low quality data problems within their organisation?
- What can be improved according to the interviewees?
- What are they doing with the SJVs?
- A further disection of 3 EANs

Figure 4-1: Overview of the different topics that are handled during the interviews

4-2 Introduction to the Dutch Electricity market

Below an introduction to the most vital elements of the electricity market is given. Figure 4-1 is a representation of the different responsibilities in the electricity chain. A more detailed explanation of the different parties can be found in 1-3. Something that was not mentioned there is the importance of the different types of connection. These categories of connections play an important role in the handling. Depending on the type of connection different parties are executing services. At the end of this paragraph an overview of the connections will be given.

Possible electricity connections in the Netherlands

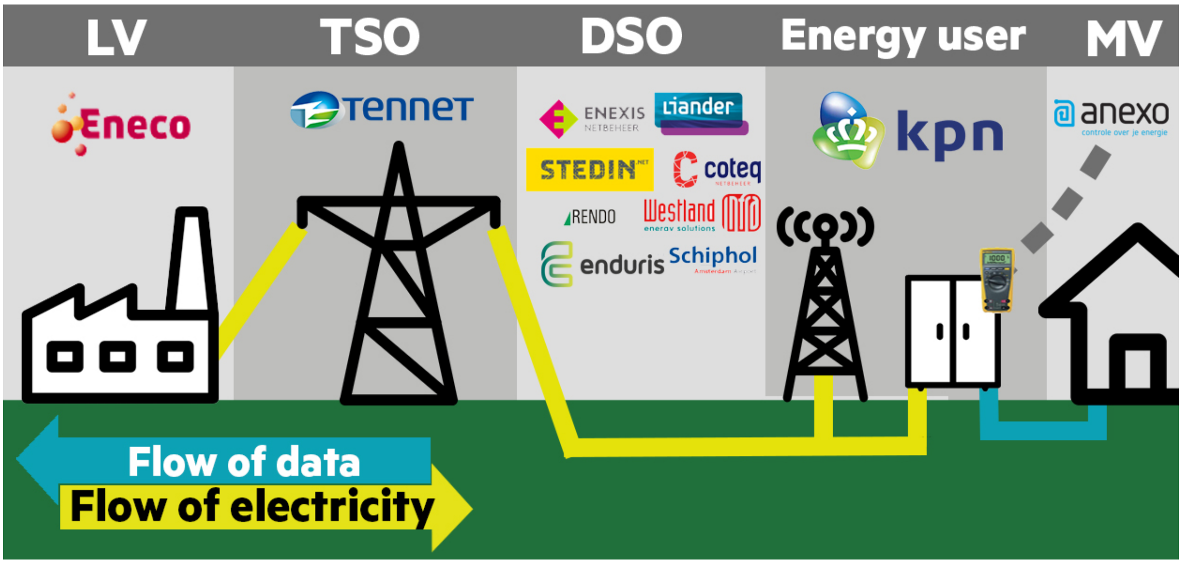


Figure 4-2: Important parties involved in the delivery of electricity in the electricity chain

The majority of the electricity connections in the Netherlands use a relatively small amount of energy, like households. To illustrate this Figure 4-2 shows the connections from both KPN and Stedin. Stedin had 2.028.778 connections that fall in the category Small consumer of electricity (Kleinverbruik) (KV), 18.863 connections that are Bulk consumer of electricity (Grootverbruik) (GV) and 11.217 connections that are Artikel 1 lid 2/3-aansluiting (A1) in 2016 [35]. The majority of the connections therefore are households. KPN on the other hand has an entirely different ratio when it gets to the amount of connections. KPN′s connections are mainly consisting out of A1. Here below the different connections and its characteristics will be further explained.

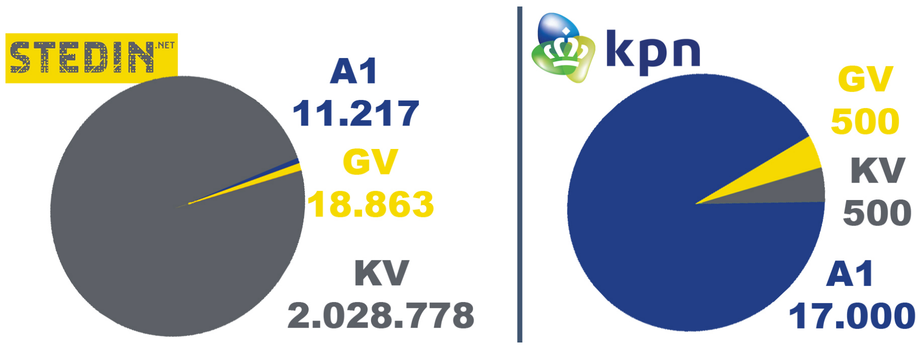


Figure 4-3: Overview of the electricity connections at KPN and Stedin [35]

By law (Energiewet) there has been a defined structure with the responsibilities. According to the law there are two main types of connections to the grid: KV and GV. An exception that is applicable to only few connections is Artikel 1 lid 2/3-aansluiting (A1) An important overview of the differences between the three connections can be seen in the figure 4-3.

KV-connection

Almost all households in the Netherlands are Small consumer of electricity (Kleinverbruik) (KV) since their connection does not exceed 3x80 Ampere. KV connections are part of the so called 'Leveranciers-model'. This means that the consumer only pays the LV. The LV pays the DSO on behalf of the customer. This makes it easier for the consumer since they only have to pay one party. The measurements for KV are done by the DSO. The DSO in this case is also the Metering company (Meetverantwoordelijke) (MV). KV-connections only need to pay the LV and they take care of the payments to the DSO for the transport and measuring services.

GV-connection

Bulk consumer of electricity (Grootverbruik) (GV) is mainly used by larger companies since they have a connection that is larger than 3x80A. They need to pay different parties: the LV for the electricity, the DSO for the local transportation of the electricity and a MV for the measurements that are carried out. GV-connections need to measure their electricity-use every 15 minutes and send this to a portal. This information is used to look what the actual use is and at the end of the day the electricity company knows what the imbalance on the grid was.

A1-connection

An Artikel 1 lid 2/3-aansluiting (A1)-connection is an exception according to the law. A1 needs to be used for connections of municipalities, telecom operators, public transport agencies (like NS and RET) and utilities (like water transport), mining, traffic lights or public lighting. All the connections of these organisations can be seen as one connection which leads to tax reductions regarding electricity. So all the telecom operators have to switch their connection to A1. A1-clients do not have to pay Energiebelasting based on the amount of connections. Therefore it is financially beneficial for companies to do this. A1 is a connection that is smaller than 3x80 however according to the law it is a GV-connection. So it can select a MV by itself. So in the case of KPN, Anexo is the MV that measures the electricity use at the connections throughout the country.

	KV	A1	GV
Measurements	Once a year		15 min
Connection	<3x80A		>3x80A
MV (meetverantwoordelijke)	LV	MV	
According to the law	KV	GV	

Figure 4-4: Overview of the different types of connections

4-3 Data in the electricity chain

Companies in the electricity chain used to be much bigger 20 years ago. Eneco for instance was a LV, MV (Joulz) and DSO (Stedin) at the same time. According to the EU-law this had to be split up. So over the years all the different activities in the value chain were not executed anymore by the same party.

A highly standardized way of sending information has already been established between most parties in the electricity chain, the so-called ′Berichtenverkeer′. There should be a clear distinction of responsibilities so that the customer is always receiving electricity and paying for it.

DSOs do not have a uniform way of working, they all have their own procedures, instrumentation and administrations. A striking example where these different modes of operation of DSOs is shown is the exchange of Weert. Stedin and Enexis needed to execute this change in order to divide the country based on the borders between provinces. Currently in Weert all currently installed smart meters need to be replaced since the smart meters from Stedin cannot be read out by Enexis. This example shows that parties in the electricity sector all have their own way of working.

4-3-1 Data division between the organisations

Every connection in the Netherlands has an European Article Number (Europees Artikel Nummer) (EAN). This EAN is a unique 16 digit number that contains information about the connection. Within the entire chain all parties understand what a EAN is and what kind of information is connected to it. In order to organize all the data a more standardized approach should be taken. The Central Connectiondatabase (Centraal Aansluitingenregister) (C-AR) is a place where information is saved about the EAN-code. Within the C-AR a lot of information is stored from all the different connections. DSOs are the information suppliers to the C-AR. Energie Data Services Netherlands (Energie Data Services Nederland) (EDSN) is in charge of the C-AR because EDSN is owned by the different DSOs.With the C-AR the EDSN act as an Information manager and put together all the information within the chain. EDSN is non-profit and a trusted party in the chain. In figure 4-3 an overview is shown of what data is known in the electricity chain. As can be seen is that the C-AR contains the most information about an electricity connection. DSOs have most of their information already put in the C-AR. The DSOs have maps of the electrical system on how cables are placed within the infrastructure. But a lot of data is duplicated in the electricity chain. An example of the is adress and GPS. Both C-AR and KPN have different values for these 2 variables. All the actors are working with their own administration that synchronizes daily with the C-AR.

The parties communicate the master data of the different connections each time to each other. The master data has been specified in the information code. This code has been described in detail and by law the parties have to follow this.

Within each of these organizations there are also multiple databases, KPN for instance has different systems that ′know′ something about the connection. There are systems, for instance, that know everything about the cell phone towers and there are systems that know everything about the street cabinet (straatkast) locations in the street.

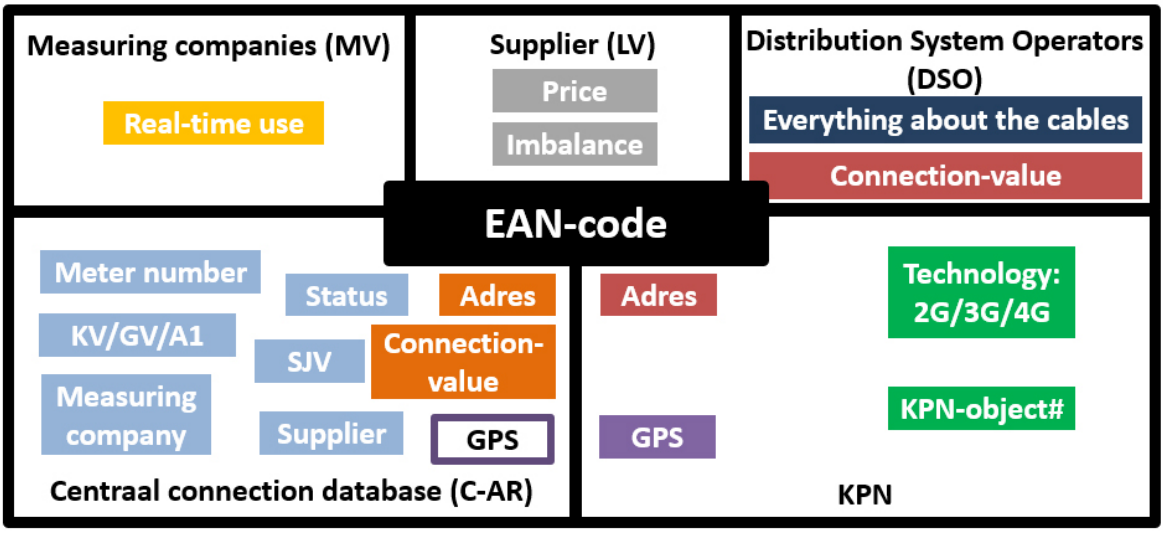


Figure 4-5: Who knows what?

4-3-2 Permissions of organisations to work with data

Not every organization in the electricity chain has the same access rights to the C-AR. Figure 4-5 gives an overview of the respective access rights. The DSOs have full access to the C-AR, they can create, read, update and delete entries. The DSOs have the most extensive rights because they are the owners of the EDSN and C-AR. The LV is rather limited in his options, they can only read information about the connections. For updating or deleting entries the LVs can make a suggestion to the DSO to do so. DSOs are responsible for the data and the mutations. A MV has access to the C-AR and can see information about all the connections.

The other DSOs, MVs and LVs that are not serving a particular EAN are not allowed to do anything. So a MV is not able to check data about connections that they are not measuring electricity for. LVs are also limited in their access to the C-AR, they cannot see events or data about what happened before they became the LV for that connection.

KPN does not have access to the C-AR. Users of electricity are never allowed entry to the C-AR. KPN can ask to the MV, DSO or LV for information from the C-AR if they want to. However the LV or MV are never able to make adjustments directly, this always goes through the DSO. An overview of the distribution of the access rights to the C-AR can be found in figure 4-5.

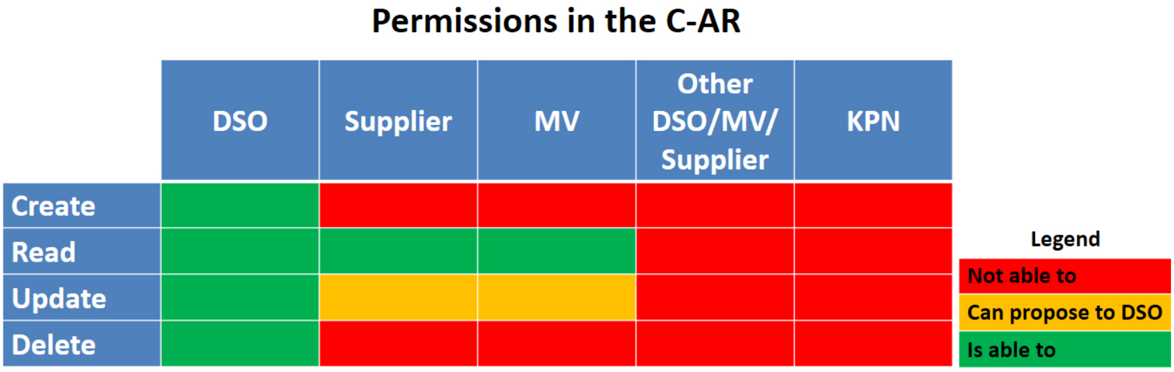


Figure 4-6: The different permissions that have been allocated to the different parties

4-4 Processes in the electricity chain

The electricity chain is complicated and multiple processes take place between the actors. In this paragraph the two most important processes concerning data are identified. These processes will play an important role later in the report.

4-4-1 A1-process

There are multiple data exchange steps that are carried out between the partners in the electricity chain (see figure 4-6). An analysis of the data for the A1 chain is given since this was most interesting for KPN. The majority of the connections at KPN are A1. A1 differs only a little bit compared to KV, the differences between these processes can be found in 4-1.

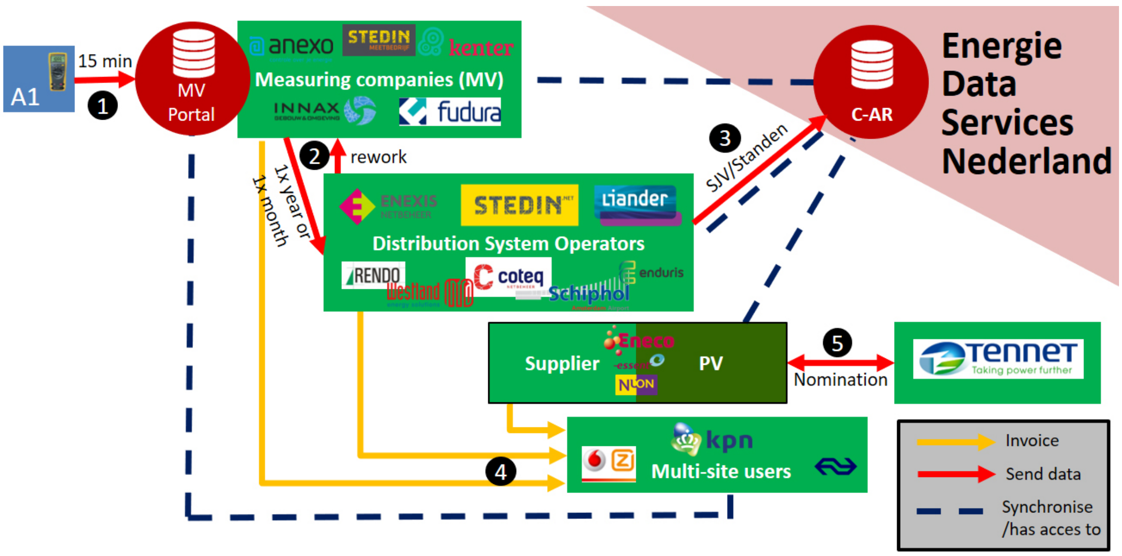


Figure 4-7: Overview of the current situation.

1. For KPNs A1 connections every 15 minutes the electricity use of the KPN-object is measured by the MV. This data is send from the meter to the MV, they save the data.

2. The data is sent once a year by the MV to the DSOs. DSOs need to check whether the usage data could be feasible. This process is done based on their historical data. When there is a dispute there will be a rework in which both parties look at the connection again.

3. After the MV and DSO have agreed on the usage data, the DSO updates it in the Central Connectiondatabase (Centraal Aansluitingenregister) (C-AR). Since the data is only being send once a year it is also only being updated once a year in the C-AR. The DSO sends the master data (stamgegevens), these are multiple variables about the connection (see figure 4-5). A history of the measured data is stored in the C-AR. The DSO also calculates a Standaard JaarVerbruik (SJV). This is a predictive value of the usage of electricity based on historic data. When there is no data available the SJV will be estimated by the DSO. According to the DSOs there is only a possibility to recalculate the SJV after 1 year of measurements. This is done in order to prevent a wrong calculation. Rules have been established in the Informatiecode [2] on how to calculate a SJV.

4. Based on what has been put in the C-AR, KPN receives invoices from the three parties (MV, DSO, LV) that have delivered services. For DSOs it is a matter of multiplying the connections times a fixed amount. KPN pays them monthly. The Measuring organisation charges per connection that is being measured. To the LV a monthly amount is transferred. Only 18 months after the year has ended Eneco is able to correct the payments to the correct amount of electricity that has been used in reality. Incorrect data in this steps leads to paying too much or too little.

5. The fifth step is where electricity has to be nominated. The nomination is needed for TenneT in order to stabilize the grid and forecast the demand. The nomination will be explained further into detail in paragraph 4-3-2.

Some important findings can be seen here already. The LV and the DSO have much less information about the electricity use compared to KPN and the MV. There is information asymmetry between the parties.

4-4-2 Nomination-process

In figure 4-7 the process of nomination can be seen. Every day this process is executed by the Programme-responsible (Programmaverantwoordelijke) (PV), in KPNs case this is Eneco. The PV has to report every day to Tennet how much electricity will be used by the different connections of the PV. Only the GV-connections can be based on real data. The KV and A1 connections are based upon estimations of what the electricity use will be. When there is a difference at the end of the day this can only be caused by GV-connections since there is no data available from KV and A1. However that does not mean that there was a difference in the values. Differences in the estimations (netverliezen) need to be paid by Eneco and thereby also indirectly by the customers like KPN. A instability on the grid was created since not all the data was used.

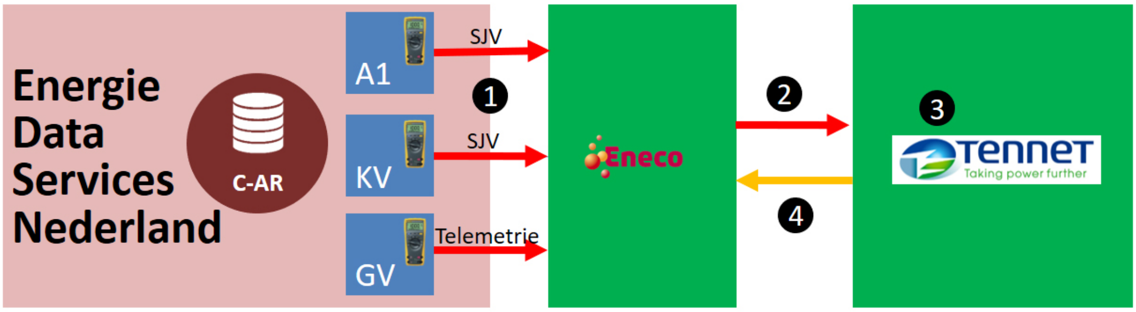


Figure 4-8: Overview of the nomination process

1. Eneco collects all the SJVs(estimations of electricity use) for the A1 and KV connections. The SJVs are estimates and can be retrieved from the C-AR. The PV uses different usage profiles in order to know how much energy will be used per fifteen minutes per connection. These connections are predefined and based upon average electricity use. Eneco uses telemetrie (real time measurements) to indicate how much electricity the GV connections will use.

2. The data is sent by the PV (Eneco) to TenneT. Tennet runs this process for all the PVs and knows how much power is needed for that day. They put the energy on the markets to be filled in by different producers of electricity.

3. At the end of the day TenneT calculates the difference between the plans and the actual electricity use. For KV and A1 connections there is no way to know what the actual electricity use was. The SJV is just a yearly estimation of how much electricity will be used by a particular connection. The LV/PV and DSO only receive information about the yearly electricity consumption (at the end of the year). When differences occur according to TenneT this can only have happened with GV connections since there is more information available.

4. TenneT makes the PV pay for the incorrect estimation that has been given for the GV connections. For the A1 and KV connections there is no actual data available and therefore no difference in the usage profiles can be found.

4-5 Conclusion

The second sub-question **What is the current state of the electricity chain data and processes?** can be answered. First the structure of the market is defining how the system operates. Multiple parties are involved in the delivery, measuring and production of electricity. There are different processes in the electricity chain that are executed among the parties. Two important processes are the nomination process and the process for A1-connections. Something that can be observed is that the different parties in the chain do not have access to the same amount of data. In the C-AR there is a lot of data available however the DSOs have a powerful position. The data can not be changed without their permission. For this reason other parties have their own administration. Within the chain also a lot of data is residing only at one party. DSOs for instance know a lot about the connection and about the cables. However the electricity usage rates are less known among the parties. The client and the MV have much more information than the LV/PV or DSO. The MV knows how much energy is being used every 15 minutes while a party like a DSO only gets the usage of electricity at the end of the year. Something that can be concluded is that DSOs have a important strong role with the access to a lot of data.

Chapter 5

Case study analysis: data quality problems in the electricity chain

Within this chapter an introduction to different data quality problems will be given. Different problems have surfaced after the input from interviews, data and insights from chapter 4 were combined. These different data quality problems will be further explained in 5.1. The dissection of the data quality problems will lead to the answering of the third research sub-question: **What are data quality problems in the electricity chain?**

5-1 Problems in the electricity chain

Since all parties in the electricity chain (LV, MV, DSO, Electricity user) have their own administration there are different truths out there. Consequently there are multiple obstacles to improve data quality and improve the exchange of data in the chain. All the parties have a large amount of information in their systems. The way to select relevant problems was through interviews and data analysis. For this sub-question multiple different stakeholders have been interviewed (see appendix A). During those conversations different problems surfaced that were a problem to them. All the information KPN, Eneco, MVs and DSOs have about KPN-connections was made available for analysis. The data was checked to see if there are differences between the administrations. Between the datasets also comparisons were made to check statements from the interviewees. All the important and relevant problems are mentioned below and divided in two categories. Data quality problems A to G are having an effect on the data quality. Data quality problems H to J are not impacting the data quality directly and therefore they can be found in appendix D.

5-1-1 Data quality problems in the electricity chain

A: Profiling of the connections

For KV and A1 connections there are five possible electricity usage profiles (E1A, E1B, E1C, E2A, E2B). These usage profiles reflect five dominant and characteristic use patterns of electricity. The DSOs are responsible to assign a profile to every KV or A1 connection. A large portion of the KV and A1 connections are all the same at KPN as they represent the same street cabinets or cell phone towers with identical electricity usage patterns. However the DSOs have assigned different profiles to KPN connections with the same standard pattern of electricity use. As can be seen in figure 5-2 the E1B profile has a big peak around 18:00. An average household behaves like this however A1 clients (like KPN) are different. Telecommunication companies have a flat user profile as can be seen in the figure 5-2. The networks need to be operational the whole day and more data use during daytime has little influence on how much electricity the equipment uses. So when it comes to KPN′s network there is a rather constant rate of electricity consumption. The implications are that for the nomination of electricity to be generated (see paragraph 4-4-2) the wrong amount of electricity is put on the grid through nomination. The blue surface in figure 5-2 displays the surplus amount of electricity that is allocated, purple shows the period with a shortage in allocation. So an imbalance is created on the grid and TenneT does not know exactly how this was created. TenneT does not have real time data on A1 and KV connections. This data is also unknown by the DSOs, LV and PV (see paragraph 4-4-1), only KPN and the MV know real time how much electricity is used on the grid.

The profiles also assume a different distribution when it comes to electricity use. As can be seen in figure 5-1 the electricity used through a E1B is much higher than through the average of KPN. A difference of almost 20% can be observed since the profiles assumes you use more electricity. The profiles assume that in winter (on 02-01) there is much more electricity consumption compared to a summer day. At KPN the contrary is the case. Most electricity of the locations is used for cooling the equipment. Therefore winter is a season where less electricity is used compared to summer. This phenomenon also contributes to a difference in the nomination of electricity for A1 and KV connections.

Profile	% of yearly use on 02-01
E1B	0,3335%
KPN	0,2793%
1/365	0,2740%

Figure 5-1: Overview of the electricity use per day

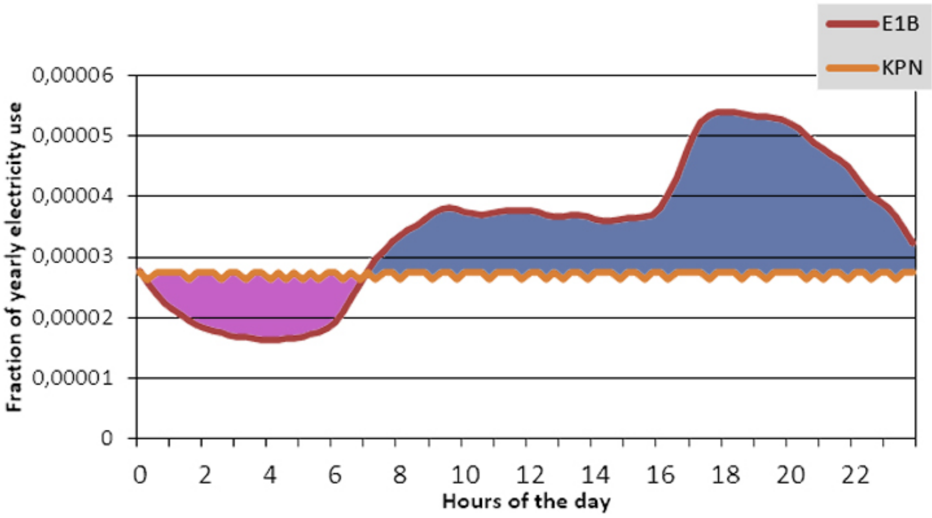


Figure 5-2: Overview of the electricity use per 15 min on 02-01-2017

B: Physical vs Digital truth

	Connections	Wrong value	Total cost
Stedin	1602	54	€ -23.375
Liander	2202	50	€ -18.205
Enexis	2324	60	€ 20.770
Total	**6128**	**164**	**€ -20.810**

Figure 5-3: Overview of the differences between the real and digital truth of connection values

The C-AR is the central data system containing information with specifications about the physical locations. All this information has been put in there. However in real life the connection value (aansluitwaarde) can be different from what has been reported by the DSOs. Since KPN has changed all their smart meters in the Netherlands they have obtained updated and actual information on their connection values. On average 2.6% of the data is incorrect. KPN should be paying 20.810 euro more for these connections. For this case we assume that the mechanics have been correct in their observations at the location. So what is the impact of this 2.6% faulty data? The main consequence is that not the right amount of money is being paid. Another important issue is that an electricity user can use electricity at higher rates than it pays for. In the case of KPN it has turned out to be profitable. However in the future this piece of low quality data could have implications when there is a rollout of a new technology. DSO should always be sure what the connection values are in their network.

C: Incorrect SJVs

The Standaard JaarVerbruik (SJV) is important in the entire chain. The data is used by different parties like Eneco to nominate electricity with Tennet. During the analysis some interesting facts were found. 4.8% of the connections have a use of 0 kwh. This is very likely to be incorrect. For these 805 connections (4.8%), KPN is paying currently too little. The other parties in the chain do not how much is used however 0 is rather unlikely. The implication of this is that the companies have to settle about hundreds of connections at a later point in time. Besides SJVs that were too small there were also a lot of SJVs calculated wrongly. Some of the connections that were found had even doubled. The SJV of a connection at Enexis was not calculated correctly and the numbers had doubled while the real electricity use stayed the same. After an inspection there is a likelihood that 80 other connections are having the same problem. KPN has been paying double the amount of energy for these connections. Also here the companies need to settle in a later moment in time. The SJVs are calculated separately at the DSOs. After having conducted the interviews the conclusion can be drawn that the calculation of the SJV is not uniform. The DSOs had their own way of calculating the SJV and there were no checks to whether SJVs had doubled over the course of just a few months.

D: Hijacking an EAN (Switches)

The Jonkersdam 2 in Oosterhout is an address with two EANs. Outside the house there is a street cabinet that provides internet to the street. So one EAN is from KPN and the other EAN is from the people that live in the house. But the people moved and therefore a new contract with a new supplier is chosen by the new inhabitants. The new supplier though takes over the EAN from KPN and connects it to the new inhabitants. The KPN-EAN now has a new owner and only after a few months parties find out this is wrong. Then a new connection has to be created. So within a system there is not a control about a new connection. So everybody can switch connections of others. Connections can be switched at all times. Only the connections of the DSOs can have a protection since they are a vital infrastructure.

Figure 5-4: The house with the street cabinet next to it.

E: Wrong customer, ZIP-code, Connection value

EAN 871694831000392159 is located in Rijssen and is owned by Reggefiber (daughter company of KPN) according to the C-AR. According to Enexis this connection is owned by KPN (not by Reggefiber). Enexis is wrong in this case with their own administration. The costs are being paid by KPN while Reggefiber uses the connection. Multiple of these errors have been found. The DSO has a different administration.

EAN 871691600001345043 is in Almelo on the Nachtegaalstraat according to the C-AR and Coteq. However the parties disagree about the house number and ZIP-code. According to Coteq it is 102-CAI001 and 7605 BR, while the C-AR reports 102 NST and 7605 BJ. Where CAI is a type of tv-connection and NST means next to it (naast). According to DSO Coteq EAN 871691600000002411 is located at Valduif 7, while C-AR thinks it is at Valduif 66. So both parties have different information in their registries. According to Rendo (a DSO) the connection value is 1x40 while the C-AR says it is 1x35. In this particular case there is a not a high impact however this shows that the parties have different administrations.

F: Incorrect values in the chain

Checks are not always performed in the entire chain. An example of this is EAN, this specific example is located at the Dorpsweg 4 in Reeuwijk. According to Stedin this location has used 3.000.000 kWh in 2015. They have send this data to the C-AR. Eneco was checking data and they found this to be an unusual number. This small street cabinet should only have an electricity use of about 3.000 kWh; a factor 1000 deviating from the calculated use by Stedin. If Eneco would not have noticed this mistake KPN would have paid the amount. Examples like this one can be found at all the DSOs. All the DSOs are separately responsible for the calculation of the SJV. Rules have been established in the Informatiecode [2] on how to calculate this. However in practise big differences can be seen.

Figure 5-5: The house with the street cabinet that almost used 1000 times too much electricity

G: Scaling down connections

When data is put together more value can be obtained. The connection values are contracted at the start of the period. Over time changes can be made in the equipment that KPN uses. Even though numerous parties have a big quantity of data nobody has advised KPN to do so. The customer should be proactive in the electricity chain otherwise nothing will happen. DSOs know how much electricity is used and they know the

Figure 5-6: Location in Houten

value of the connection. However DSO does not see itself as a responsible party to discuss opportunities with the customer. The customer should all do this themselves, the customer is responsible for the enlarging or making smaller of their connection. When the DSO advises something they are afraid that they will be prosecuted by the customer when something goes wrong and the electricity goes down for that particular connection.

5-2 Conclusion

The different problems that have been mentioned in this chapter are answering to the third research sub-question: **What are data quality problems in the electricity chain?** A variety of problems (A-G) has been identified after the interviews and analysis of the data. While some of the problems can be solved with one call others require more work to solve. All parties in the electricity chain (LV, MV, DSO, Electricity user) have their own administration. Because of these different administrations there are also different truths out there about an EAN. Consequently there are multiple obstacles to improve data quality and improve the exchange of data in the chain. Other problems include the potential that is not being utilized because data simply is not shared outside of the own organisation. There are different laws that are describing the electricity chain however details are lacking on how responsibilities should be picked up.

Chapter 6

Expert workshop: evaluation of recommendations for actors in the electricity chain

Together with experts from the energy field the problems are reviewed in chapter 6. So the fourth sub-question is answered: **How do data quality problems affect the electricity chain?** What elements that have been identified from literature can be applied to the case? After this question has been answered also the effects can be mapped.

6-1 Expert workshop on the data quality problems

The problems described in chapter 5 have been checked by experts in the field. In order to gain understanding different parties in the electricity chain have been consulted. Between 18th and 20th of September these experts were contacted again for a verification of the proposed remedies. The people that have attended the meetings can be found in table 6-1. The presentation that has been send to them can be found in Appendix C. The progress of the research has been shared and they were asked about the observations and solutions. Here below the different finding from the chain were discussed again with the participants. A lot of the problems that have been shown feel recognizable for the different participants. Their response is important to check the findings and is shown here below.

Table 6-1: Participants to the expert workshop

Role interviewee	Company	Category
Energy connection manager	Vodafone/Ziggo	Multi-site user
Energy expert	Rijkswaterstaat	Multi-site user
Energy expert	Anexo	Measuring company
Strategic Partner Manager	Eneco	Supplier
Accountmanager	Stedin	DSO
Accountmanager	Enexis	DSO
Energyteam-member	KPN	Multi-site user
Energyteam-member	KPN	Multi-site user
Energyteam-member	KPN	Multi-site user
Energyteam-member	KPN	Multi-site user

Here below the seven data quality problems were scored on the eight different data quality dimensions.

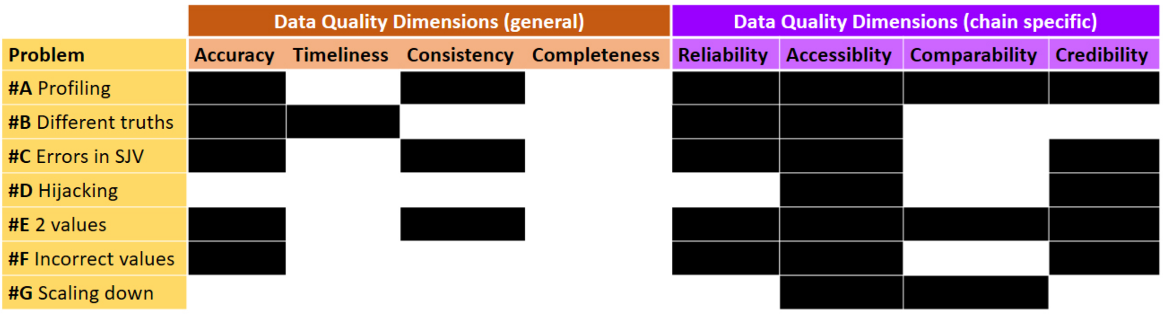


Figure 6-1: Data Quality Dimensions versus the different problems in the electricity chain

A: Profiling of the connections

The parties involved do mention that 75% of the electricity use is by GV-connections and this nomination goes well. But regarding the nomination for A1 and KV connections Tennet is faced with the consequences. They agree to the statement that the market is creating the losses themselves with this system. The profiling is not interfering with the accuracy. Also the consistency since different connections are assigned a different profile. Between the different DSOs there is not a uniform process to assign the profiles leading to inconsistencies. Accessibility plays a role here since KPN has to acces to the data and has to ask another player in the chain to provide the data. The credibility of the DSOs is therefore under attack since the multi-site users are not believing the profile data since is does not make any sense for them.

B: Physical vs Digital truth

Establish a process in order to reduce the amount of differences between the physical and digital truth. The analysis conducted in chapter 5 has shown that there is a difference in the digital truth and the physical truth. Parties in the electricity chain should join their forces since data is not being looked after, after it has entered the system. The data user is not the same as the data creator. Before a project starts everybody should be consulted. The connections from multi-sites are not checked regularly because they are hard to reach geographically. It is however essential to update certain values. At a later moment in time it is not possible to check the connection since it will be too costly. Bringing together the different stakeholders can have a good influence in order to quickly contact. At every moment in time there should be an agreement on the data and connections. All parties in the electricity data chain should be communicating with each other and also KPN should put in an effort to improve the data quality. This data quality problem has an impact on the accuracy, timeliness, accessibility and reliability. Multi-site users are not sure if the data is collected consistently over time. Through different mergers it could be that data has never been looked at for many years. This incurs a loss for at least one of the two parties. As has been shown in the calculations of chapter 5 this could add up to thousands of euros per year that KPN could potentially save.

C: Incorrect SJVs

Multiple parties in the electricity chain do not trust the DSOs in making these calculations. The experts confirm that the SJVs used for the allocations currently are not always correct. SJVs are calculated and updated by the DSOs. Something that is not taken into account is that they are not the actor who is relying on the data. TenneT for instance relies on the data that has been given to them. Likewise Eneco relies on these data in order to charge the customer. There is no incentive for the DSO to find the correct data. Not everybody is happy with the task that the DSOs are fulfilling. Participants suggested to liberate the market for specific actions. "The pain should be put where it is felt the most." was said by a participant. An example of this is where the calculations of the SJV do not bring benefits for the DSOs. There is no intrinsic motivation to improve this. So an extrinsic motivation could help much better. DSOs can earn money by making connections bigger than they actually are. So it is not in their interest to lower them or carry out more inspections.

After being faced with the errors on the SJVs they acknowledge it but do not feel the responsibility to improve this. Another issue before is that the SJV is being calculated by all the different DSOs independently. There is a shift to paying based on real numbers according to the DSOs so no need to worry about this. According to the DSOs the problems with the SJVs will be something of the past. With the smart meters there is not a need to improve the calculations. The DSOs understand that connections of Multi-site electricity users are rather difficult for multiple reasons.

There is no free market on electricity infrastructure. Some participants mentioned that this obstructs the market. The monopolistic character of the market in the case of distribution is necessary since it is inefficient to let multiple parties connect to the homes. DSO needs to keep losses to a minimum but it has not been privatised and the government is the owner. So therefore they are not receiving any incentive to do anything about it. The errors in the SJVs make that the accuracy of the calculation is doubted. Since the DSO have problems explaining the calculation and to find mistakes, multi-sites remain sceptical of the values that are stored in the C-AR.

D: Hijacking an EAN (Switches)

A solution for the switching could be to lock the A1 connections. A1 connections almost never change owner. When the connection is closed off a special permission from the supplier or electricity user could be asked. By doing this a lot of work can be avoided. Asked about the switching DSOs see the problem. Independently the DSOs do not have the power to change. Forces between DSOs and organisations like VEMW should be joined in order to make changes in the future. So actually you should be able to approve/disapprove a switch and not let the other party take over. The credibility of the data is important since the messages of the hijack are easy to distinguish from being a mistake. However the system does not pick this up and it has to be corrected manually.

E: Wrong customer, ZIP-code, Connection value

There is no active maintenance on the data. DSOs do not feel responsible however they are open for suggestions by the customer on what should be changed. However in practise this was not as easy as they said. Multiple times I have asked them to make changes regarding incorrect data about a KPN connection. Sometimes they did not respond or it was a lot of work to change one variable from one connection.

The responsibility to make changes is put outside of the DSOs. It is not in the interest of the DSOs to work on this. This will only mean that they lose money, therefore there is no incentive to improve the data about connections. Ziggo is having the same issues where postal codes are not working. While things like ZIP-code should be easy to correct on a central level it is not done.

Comparability is the most important here since it is easy to spot that two organisations have 2 different values. The process however of correcting it costs much more time and therefore is not initiated by the DSO itself. A lot of discussion among the experts was about this point about responsibility. Where does it lie to correct this? With EDSN? With the customer? The customer of course is not possible since the customer has no access to the incorrect data. But these differences have been created since the parties keep sticking to their own administration even though there is higher quality data available.

F: Incorrect values in the chain

There are multiple ways to ensure that checks are performed. Currently this is not always the case. There should be checks however there are still possibilities for failures. Since there are eight DSOs it is hard to control and unify the process. There are teams to control this even though obvious mistakes could be prevented. So the experts are aware of the effect that certain checks are not being executed. It was worrying to hear that in the opinion of the DSO the customer is responsible while it does not have access. The accuracy and reliability are 2 important factors that are at risk when there are no additional checks executed.

G: Scaling down connections

Ziggos policy is to have a higher connection value in order to be safe. At KPN there is a margin of 20-30% in order to stay safe. For both this is causing a lot of costs. The example of a single connection in Houten where 3000 euro per year could be saved was interesting. Centralizing data will help to spot the differences between processes. The project that is coming up (Central ARM) could be helping out a lot in improving data quality. A centralization will only work if the other parties let go of their old information that could potentially be conflicting. There is a central system but the data shows that all parties still keep their own data that could potentially be different from the central data. So currently it is still necessary to obtain information from different parties. Here it is important that the comparability should be facilitated and data should be analysed. Once data is brought together it is easy to see what should be done. All parties do not feel responsible to tell KPN that they could save a big amount of money at certain connections. They will only lose income when they do so. The experts in the call agreed that the energy industry is not the most client friendly markets. A lot of responsibility from the client is expected.

6-2 Conclusion

The different problems that have been mentioned in this chapter are answering to the and fourth research sub-question: **How do data quality problems affect the electricity chain?** The different categories of data quality problems affect the electricity chain in different way. The assigning of profiles for instance is a process that directly influences the electricity that is allocated on the grid. Other data quality problems are having a direct effect on the costs. Incorrect data is leading to higher or lower bills and companies need to settle later. Through the research on the data it can be said that multiple mistakes are in the electricity chain since very little checks are performed once the data is in the systems.

Chapter 7

Recommendations for policy makers

The fifth and final research sub-question is: **How to improve data quality in the electricity chain?** This leads to the policy recommendations in 7.1 that could improve the Data quality and processes within the entire electricity chain.

7-1 Policy options

What are policy recommendations that could improve the data quality in the chain? What is needed to align the parties and demarcate the responsibilities? Within this a few policy recommendations are drafted based upon earlier identified problems.

The effect of the policy option is on the different Data Quality Dimensions that have been in defined in chapter 3.

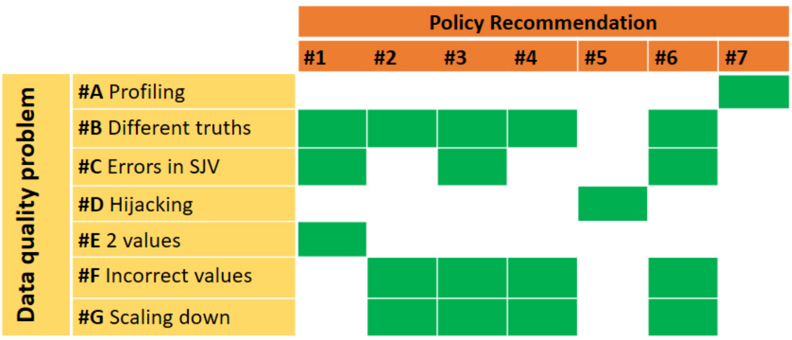
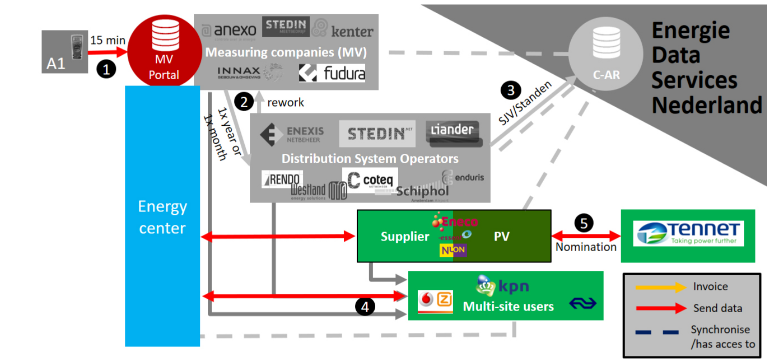


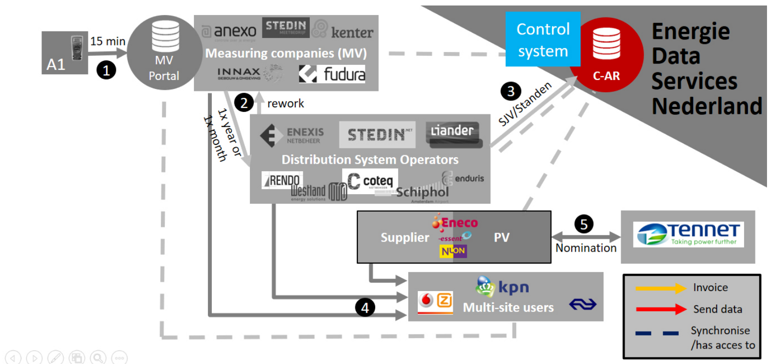
Figure 7-1: Impact of the policy recommendations on the different data quality problems

Policy recommendation 1: Start a shadow administration



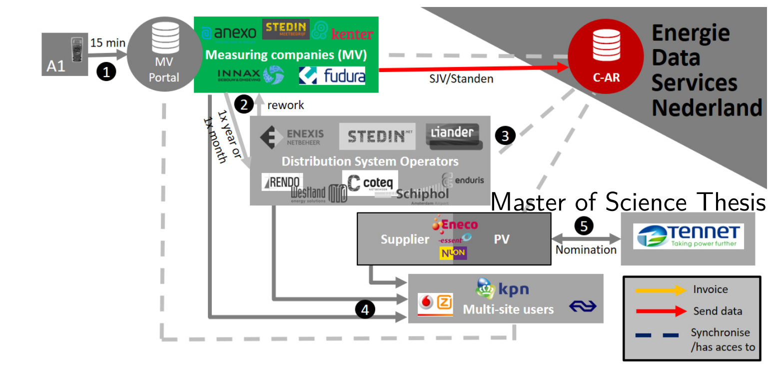
Currently calculations of the SJV is done by the DSO. The DSO could potentially be bypassed and a new administration could be started. This new administration could make sure the calculation of the SJVs will be improved, leading to a better working of the system. KPN and Eneco could join forces to improve the data and the calculations. KPN has access to a lot of usage data of their connections. Eneco has a lot of expertise on the electricity market. Why would Eneco be willing to execute more work that is outside of their responsibilities? Correct numbers could save Eneco a lot of money. When SJVs are incorrect Eneco has to pay. Eneco is currently not sure of the real usage data and they have to settle the bills with KPN after a lot of time. Eneco could also save a lot of money on imbalance-premiums that they currently have to pay to Tennet. This construction is not illegal since the (other) data will still be stored in the C-AR however they will not be used to balance the Dutch electricity grid. A better function could be that the responsibilities are transferred to the supplier and energy user.

Policy recommendation 2: Develop a central control at the C-AR



The DSO is doing the calculations for SJVs while it is not in their interest. Another possible recommendation therefore is to let analyses run on the C-AR. An issue with DSOs is that there are many of them and that they do not treat data the same way. When there is a department at EDSN that is in charge of improving the data quality all members could potentially benefit. This solution is already in the making. The EDSN is creating a system that could uniform and standardize processes.

Policy recommendation 3: Allow the Measuring companies to enter data in the C-AR



A third option is to also make the Measuring companies responsible for the correct input of data. Currently they are not allowed to enter this data into the C-AR. Measuring companies have all the data that is needed to make correct calculations. These companies also have a context on how much data when is used. When this data is updated properly it becomes easier for Eneco to work with the data. It is not in the interest of the measuring company to increase or decrease the numbers, they are paid per connection, not for the amounts of electricity.

Policy recommendation 4: Let the MV and Supplier connect

Currently it is not allowed according to the law to let the supplier and MV connect and share data. Eneco is not receiving the data directly from the MV, first it needs to pass the DSO. This is a odd situation since the DSO does not have a high interest in how much electricity you use. With this policy option it becomes easier to just pay for your own electricity usage. The ties between the supplier and the MV could lead to a quick process where it is in the interest of both to provide data that is of high quality.

Policy recommendation 5: Lock all A1-connections in the C-AR

Currently it is not possible to stop parties from switching A-1 connections from supplier. There is already the possibility with the C-AR to lock A1-connections, which can be an advantage as A1 connections are connections that rarely change owner. DSOs, however, are currently not allowed to do so. Therefore it is suggested to lock these A1-connections so the EAN can not be changed to a new supplier without approval from the current supplier. The extra work that this policy recommendation will imply has to be executed by the suppliers. It will also save the parties a lot of work when incorrect EANs are not hijacked anymore. So DSOs should be allowed to lock these connections. They know which connections are A1 since they are the owners of the data.

Policy recommendation 6: Give multi-site users access to the C-AR

Currently initiatives have to come from KPN (the customer). When the customer does not take action nothing will change. Therefore collaboration and service should be stimulated within a chain. The benefits and costs are not equally spread in a chain and therefore it can be harder to move an actor in a desired position. The data in the C-AR is owned by the EDSN and EDSN is owned by the DSOs. In the C-AR users of electricity do not have access to see anything while DSOs are the only parties able to make changes. Bigger customers like VodafoneZiggo, NS and Rijkswaterstaat are currently not involved in the processes. Lots of efforts and money could be saved when they are involved since they can have a big influence. The multisite users already have the data and are willing to correct information. However this is currently not possible since the customer is not able to see the data without consulting one of the other parties like a MV, LV or DSO. Redefining the access policy for the C-AR will be making sure that multi-sites will be more involved in changing their own data.

Policy recommendation 7: Allow flexibility in profiles for A1-connections

The current choice of profiles is based upon households or public lighting. A constant rate at which KPN uses its electricity is not facilitated in the current profiles. So what is needed? A new profile that fits the needs of A1 customers. Freedom should be given to the parties in the chain to create new profiles. These new profiles needs to be implemented and assigned to the connections by the different DSOs. Tennet and the suppliers will be very supportive towards this subsection. Tennet will be the biggest beneficiary of this measure. The data from the suppliers will be more reliable and closer to the real usage. Therefore Tennet will be benefitting since the imbalance on the grid will be smaller. An issue that might arise is that DSOs have to execute more work however are not receiving any benefits from this. The extra profile will not lead to much more work, the assigning of the profiles will happen anyway. With the energy transition also the amount of irregular electricity use will grow. Currently there are pilot projects in the Netherlands to use electric vehicles (EV) or smart washing machines to peakshave. By peak shaving the electricity consumption is spread more evenly over the day and in combination with solar panels profiles might even be delivering back to the grid. So besides the necessity for the industrial A1 users it is also essential to implement this for the regular customers.

7-2 Conclusion

The fifth and final research sub-question is: **How to improve data quality in the electricity chain?** In this chapter 7 policy recommendations are drafted up that could have a positive impact on the data quality problems that have been explored earlier. Parties in the electricity chain do not talk a lot about data quality. Contacts between organisations in the chain are rather practical and usually concentrate on singular EAN-connections. Taking a step back and looking at how data process can be improved can be beneficial for the different parties. Furthermore there are not a lot of checks that are executed. Multiple mistakes are in the system since very little checks are performed once the data is in the system.

Chapter 8

Conclusions

In this chapter the conclusions will be drawn based upon the research. The research questions will be answered. The conclusions will be applicable to users (KPN and other Multi-sites) as well as facilitators (DSO, TSO, Supplier) in the chain. In 8.3 multiple points will be brought up for future research. In 8.4 multiple points of discussion will be brought up about this research. A personal reflection will be concluding this chapter at 8.5.

8-1 Answers to the research sub-questions

How can data quality be expressed in a chain? Expressing data quality within one single organisation is a tough job. Expressing data quality for multiple organisations add additional complexity. In chains this extra complexity consists out of multiple interests, responsibilities and interdependencies between organisations. The value of the data increases when it is of higher quality since more information can be deduced from it. This is vital for the costs, innovation and stability of the electricity infrastructure. There are multiple ways to describe data quality. For this research the first four data quality dimensions are: accuracy, completeness, timeliness and consistency. However what has been seen in literature is that a lot of literature does not take into account the multiple organisations that can be involved. Therefore four other Data Quality Dimensions (credibility, accessibility, reliability and comparability) have been added to the framework. These four extra DQDs that make data fit for use should be connecting better to the data quality problems in the chain.

What is the current state of the electricity chain data and processes? First the structure of the market is defining how the system operates. Multiple parties are involved in the delivery, measuring and production of electricity. There are different processes in the electricity chain that are executed among the parties. Two important processes are the nomination process and the process for A1-connections. Both processes are about data exchanges that helps in the executing of producing and stabilizing the Dutch electricity grid. Something that can be observed is that the different parties in the chain do not have access to the same amount of data. In the C-AR there is a lot of data available however the DSOs have a powerful position. The data can not be changed without their permission. For this reason other parties have their own administration. Within the chain also a lot of data is residing only at one party. DSOs for instance know a lot about the connection and about the cables. However the electricity usage rates are less known among the parties. The client and the MV have much more information than the LV/PV or DSO. The MV knows how much energy is being used every 15 minutes while a party like a DSO only gets the usage of electricity at the end of the year. Something that can be concluded is that DSOs have a important strong role with the access to a lot of data.

What are data quality problems in the electricity chain? A variety of problems has been identified after the interviews and analysis of the data. Problems regarding the calculation of electricity use, nomination of electricity on the grid, sharing of important data are a few examples. While some of the problems can be solved with one call others require more work to solve. All parties in the electricity chain (LV, MV, DSO, Electricity user) have their own administration. Because of these different administrations there are also different truths out there about an EAN. Consequently there are multiple obstacles to improve data quality and improve the exchange of data in the chain. Other problems include the potential that is not being utilized because data simply is not shared outside of the own organisation. There are different laws that are describing the electricity chain however details are lacking on how responsibilities should be picked up.

The different problems that have been mentioned in this chapter are answering to the and fourth research sub-question: **How do data quality problems affect the electricity chain?** The different categories of data quality problems affect the electricity chain in different way. The assigning of profiles for instance is a process that directly influences the electricity that is allocated on the grid. Other data quality problems are having a direct effect on the costs. Incorrect data is leading to higher or lower bills and companies need to settle later. Through the research on the data it can be said that multiple mistakes are in the electricity chain since very little checks are performed once the data is in the systems. Calculating the SJV is at the wrong place currently with a party (DSO) that gains nothing from the correct execution. In this thesis multiple fields have been identified that could improve the overall data quality in the chain. Once data is entered it is rarely checked. This can be partly explained by the fact that it is rather costly to check this manually.

How to improve data quality in the electricity chain? Policy recommendations are drafted up that could have a positive impact on the data quality problems that have been explored earlier. The majority of the policy recommendations are all aimed at redefining the responsibilities for the parties in the chain. Putting the responsibility closest to the party that has the biggest interest in good data quality is therefore essential. Parties in the electricity chain do not talk a lot about data quality. Contacts between organisations in the chain are rather practical and usually concentrate on singular EAN-connections. Taking a step back and looking at how data process can be improved can be beneficial for the different parties. Furthermore there are not a lot of checks that are executed. Multiple mistakes are in the system since very little checks are performed once the data is in the system.

8-2 Limitations

This case study has been applied to the Dutch energy market. Data quality is essential for all parties and universal. However this thesis is aimed at the dutch energy sector. Regional differences therefore are not taken into account since markets in Europe are structured differently.

In further research the selection of experts could be expanded. The experts that have been consulted for this thesis were mainly working with KPN. Therefore a biased view could arise since they might provide different insights and perspective on the case. Advice for further research would be to involve more suppliers, TenneT and EDSN. Because of limitations in the agenda it was not possible for me to include these actors.

A common critique on a case study is that the data is not systematically treated. The writer is able to influence the process heavily. Therefore it is not possible to generalize the research scientifically. These critiques can also be assigned to this thesis since it is hard to reproduce scientifically.

8-3 Recommendations for further research

Three different suggestions for further research will be given. These three topics for further research are a selection of the most important things that have been mentioned during the interviews.

Data quality as a service

When a party puts in effort in improving data others will reap the benefits. In this research no emphasis was placed on the costs. Who should pay the costs that are incurred while improving data? The user, the facilitator or the handler for instance. This research has shown that not everybody has a focus on data quality.

Blockchain

In the discussions with experts the question arose whether new technologies could provide a solution. Blockchain was raised as an opportunity. Blockchain has become popular since there are certain characteristics that make it interesting. The system is a network consisting out of multiple locations recording transactions. These transactions are saved in a single way so there cannot be a dispute among parties. Within the energy chain where multiple organisations are working together this could provide an outcome.

8-4 Connection to COSEM

The energy-industry is a complicated industry to comprehend. There lies a world of data exchange, information gathering and accounting behind the production and delivery of electricity. Therefore it took more time than expected to understand the individual responsibilities and dependencies of the energy chain. Also the standpoints of the different actors was confusing at times since two people (within the same organisation) could have an entirely different opinion.

Something I learned during the conducting of the thesis is that it can be hard to have a common understanding among different-minded parties. Interviewees are biased in the answers they gave. Some organisations said all their data and processes were flawless but data was telling a different story. Since multiple parties have been interviewed multiple perspectives/-claims were heard. These perspectives made it easier for me to understand the complexity that is present. Something like delivering electricity may seem easy but there is more than meets the eye.

8-5 Personal reflection

This thesis has been a big challenge and has made me much more aware of my who I am and what I am good at. Executing the research and talking to the different people in the energy chain was something that energized me and was an enjoyable process. Being able to structure the research and being the writer of your piece was a challenge. I have learned a lot from all the comments I got throughout the process. To do a case study that involved different parties was very important to me because only then I would able to put all the things, I learned during the study, into practise. I am proud about the fact that I have been able to get an insight in the energy chain. Looking back I think I raised the awareness among the parties that there is a need to improve data. Where I initially was met with a little scepticism this decreased throughout time. So all in all it has been an experience that has not been easy but I definitely learned a lot from it.

Bibliography

[1] Russell L Ackoff. From data to wisdom. *Journal of applied systems analysis*, 16(1):3–9, 1989.

[2] ACM. Informatiecode electriciteit en gas. https://www.acm.nl/nl/publicaties/publicatie/14457/Informatiecode-Elektriciteit-en-Gas, 2016.

[3] H. Afshari and Sh Khosravi. Information resource Management Maturity Model. *World Academy of Science, Engineering and Technology*, 37(1):258–262, 2009.

[4] APX. Electriciteitsmarkt, 2017.

[5] Nicola Armaroli and Vincenzo Balzani. Solar electricity and solar fuels: status and perspectives in the context of the energy transition. *Chemistry-A European Journal*, 22(1):32–57, 2016.

[6] Danielle GT Arts, Nicolette F De Keizer, and Gert-Jan Scheffer. Defining and improving data quality in medical registries: a literature review, case study, and generic framework. *Journal of the American Medical Informatics Association*, 9(6):600–611, 2002.

[7] Carlo Batini, Cinzia Cappiello, Chiara Francalanci, and Andrea Maurino. Methodologies for data quality assessment and improvement. *ACM computing surveys (CSUR)*, 41(3):16, 2009.

[8] ING In Business. Green data: Interview with dutch tso tennet's head of digitisation renÃľ kerkmeester. https://www.tennet.eu/fileadmin/user_upload/ArtikelTenneT.pdf, 2017.

[9] Li Cai and Yangyong Zhu. The challenges of data quality and data quality assessment in the big data era. *Data Science Journal*, 14, 2015.

[10] CBS. Elektriciteit in nederland. https://www.cbs.nl, 2015.

[11] CBS. Aandeel hernieuwbare energie 5,9 procent in 2016. https://www.cbs.nl/nl-nl/nieuws/2017/22/aandeel-hernieuwbare-energie-5-9-procent-in-2016, 2016.

[12] José Pablo Chaves-Ávila, Reinier AC van der Veen, and Rudi A Hakvoort. The interplay between imbalance pricing mechanisms and network congestions–analysis of the german electricity market. *Utilities Policy*, 28:52–61, 2014.

[13] Wen Chen, Kaile Zhou, Shanlin Yang, and Cheng Wu. Data quality of electricity consumption data in a smart grid environment. *Renewable and Sustainable Energy Reviews*, 75:98–105, 2017.

[14] Brendan Cook, Jerrome Gazzano, Zeynep Gunay, Lucas Hiller, Sakshi Mahajan, Aynur Taskan, and Samra Vilogorac. The smart meter and a smarter consumer: quantifying the benefits of smart meter implementation in the united states. *Chemistry Central Journal*, 6(S1):S5, 2012.

[15] Ian Dobson, Benjamin A Carreras, Vickie E Lynch, and David E Newman. Complex systems analysis of series of blackouts: Cascading failure, critical points, and self-organization. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 17(2):026103, 2007.

[16] EU. Smart metering deployment european union. http://ses.jrc.ec.europa.eu/smart-metering-deployment-european-union, 2017.

[17] Ragib Hasan, Radu Sion, and Marianne Winslett. Introducing secure provenance: problems and challenges. 2007.

[18] Benjamin T Hazen, Christopher A Boone, Jeremy D Ezell, and L. Allison Jones-Farmer. Data quality for data science, predictive analytics, and big data in supply chain management: An introduction to the problem and suggestions for research and applications. *International Journal of Production Economics*, 154:72–80, 2014.

[19] Forest Woody Horton. *The Information Management Workbook: IRM Made Simple: a Practical Step-by-step Guide that Explains how to Translate Information Management Concepts Into Practice, and to Make Them Work for Your Company, Government Agency, Hospital, College, Association Or Other Enterprise*. Information Management Press, 1981.

[20] IBM. The four v's of big data. http://www.ibmbigdatahub.com/infographic/four-vs-big-data.

[21] Health Information and Quality Authority (Ireland). International review of data quality. 2011.

[22] Eerste Kamer. 34199 electriciteits en gaswet. https://www.eerstekamer.nl/wetsvoorstel/34199_elektriciteits_en_gaswet?start_docList=100#p3, 2017.

[23] John W Kingdon and James A Thurber. *Agendas, alternatives, and public policies*, volume 45. Little, Brown Boston, 1984.

[24] Barbara Kitchenham. Procedures for performing systematic reviews. *Keele, UK, Keele University*, 33(2004):1–26, 2004.

[25] Corinna Klessmann, Christian Nabe, and Karsten Burges. Pros and cons of exposing renewables to electricity market risksâĂŤa comparison of the market integration approaches in germany, spain, and the uk. *Energy Policy*, 36(10):3646 – 3661, 2008.

[26] KPN. Jaarverslag 2016. https://jaarverslag2016.kpn, 2016.

[27] Rolf Künneke and Theo Fens. Ownership unbundling in electricity distribution: The case of the netherlands. *Energy Policy*, 35(3):1920–1930, 2007.

[28] D. Laney. 3d data management: Controlling data volume, velocity, and variety. http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-VolumeVelocity-and-Variety.pdf, year = "2001 (accessed September 23, 2017)".

[29] In Lee. Big data: Dimensions, evolution, impacts, and challenges. *Business Horizons*, 60(3):293–303, 2017.

[30] Vasco Lenzi, Andreas Ulbig, and Goran Andersson. Impacts of forecast accuracy on grid integration of renewable energy sources. In *PowerTech (POWERTECH), 2013 IEEE Grenoble*, pages 1–6. IEEE, 2013.

[31] Stuart E Madnick, Richard Y Wang, Yang W Lee, and Hongwei Zhu. Overview and framework for data and information quality research. *Journal of Data and Information Quality (JDIQ)*, 1(1):2, 2009.

[32] Thomas C Redman. The impact of poor data quality on the typical enterprise. *Communications of the ACM*, 41(2):79–82, 1998.

[33] J.E. Ross and S. Perry. *Total Quality Management: Text, Cases, and Readings, Third Edition*. Business/Quality management/Manufacturing. Taylor & Francis, 1999.

[34] Monica Scannapieco, Paolo Missier, and Carlo Batini. Data quality at a glance. *Datenbank-Spektrum*, 14(January):6–14, 2005.

[35] Stedin. Jaarverslag 2016. https://www.stedin.net/over-stedin/jaarverslagen-en-publicaties, urldate = 2017-08-30, 2016.

[36] Tennet. Jaarverslag 2016. http://annualreport.tennet.eu/2016/annualreport, 2016.

[37] VEMW. Energietransitie wet ve-klimaatwet. https://www.vemw.nl/Nieuwsoverzicht/2017-04-06-Energietransitie-WetVET-Klimaatwet.aspx, 2017.

[38] Richard Y Wang. A product perspective on total data quality management. *Communications of the ACM*, 41(2):58–65, 1998.

[39] Richard Y Wang and Diane M Strong. Beyond accuracy: What data quality means to data consumers. *Journal of management information systems*, 12(4):5–33, 1996.

[40] Richard Y Wang and Diane M Strong. Beyond accuracy: What data quality means to data consumers. *Journal of management information systems*, 12(4):5–33, 1996.

[41] Bo Pedersen Weidema and Marianne Suhr Wesnaes. Data quality management for life cycle inventories-an example of using data quality indicators. *Journal of Cleaner Production*, 4(3-4):167–174, 1996.

[42] Coralyn W Whitney, Bonnie K Lind, and Patricia W Wahl. Quality assurance and quality control in longitudinal studies. *Epidemiologic Reviews*, 20(1):71–80, 1998.

[43] Robert K Yin. *Case study research: Design and methods*. Sage publications, 2013.

[44] Robert Zavadil. Renewable generation forecasting: The science, applications, and outlook. In *System Sciences (HICSS), 2013 46th Hawaii International Conference on*, pages 2252–2260. IEEE, 2013.

Appendix A

Conversations with involved parties in the energy chain

People I have spoken to:

Table A-1: My caption

Company	Category	Role interviewee	Date of interview
NS	Multi-site user	Energy advisor	12-06-2017
Vodafone/Ziggo	Multi-site user	Energy connection manager	21-06-2017
Rijkswaterstaat	Multi-site user	Energy expert	07-06-2017
Anexo	Measuring company	Energy expert	01-06-2017
Eneco	Supplier	âĂŐ Strategic Partner Manager	10-08-2017
Stedin	DSO	Accountmanager	25-07-2017
Liander	DSO	Accountmanager	22-08-2017
Enexis	DSO	Accountmanager	18-07-2017

A-1 Multi-site users

A-1-1 Nederlandse Spoorwegen (NS)

The NS is using a lot of power for their trains. Like KPN this company has connections to the electricity grid all over the Netherlands. My reason to interview the NS was to check their role as a multi-site user in the Energy-chain.

Nederlandse spoorwegen informatie

Christel Franken is al 10 jaar met Energie bezig bij de NS. De NS is organisatorisch opgesplitst in 2 delen: NS Stations en NS Reizigers. NS Reizigers heeft een contract voor tractiestroom (om de treinen te laten rijden) met XXX. NS Stations hebben een contract met XXX voor verlichting, kantoorgebouwen en stations. Bij beiden is de stroom 100% groen. Christel zit bij NS Stations (stations ontwikkeling en onderhoud) en daar is energiemanagement bij ondergebracht. 3 mensen onderhouden de contracten met de leveranciers (XXX voor Gas en XXX voor Electriciteit). Daarnaast verlenen ze energiedienstverlening aan de verschillende afnemers (Prorail, Nedtrain, NS, Structon). Vanuit de geschiedenis heeft de NS de locaties van de Stations daardoor is dit ook een aparte entiteit.

Slimme meters, Verbruik en Aansluitingen

NS Stations heeft ongeveer XXX aansluitingen: XXX voor elektriciteit, XXX voor gas, XXX voor water. Er zijn XXX stations die allemaal al slimme meters hebben. 2007-2015 is de grootste bulk aan slimme meters geplaatst, nu worden de laatste meters vervangen. Veel is afhankelijk van de afspraken met de netbeheerder. Grootste gedeelte van de aansluitingen is A1.

Mutaties

NS heeft vrij veel mutaties (XXX tot XXX per week op elektra), dit komt door nieuwe aansluitingen op stations, verbouwingen op stations (voorbeeld: Eindhoven is XXX meters). Veel tijd gestopt in het op order krijgen van het aansluitingenbestand. Dat werpt nu de vruchten af. Ze weten nu precies welke aansluitingen ze hebben. Er is nu een duidelijke manier van mutaties doorgegeven. Alles in genomen waardoor mutaties altijd door NS worden geregeld. Via Excel leveren ze de mutaties aan. Sinds een paar jaar hebben ze afspraken over wanneer ze mogen beginnen of eindigen met factureren. Elke maand een maandrapport van wijzigingen, volumes en aansluitingen van XXX en XXX. Ook al is facturatie uitbesteed, ze zijn er alsnog tijd aan kwijt. Duidelijke instructies naar XXX toe over afwijkingen en hoe ze moeten werken. Ze hebben een systeem waar alle gegevens samenkomen, facturen, verbruiken etc.. Zo kunnen ze het werk van XXX toetsen. In de keten gaat er iets fout omdat er soms waardes niet kloppen aan het begin en aan het einde.

Bij de verandering van gasleverancier duurde het bijvoorbeeld lang voordat de keten weer alle informatie goed voorhanden had. Het is dus belangrijk daar dicht bovenop te zitten. Wat ze doorbelasten naar eindverbruikers is 1 op 1 wat er verbruikt word volgens XXX.

Verschillende methodieken zijn er gebruikt om hier toe te komen. Het factureren staat los van wat ze aan de eindverbruiker factureren. XXX verzamelt alle contracten en rekenen uit wat het het kost. Achteraf kijken ze of het klopte met de netbeheerders etc.. NS weet dingen sneller en secuurder dan partijen midden in de keten. Ze praten met alle netbeheerders maar niet tegelijkertijd want ze willen vaak niet met elkaar praten. Dat kost wel tijd.

Heeft 4 jaar geduurd voordat ze met Netbeheerders Nederland aan tafel konden voor een aanpak van A1 aansluitingen. Dit laat maar zien dat iedereen andere systemen heeft en dit anders inricht.

Facturen

Alle aansluitingen zijn A1 of grootverbuik, NS krijgt die per maand gefactureerd (afgesproken met Leverancier). Netbeheerder krijgen ze soms een verzamelnota, soms niet, soms een acceptgiro. XXX meet bijna alles voor de NS en zij doen ook de factuurverwerking. XXX besteed dat weer uit aan XXX. Hier zijn duidelijke afspraken over gemaakt. NS kan iets anders handig vinden dan bijvoorbeeld XXX. NS stations betaalt en XXX ontvangt alles. Met XXX en XXX hebben ze afspraken gemaakt over de kosten. Zij kregen de instructie om alles in te delen op de verschillende kostenposten (leverings, transport, trafo en overig). Dat stemmen ze periodiek af als er iets niet past. NS speelt in actieve rol in de energieketen. NS prefereert ook vaak de pdf-factuur aangezien ze dat soms moeten aantonen naar de uiteindelijke afnemers. Daarnaast zijn er nog een XML/CSV bestanden die worden gestuurd. XXX de core business is Energieadministratie en doorbelasting, nog veel met de hand. Het kan bijvoorbeeld stuk per stuk qua facturen nodig. NS heeft niet de expertise en kennis om dit allemaal zelf te gaan doen. NS is meer de linking pin/contractmanager hierin. Elke 2 weken zitten XXX en XXX 1 hele dag bij de NS. XXX en NS zitten eens per maand om tafel. De netbeheerders zitten elk kwartaal met de NS aan tafel. Onderwerpen zijn dan vooral mismatches die plaatsvinden. Waar kijk je naar en hoe doe je dit? Door de kennis die Christel heeft opgebouwd kent ze alle eigenaardigheden van het netwerk. Er word nog op geschatte jaarverbruiken afgerekend. Met XXX is afgesproken dat ze elke maand afrekenen. Ze hebben ook een connectie met het C-AR. Via XXX zet die om dat afrekening maandelijks moet gaan plaatsvinden. Gaat op dit moment nog niet lekker terwijl NS de echte gemeten waardes heeft. Er word in de keten soms nog steeds op de SJVâĂŹs afgerekend. Zorg er voor dat je dit maandelijks afrekent en laat het checken. Je moet wel in de gaten houden dat de data van het meetbedrijf controleert met de rekeningen.

A123

Als er een dienstwoning is die bijvoorbeeld binnenkort verkocht gaat worden doen ze de omzetting niet. Als je dan omzet naar A1 dan moet het fysiek worden teruggezet naar kleinverbruik. Daarna kan je het nooit meer naar A1 brengen. Gevolg: regionale lijnen (Hoekse lijn) blijven op kleinverbruik.

Problemen

- KPN en NS hebben geen toegang tot het C-AR. Er is geen duidelijke afbakening van wie wat mag veranderen in het C-AR. Concreet voorbeeld: bij alle grootverbruik aansluitingen staat bij fysieke capaciteit 1x25, dit klopt natuurlijk niet. XXX en Netbeheerders hebben de juiste contractwaardes, echter word hij niet aangepast.

- Aansluiting heeft status âĂŸin aanlegâĂŹ terwijl er gewoon al stroom over loopt en die in gebruik is.
- Elke Netbeheerder heeft een eigen procedure om een A1 aansluitingen aan te melden.
- Verkeerde aanmeldingen in het aansluitregister omdat ze soms als kleinverbruik aanmelden. Dan is de aansluiting en meter al gerealiseerd. Daarna komt het pas bij de NS en dan moeten ze een nieuwe meter installeren en de EAN weer omzetten. Met alle netbeheerders is hierover gepraat om dit uniform te maken: . Mijnaansluiting.nl werkt wel bij Stedin/Liander.
- Bij Enexis moet je 3 maanden doorbetalen bij een afsluiting, bij Liander niet.
- Storingen. Ze hebben een pilot met Liander. Alles sturen ze naar 1 adres wat gemonitord word door 2 personen. Die personen kijken dan waar het naar toe moet (NS/ProRail etc). Liander laat het op het zo vroegst mogelijke moment weten. Door vertakkingen is het lastig voor de netbeheerders om de NS te informeren.
- Gebeurd wel eens dat een aansluiting word afgesloten omdat ze niemand hadden kunnen vinden. Of onderhoud en dat toen de OV paaltjes het niet deden.

Saldering

NS Stations gaat binnenkort weer aanbesteden en neemt zonnepanelen mee in het contract. De leverancier moet minimaal dezelfde prijs teruggeven (saldering). Door de ontwikkelingen gaan ze een relatief kort contract aan. Dit om een gevoel te ontwikkelen welke rol ze kunnen spelen.

MJA3 afspraken

Ze hebben deze afspraken en MJA3 zegt twee procent maar NS is ambitieuzer. Ze trekken veel op met ProRail die nog meer wil bereiken. Alle zonnepanelen op de perrons zijn bijvoorbeeld van ProRail. NS heeft jaren lang in warmte-koude systemen geïnvesteerd en zon staat heel erg hoog op het lijstje. Ze hebben immers erg veel ruimte op daken door het hele land heen. Business cases sneuvelen ook altijd bij NS maar ze vinden het halen van de doelstellingen belangrijker. Dan rekenen ze maar met een hoger tarief. NS heeft ook politieke druk om groen te zijn.

Toekomst

Circulariteit is bijvoorbeeld ook voor NS een belangrijk onderwerp. Dingen als 2e EAN op een aansluiting kan een grote rol spelen. De verschillende kostenposten is wellicht een mooie eerste stap. Met de verschillende grote partijen dit afspreken. Daarnaast een ander belangrijk punt is dat er nieuwe aansluitingen en verwijderen kunnen worden gestandaardiseerd.

A-1-2 Ziggo

Ziggo is a competitor of KPN but when it comes down to Energy they face the same problems. They both need to keep their systems up and running in order to keep the customer satisfied.

Ziggo informatie

Facturatie word gedaan door XXX. Zij doen alle controles op eindstanden en ontvangen alle facturen namens Vodafone Ziggo. Roland heeft alleen met Ziggo te maken. Roland houd zich vooral bezig met aansluitingenbeheer, hij controleert of er betaalt word en uit de systemen gehaald. Daarnaast is er ook een groot deel wat er moet worden opgehelderd. Ziggo heeft door zijn geschiedenis veel overgenomen. Zo zijn er aansluitingen waar voor betaald worden maar er zijn geen aansluitingen.

Nieuwe aansluitingen

Alle nieuwe aansluitingen worden door externe partijen (partners / afdeling bouw) aangemeld. De aansluitingen komen dan alsnog bij Roland terecht. Er is een escape, als een aanvraag bij het meetbedrijf binnen komt dan krijgt Roland een blind copy en dan weet hij het alsnog. Er zijn meerdere mensen bezig met de inhuizen/uithuizen. Het is echter lastig iedereen te bereiken qua process maar daar word nu aan gewerkt.

Rolands taak om aansluitingen aan te melden bij de leverancier. Roland houd zich ook bezig met het registreren van alle aansluitingen in een externe database. Ziggo heeft ook wijkkasten en daar zijn meerdere partijen bij betrokken. Dan is er de netbeheerder die een EAN toekent maar er is een mismatch tussen het EAN-adres en het bezoekadres. Registratie bij netbeheerder word door Ziggo gedaan, en dan meld Roland het aan bij de leverancier. Elke nieuwe registratie/aansluiting word voorzien van een nieuwe meter door het meetbedrijf. Soms word er wel eens een kleinverbruik aansluiting aangevraagd. De kleinverbruik zijn de kantoren omdat die buiten het complex (Technisch Netwerk) vallen. Het word snel gezien door het meetbedrijf en die zeggen: dat kan niet.

Facturen

Er is een externe partij en daar komen geen bijzondere vragen uit. Roland is niet op de hoogte van welke vragen naar de leverancier gaan. Soms wel eens een vraag over een nieuwe EAN-code om te kijken of het klopt. Alle aansluitingen die erbij komen of verdwijnen word ook doorgegeven naar de externe partij. Sporadisch zijn er vragen van de netbeheerder: we hebben een EAN-aansluiting zonder leverancier of meetverantwoordelijke. Soms was er onduidelijkheid of het wel een aansluiting was. Er was soms geen match omdat er vanwege de locaties een mismatch was. Het contact loopt via email vooral.

A1

Enexis heeft eigen excel-formulieren voor A1. 2 maanden geleden over gesproken en graag willen ze dat op termijn integreren. Maar dat is nog een lange-termijn-verhaal. De grote 3 zijn belangrijk, de andere zijn kleiner. Als ik een account heb (van Ziggo) op www.mijnaansluitingen.nl dan kan ik daar verwijderingen doorgeven. Als de partners dat doen dan hebben zij niet de keuze voor A123. Ideale wereld is dat als Ziggo iets aanmeld dat het dan sowieso een A1 aansluiting. Enexis is bijvoorbeeld erg flexibel als er administratief iets niet klop en Ziggo met bewijzen komt. Bij Stedin kan dat bijvoorbeeld toch iets langer duren. Maar iedereen werkt op zijn eigen manier.

Storingen

Met de netbeheerder veel afgesproken qua storingen. Gepland onderhoud gaat dan ook naar het NOC National Operating Center. Alle kasten kunnen op afstand worden gepingd eom te kijken of het werkt. Zo kunnen extra telefoontjes vanuit de klant worden voorkomen. Meetverantwoordelijke en Externe partijen Zij managen meerdere meters, ook oud Stedin Meters. Alle afstemming kost extra tijd doordat er een extra iemand aan tafel zit. Daarnaast zijn er veel externe partijen die verantwoordelijkheid dragen in de keten. Er is geen toegang tot het C-Ar om de eigen aansluitingen te zien.

Toekomst

Ziggo is wel bezig met vergroenen. Terugleveren aan het net heeft op dit moment nog niet prioriteit. In het verleden heeft Ziggo (voor fusie met UPC) meegedaan aan de noodstroompoel voor het balans van het net. De noodstroomagregaten konden worden ingezet maar het eigenbelang ging natuurlijk boven. Uiteindelijk is het gestopt vanwege de risicoâĂŹs qua storingen. De servicegraad kan daardoor in gevaar komen. Net zoals KPN, de apparaten worden efficienter maar er worden simpelweg veel meer diensten afgenomen. Per saldo kan het verbruik dus omhoog gaan.

A-1-3 Rijkswaterstaat

Rijkswaterstaat informatie

Coördinatie Bureau Energie (CBE) waar Fred werkt is verantwoordelijk voor de inkoop van Energie. Het CBE koop in voor Ministerie van Infrastructuur en Milieu (I&M), daaronder vallen: Rijswaterstaat (RWS), KNMI, Dienst Wegverkeer (RDW) en voor de kantoren van Ministerie I&M in Den Haag. CBE werkt nu meer samen met andere overheidspartijen: Defensie (Kazernes), Justitie (Gevangenissen) en de Rijksvastgoedbedrijf. Rijkswaterstaat is de meest actieve. Ze kopen samen gecoördineerd in. Tegelijk aanbesteed in 4 percelen (Defensie, Justitie, Rijksvastgoed, I&M). Uitkomst kan dezelfde leverancier zijn maar 4 verschillende. Voorbeeld: RWS gaat van XXX naar XXX.

Slimme meters, Verbruik en Aansluitingen

RWS heeft XXX aansluitingen (grote en kleine aansluitingen). XXX van de XXX zijn kleinverbruik (leveranciersmodel). 500 zijn grootverbruik en daar krijgen ze aparte facturen voor. Lampen langs de rivieren kunnen ook een aansluiting hebben (XXX-XXX kWh) en grote gemalen bijvoorbeeld (XXX GWh). Verlichting van snelwegen + matrixborden (XXX procent van totaal). Verlichting brand bijvoorbeeld alleen in de avond. Gemalen werken alleen bij hoog water, in Eijsden is dit bijvoorbeeld XXXMW die een paar dagen aanstaat. Daarnaast is er een groot gedeelte kantoren met een standaard gebouwprofiel. Er zijn ook punten op het water die alleen een accu en zonnepaneel hebben, hier is geen aansluiting/EAN voor nodig. Al enkele jaren zijn de meerderheid van de aansluitingen voorzien van Slimme meters. Het loopt redelijk soepel maar het is natuurlijk per netbeheerder anders. XXX heeft een product XXX, deze is aangekocht door RWS om te gebruiken voor het aflezen van de standen. Dit maakt profielen, rapportages en ook voor verschillende afdelingen (Zuid Holland bijvoorbeeld). Het meetbedrijf is anders aanbesteed.

Mutaties

Belangrijk voor Rijkswaterstaat is bijvoorbeeld mutaties. Mutaties is het een uitdaging aangezien oude aansluitingen worden weggehaald en opnieuw geplaatst. Of de aansluitingen worden verzwaard. Daarnaast ook de aansluitgrootte is belangrijk omdat ze daar ook geen compleet zicht over hebben. Het verschil tussen KPN en Rijkswaterstaat is dat Artikel 123 bijna geen rol bij RWS speelt. Er kleven voordelen (bescherming, 1 factuur met alles bij elkaar) maar ook nadelen van kleinverbruik (geen meetbedrijf). Die XXX hebben standaard kosten bij de netbeheerder, welke per netbeheerder weer afwijken.

Facturen

In XXX kan je de factuur als PDF opvragen en daarnaast is er nog UML. Via XXX zit er geen controle meer vooraf, alleen achteraf. Er is genoeg vertrouwen. Daarnaast is er een check om te kijken of de aansluiting bij RWS hoort. Daarnaast word er een check uitgevoerd of het een afwijkend bedrag is. De gemeten waardes komen niet altijd correct aan het einde van de keten aan. 4 facturen en aan het einde veel gedoe. Dingen kloppen niet omdat het verbruik dan toch niet klopt (van de XXX kloppen er ongeveer XXX niet) Dit zijn grote afwijkingen. CBE faciliteert hierin. Dan gebruiken ze XXX om te kijken of het klopt en wat de data is. Er zitten elk jaar fouten in. Redenen hiervoor kan een meterwissel bijvoorbeeld zijn. Factuurdata kan ook steeds vaker als stroompje worden verstuurd via XXX bijvoorbeeld. Grote systemen doen dan zaken met elkaar echter is hier wel genoeg vertrouwen nodig in de keten.

A123

RWS heeft contact accountamanagers bij de netbeheerders maar niet veel aangezien alle facturen worden doorgegeven (kleinverbruik - leveranciersmodel). Hierdoor kan je het leveranciersmodel gebruiken en alles van de netbeheerder staat op de leveranciersfactuur. Rijkswaterstaat maakt geen gebruik van A123 (Zie wet). De leveranciers betalen de netbeheerders van de voorschotten. Wel gebruikt RWS de complexheffing. Zo zijn er meerder complexen (bijvoorbeeld snelwegen of vaarwegen). Je mag binnen een complex de verschillende verbruiken bij elkaar optellen voor A123. De leverancier heft deze namens de Belastingdienst. Je betaalt dan ook BTW over de energiebelasting (Wet belastingen op Milieugrondslag). Netwerkkosten, leveringskosten, Energiebelasting en BTW zijn de 4 componenten voor de kosten van energie. Voor de grootverbruikers sturen ze alles per maand naar XXX, zij zijn een serviceprovider (facturatie). Postbusnummer van Rijkswaterstaat en daarna word het ingescand.

Sommige Netbeheerders sturen met Universal Business Language (UBL) alle data door de keten. Rijkswaterstaat heeft een speciaal XXX systeem dat automatisch betaald. Steekproefsgewijs kijken ze aan het einde van het jaar. Totale kosten: XXX, en Rijkswaterstaat betaald voorschotten aan de leverancier. In het vierde kwartaal volgt de uiteindelijke afrekening. Het is eigenlijk handiger om een partij als XXX eruit te halen om kosten te besparen. Het meetbedrijf is echter een ander bedrijf: XXX. Groter dan 100kW dan moet je elke dag de verzamelde data doorgeven aan de netbeheerder. Dit is volgende de Meetcode. Daarnaast is er ook de Informatiecode.

Toekomstige mogelijkheid 2 EAN's op 1 aansluiting

Binnenkort komt er de mogelijkheid om 2 EAN′s op 1 aansluiting te hebben. Zo kan je stroom afnemen bij de NUON en terugleveren aan Essent bijvoorbeeld. Met nieuwe hernieuwbare energie kan dit belangrijk zijn, het is natuurlijk de vraag hoelang de salderingsregeling blijft bestaan. Voorbeeld: een zoutsilo waar zonnepanelen op worden geplaatst door de lokale buurt die korting krijgen op de rekening. Door de 2 aansluitingen gaat dat veel nieuwe mogelijkheden opleveren. Je mag echter niet zonder vergunning aan huishoudens gaan leveren. KPN Energie kan bijvoorbeeld wel nu heel makkelijk aan RWS gaan leveren. Rijkswaterstaat haar verantwoordelijkheden zijn de vaarwegen, wegennet en het droog houden van Nederland. Maar toch wil Rijkswaterstaat verschillende opwekideeen uitzoeken. Wellicht is de ruimte ter beschikking stellen efficienter aan derden die vaker windmolens exploiteren. Die energie die opgewekt word kan je weer kopen. RWS heeft een aansluitingenregistratie gekocht met EAN-code, GPS, locatie en fotos omdat ook RWS hetzelfde probleem heeft met rare locaties. Er is ook een link naar XXX.

A-2 Distribution Grid operators

A-2-1 Enexis

Waarom treden er verschillen op in de administraties?

Iedereen houd zijn eigen administratie er op na, daardoor kunnen er fouten optreden.

Facturen

Aan het begin van de maand dan stuurt Enexis een factuur voor de vorige maand. Voor de netwerkkosten sturen ze een factuur gebasseerd op het CAPTAR-tarief. Op basis van het aansluittarief betaal je je vastrecht. Deze gegevens staan geregistreerd in het C-Ar. Deze data kloppen tot op zekere hoogte. Zo kan het voorkomen dat er een capaciteitsverlaging op de aansluiting is geweest. Als Enexis hier niet bij betrokken is dan weten ze ook niet dat er op de fysieke locatie iets gewijzigd is. Dit kan het geval zijn van een aannemer die buiten Enexis om werkt. Ze melden het wel naar de opdrachtgever maar niet aan het C-Ar. Dit zou niet moeten kunnen terwijl dat toch sporadisch voorkomt. Vooral bij sloop komt dit voor, minder bij aansluitwaardeaanpassingen. Het is ook voor Tennet belangrijk dat het Standaard JaarVerbruik klopt (even uitzoeken).

Wat doen jullie nou precies met het C-AR?

Het aansluitregister waar Wytske en de collega's mee werken synchroniseert telkens met het Centraal Aansluitingenregister. EDSN beheert dat systeem en de andere partijen in de keten hebben hier toegang tot, de consumenten niet. Een privacy-aspect is de reden waarom je als klant geen toegang krijgt tot de aansluitingen van concurrenten bijvoorbeeld. Als je een overzicht van Enexis krijgt dan is het eigenlijk ook een overzicht van het C-Ar. Dit is niet veel werk en kan makkelijk gerealiseerd worden.

A1

Anexo stuurt de meetdata eens per jaar de data. De afdeling facturatie controleert de meetdata op grote verschillen. Indien er iets fout is dan ze terug naar de MeetVerantwoordelijke. Uiteindelijk gaan ze akkoord en dan sturen ze de data naar het meetregister. Staat als grootverbruik in het C-Ar (keuze KV/GV). A1/lid2/3 (ja/nee). Als A1 aansluiting kan je ook maandelijks afrekenen op verbruik. Dan krijg je wel een maandelijks overzicht van je verbruik. Eneco ontvangt ook de data van Anexo. Beide partijen (NB en LV) checken de meetdata. Alle A1 zijn verplicht om die te laten vervangen door slimme meters. KPN is niet de enige en alle andere grote bedrijven hebben vaak ook niet de meterstand opgenomen. Zo kan het zijn dat de meters op wissels of overgangen zijn geplaatst (in het geval van NS) en daardoor is het erg lastig om te bekijken. Er zijn dus erg veel meters oude meters die nog niet zijn uitgelezen. Dit zorgt voor onzekerheid/verrassingen waar Netbeheerder als Afnemer niet op zit te wachten.

NS-punt over de aansluitwaardes

Waarom zijn deze waardes eigenlijk anders. Er moet iets gebeurd zijn dat de waardes anders zijn. Eerst kan er worden gekeken bij de infrasturctuur-afdeling wat er op de tekeningen staat. Als de Klant dan alsnog een andere mening heeft dan is een schouwing nodig. Dit kan uiteindelijk goed uitpakken omdat er dan duidelijkheid is voor beide partijen. Meestal is het 50/50 qua wie er gelijk heeft. Meestal zijn de aansluitingen ook wel weg, alleen het is nooit gemeld aan Enexis. Ook monteurs kunnen fouten maken. Monteurs kunnen fouten maken met A1. Het zijn natuurlijk Kleinverbruik aansluitingen maar bij activatie moet je dan niet naar Kleinverbruik bellen want daar staat hij niet in het KV-register. Monteurs weten bijvoorbeeld ook niet welke klant er A1 is. Zo is er een fysieke werkelijkheid en een digitale werkelijkheid. Juist daarom is de overeenstemming en samenwerken met de klant volgens Wystke zeer belangrijk. Door gegevens naast elkaar te leggen en te verifieren kan je als partijen beiden verder komen. In de toekomst komen er werkelijke standen aan en dan zou het proces beter moeten verlopen.

Externe partijen

Deze partijen zijn niet overbodig geworden door de digitalisering. Veel zakelijke klanten besteden de energieadministratie uit. Door de complexiteit van de energiemarkt is het begrijpelijk om een externe partij in de arm te nemen. Echter hebben die tussenpartijen minder een belang erbij om de kosten naar beneden te brengen. Deze partijen hebben niet een heel andere opstelling tov een andere losse klant. De tussenpersonen hebben minder inzage in de aansluitingen.

Wat is de procedure om een A1 aan te melden?

Enexis werkt met een aanvraagformulier voor het aanmelden van een A1 aansluiting. Via een special mailadres komt het terecht bij Wytske en dan word er een EAN code aangemaakt. KPN is bijvoorbeeld de enige die in bulk mag indienen en de rest kan dat weer niet.

Contact

Enexis heeft veel meer te maken met meetbedrijven. Ze hebben veel minder contact met andere netbeheerders. Wytske onderschrijft dat iedereen zijn eigen manier van werken heeft. Zo heeft bijvoorbeeld Liander alles in 1 register. Binnen Enexis zijn er bijvoorbeeld 2, KV en GV staan los van elkaar. Die systemen worden dan ook apart behandeld. Het verschil is verplicht. Als een aansluiting van KPN bij KV is aangevraagd, dan hoort die daar niet geregistreerd te zijn. Je moet dan weer een nieuwe EAN krijgen. Bij Liander zou dit geen probleem hoeven zijn. Voor de klant kan dat soms niet prettig zijn aangezien er bijvoorbeeld een nieuwe EAN voorbij komt. Bij Enexis moet je 3 maanden doorbetalen bij een afsluiting, bij Liander niet. Klopt dit? Voor grootverbruik is er een overeenkomt transport. De opzegtermijn is 3 maanden. Er zijn 3 opties daarna: -Overnemen -Afsluiten -Na 3 maanden word de aansluiting verwijderd

Waarom mag je van A1 niet terug naar KV en dan weer naar A1?

Het mag wel, echter willen de partijen voorkomen dat bedrijven kaan ping-pongen. Het kost immers geld omdat ze wisselen van MV. Dus nieuwe meter en een nieuwe EAN. Binnen Kleinverbruik mag er alleen maar een meter van Anexo komen. Liever niet elke maand wisselen of jaar, het kan wel.

A-2-2 Stedin

Stedin informatie

Mark is accountmanager/relatiebeheerder bij Stedin. Basti is Netbeheerder zijn natuurlijk niet gericht op winst. Hij is het aanspreekpunt voor grote zakelijke klanten zoals KPN. Stedin heeft verschillende afdelingen: aansluitingen en vervangen, meterkast en aansluiting (KV/huishoudens), klant en markt (marktprocesen zit hier in), marketing en communicatie, klachtafhandeling, klantenservice, strategie en innovatie, BVC (BedrijfsVoeringsCentrum) en er zijn nog veel aannemers die buiten werkzaamheden verrichten. Recent is Delta overgenomen en de komende jaren gaat er gekeken worden hoe dit in te passen.

Waarom treden er verschillen op in de administraties?

Verschillende systemen, manieren van werken zorgen voor verschillende manieren van werken onder netbeheerders. Van oudsher is Stedin 13 kleine netbeheerders. Alles is nu gecentraliseerd maar je hebt te maken met dingen uit het verleden. Zo heeft stedin ook niet altijd alles 100% op orde. Zo vroeg Stedin bijvoorbeeld meer informatie over bepaalde aansluitingen die al sinds 2013 waren verlaten door KPN. Stedin had ze nog niet uit de systemen verwijderd en vandaag de vraag aan KPN wat ze wisten over de aansluitingen.

SJV

Bij KPN locaties worden niet altijd standen opgenomen. Je kan niet aanbellen bij de aansluitingen waardoor je geen inzicht kan hebben. Er moeten dus aannames gedaan worden door het systeem. Als er een lange tijd niet word geschat dan mag de leverancier een schatting maken. Naar aanleiding van de aansluitcategorie en de rest van de populatie kijk je. Dan word je ingedeeld op een bepaalde waarde. Als je echte standen krijgt dan word de SJV weer opnieuw berekend. Wie doet er nog iets met deze informatie? Met slimme

Facturen

Maandelijks stuurt Stedin de facturen. Ze weten wat de aansluitingen zijn en die nemen ze mee op de factuur. Stedin rekent af op stuksbasis en dit verloopt automatisch. Dit is anders dan bijvoorbeeld Liander die 17 facturen heeft.

Meetbedrijf

Het meetbedrijf stuurt eens per jaar alle data door de hele keten heen. Stedin is compleet afhankelijk van de data die ze van het meetbedrijf ontvangen.

Wat doen jullie nou precies met het C-AR?

Elke nacht synchroniseert het systeem van Stedin met het C-AR. Als netbeheerder ben je ook afhankelijk van de leverancier aangezien zij geld innen namens de netbeheerder (KV-aansluiting). Verschil tussen digitale en fysieke werkelijkheid? Er is een team binnen Stedin die dit regelt en oppakt.

A1

Staat als vinkje in het register. Het valt onder grootbruik in de systemen van Stedin. Maar als je bijvoorbeeld 3x25A uitdraait dan komt deze waarde wel naar boven. Bij KV komen ze er niet uit, bij GV wel. Het aanmelden van een A1 heeft Stedin intern moeten inregelen. Via mijnaansluitingen.nl kan je als aannemer ook een aansluiting aanvragen namens KPN. A1 levert ook meer problemen op, je kan het fysiek niet zien. Zo zie je het er niet aan af en niet de hele organisatie weet het. Niemand checkt bijvoorbeeld ergens als je een A1 aanmeld, dus er kan er ÃľÃľn doorglippen.

Stuksbasis

4000 nieuwe aansluitingen voor KPN de komende jaren. Die aansluitingen gaan allemaal met de hand worden ingevoerd. Segmenten kan je automatiseren maar de fysieke kabel moet je op locatie gaan kijken.

Fysieke en digitale werkelijkheid

Er is een afdeling locatiebeheer die een EAN maakt en als zij twijfelen dan laten ze het bekijken. Dan gaat er een monteur fysiek naar de locatie op te kijken. Soms doe je aannames maar dat kan niet altijd. Dat komt soms ook door externe partijen die zelf opereren. Het nadeel hiervan is is dat je gewoon door blijft betalen.

Locaties

Zo zijn er wel eens andere adressen of locaties die anders in de systemen staan. Betalingen voor locaties die niet bestaan of niet betalingen voor locaties die wel bestaan. Communiceren met EAN codes is erg belangrijk. Je bent afhankelijk van de berichten van anderen.

Uithuizen

Er is geen controle en zo kan je je buurman uit laten huizen van een energieaansluiting. Je kan heel makkelijk bij een andere leverancier aanmelden.

Externe partijen

Zij hebben minder inzicht. Het is wel hun specialiteit maar hebben minder toegang tot gegevens. Na een machtiging krijgen ze toegang tot gegevens. Externe adviespartijen moeten daar iets meer moeite voor doen. Er komen wel meer vragen vanuit die partijen. Nu kan je datacontroles uitvoeren maar dat kan over 10 jaar weer oud zijn. Daarom is het noodzakelijk vaak bij elkaar te gaan zitten en dat uit te zoeken. Ze staan positief tegen een reguliere opschoning. Je kan lijsten blijven controleren maar je hebt dan een digitale werkelijkheid, geen fysieke.

Wat zijn de checks en balances in het systeem?

Meerdere afdelingen zijn bezig met data. Door verschillende verschuivingen in de organisatie is het altijd lastig om je data up to date te houden. Er zijn checks of de facturatieafdeling. Het is ook lastig want bij een gezinsuitbreiding dan kunnen er ook grote goicijfers worden gezien qua toename. A1 is niet ideaal, het is kort beschreven en niet heel expliciet uitgewerkt. In de praktijk komt A1 ook niet zo heel erg vaak voor. Bepaalde elementen zijn nog steeds niet goed geregeld in de energiewet. Centralisatie ARM Eneco hoeft bijvoorbeeld alleen nog maar met 1 partij contact te hebben voor alles mailtjes over de klanten die ze vertegenwoordigen.

Voorbeeld van samenwerking

Stedin en KPN zorgen ervoor dat beiden partijen het maandelijks eens zijn over de te factureren EANâĂŹs. Het zou in iedereens interesse moeten zijn om hier aan te werken. A123 kan je niet zomaar spanningsloos zetten. Het is een grootverbruik met directe meting. Dus je moet de meter weghalen en dan de aansluitkabel.

A-2-3 Liander

Liander informatie

Liander is afgeslitst van Nuon, sinds 01-08-2014 is dat apart gegaan [http://wetten.overheid.nl/BWBR0020608 08-01]. Ze hebben 1 SAP-systeem voor kleinverbruik en grootverbruik. Er zijn verschillende profielen voor medewerkers met verschilllende toegangsauthorisaties.

Create, Read, Update, Delete

Kunnen een nieuwe EAN aanmaken, lezen, updaten en verwijderen. De contactpersonen hebben complete toegang tot de systemen om alles uit te voeren. Ze kunnen alles uitbouwen en op gesloopt plaatsen. Doorlaatwaardes kunnen ook worden gewijzigd. Er zijn ook soms fysieke wijzigingen waarbij de aansluiting verzwaard kan worden. Eerst is er www.mijnaansluitingen.nl voor een aanvraag en een EAN.

Stappen voor aanvraag aansluiting

Eerst is er de aanlegafdeling. Liander -> Klant en Markt, afdeling die alle service voor alle klanten vericht. Service word vericht vanaf de aansluiting aangelefgd is. Administratieve afhandeling van de aanleg vind daar ook plaats. De 3 datacenters worden apart afgehandeld aangezien dit grote afdelingen zijn. Nauw contact met de klanten voor de incassokosten. Bij openbare aansluitingen word er veel verantwoordelijkheid genomen voor de aansluitingen. Binnen de A1-klanten zijn er gemeentes die aansluitingen hebben voor riool en verlichting. KPN heeft meer aansluitingen en heeft ook meer verschillende soorten aansluitingen. Het idee is dat het mandaat zo laag mogelijk in de organisatie ligt. Zo kan de klant het beste geholpen worden.

C-AR

Het C-AR synchroniseert elke nacht. Sommige velden synchroniseren meteen. Elke ochtend worden er rapportages en checks uitgevoerd om zeker te zijn dat de waardes gelijk zijn. Ook voor de allocatie en reconciliatie is het erg belangrijk dat de waardes kloppen en de systemen gesynchroniseerd. Wanneer de status afwijkt krijg je een andere allocatie. Als netbeheerder zien ze alleen de technische gegevens. Atributen zouden gelijk moeten zijn volgens de informatiecode maar iedereen houd er iets anders op na. Naam contractant staat in het C-AR. De leveranciers is verantwoordelijk voor de klantgegevens in het C-AR. Liander volgt deze regel en gaat deze niet overrulen. Liander: Janssen en Zonen B.V. Eneco: Janssen B.V. Nu word het KvK-nummer vaker gevolgd om zeker te zijn dat het klopt.

Alle data is van de netbeheerders. EDSN is de uitvoerende partij, de eigenaar is de Netbeheerder. Netbeheerders hebben de meeste rechten.

De netbeheerder kan alleen zijn eigen aansluiting inzien. De leveranciers kunnen dingen inzien van de eigen contractanten. Ze kunnen een aanvraag doen uit het TM voor meer informatie voor potentiele klanten. Voor de afsplitsing van het meetbedrijf konden ze ook dit nog aanpassen.

Switchen van aansluiting

Er zit een menselijke check op voor een switch. De kennis ligt bij de medewerkers om een beoordeling te doen. Ze huizen hem niet in, dan vraagt Liander aan KPN of het klopt. Leveranciers kunnen aansluitingen wegswitchen zonder de contractant. De Leverancier heeft niet genoeg informatie om deze keuze te maken want je moet er op afgaan bij de huisnummertoevoeging. Ze zien echter wel dat het een aanvraag op een GV of A1 is terwijl ze een kleinverbruik-contract hebben afgesloten. Liander laat alle A1 aansluitingen bij 1 team samenkomen dat daarin gespecialiseerd is.

Verschil tussen fysieke en digitale werkelijkheid

Het vermoeden is er dat je langs moet gaan. Met de meterswitch is er gekeken door het meetbedrijf wat de waardes zijn. Er waren een aantal die afweken die ze hebben aangepast. Maar ze hebben deze nog niet allemaal sturctureel nagelopen. Het is voldoende om te kijken naar de verschillen. Er zijn ook andere tarievenstructuren maar die hebben geen impact op de prijs structuren.

SJV

Dit is voor een andere partij in de markt. Eneco gebruikt dit bijvoorbeeld voor de nominatie. Uit de analyse van de data blijkt dat niet alles klopt. Liander is verantwoordelijk voor het distribueren van de meetgegevens en SJVâĂŹs. Als er niks word meegegeven bij de aanvraag maken ze een berekening. Bij GV (Telemetrie) is er niet altijd een SJV. Na een nieuwe meterstand word er weer een SJV berekend. Er is een andere afdeling binnen Liander die dat regelt. Wat zijn de speciale regels voor SJV bij Liander?Nee er zijn geen speciale regels zoals die bij Enexis zijn.

Afrekening

Het expertisecentrum Facturatie en Data Monitoring is blij als klanten samen willen werken. De gemeente Amsterdam heeft een overstap gemaakt naar e-facturatie. Nu zijn er XML facturen die naar de klant (gemeente Amsterdam) gaan.

Appendix B

Actor Analysis

B-1 KPN

Background and Contextual information

KPN is a so called Multi-site operator. This means that their connections to the electricity grid are spread out over thousands of locations in the Netherlands. Their street cabinets and bigger hubs are in almost every little village. So as a company with thousands of invoices they must deal with all the Distribution System Operators in the Netherlands. Currently all these parties have their own way of communicating to the final customer. This is not a problem for customers that have just 1 Distribution System Operator however it becomes a hassle for others. KPN is a big user of energy, therefore they are paying much less compared to regular customers. Normal prices are XX, which for KPN it is YY. When Energy is a relatively small cost it was not needed to be on top of it. KPNs vision on energy has changed and awareness of the energy transition is growing. KPN is maintaining a network and investments are made for large periods of time, therefore it is essential to make sure that choices are robust. For years Energy has not been an important topic for KPN. Everything was outsourced to their energysupplier (Eneco). Eneco would collect the invoices from all the people in the chain and present the total costs to KPN. Since 2017 there has been in change in where KPN wants more transparency for what they pay.

Main Goal

KPN wants to anticipate developments on the market and have a good strategy and policy regarding energy. Essential is a good overview of the current energy usage and a standardized way of exchanging information. A transparent insight in the usage data is important to analyse the data and improve the future use of energy. KPN is trying to have a higher quality dataset to help make decisions in the future. KPN is not unique in using energy throughout the Netherlands, competitors like VodafoneZiggo, Tele2 and T-Mobile are in exactly the same situation. Besides telecom there are also other industries that are using energy in multiple places in the Netherlands, Transporting companies like NS. Also the Government is a big user of electricity in the Netherlands. Ministry of Defense, Government-buildings and Rijkswaterstaat are examples that are also spread throughout the Netherlands. Because of this, multi-site operators are dealing with a lot of stakeholders and companies in the chain.

Goals

KPN strives for more transparency in their billing and in their connections. Gaining more insights in the data is a priority. Financial streams in the company thereby will become more predictable. KPN is willing to take the lead in things like this. KPN is a member of the VEMW. KPN like all companies are dependent on laws and regulation. When the polluter has to pay things might change for KPN. This could mean higher prices need to be paid for energy. When this is the case it will need a higher budget and becomes more important to the finance department and the rest of the company. When costs are rising there is a bigger need for transparency and accountability. Therefore KPN is very interested in improving the communication in the chain.

Data owned by KPN

KPN has lots of data about the connection. KPN knows a lot about the object

Which means do they have?

KPN is a big customer. They are allowed to select their own supplier and their own MV. So when they are not satisfied with the service they received they can switch to another MV/supplier. DSOs are fixed and therefore KPN does not have the mean to change this.

What are critical points from their points of view? KPN wants to deliver power to all their connections without interruptions. A power shortage could lead to serious problems. KPN is also in charge of keeping up the national emergency hotline 112 and also has a lot of

B-2 TSO and DSO

Tennet (TSO)

The Netherlands has one operator of the high voltage transmission: Tennet. Tennet has three main responsibilities [Tennet website]:

- Operating the high-voltage grid.
- Maintain the balance between the supply and demand on the grid.
- Facilitate the electricity market.

For maintaining the balance there is the Noodstroompool. Tennet is in contact with a group of heavy energy users. These heavy energy users are able to rely on their backup and scale down quickly. By doing that they have an influence on the dropping of the demand and thereby balancing the grid. A few times a year this happens where the balance needs to be restored again. Tennet is thereby dependent on the information that has been given by the suppliers.

Important: They want to provide a stable energy connection to the entire Netherlands.

DSO's (Distribution System Operators)

Within the Netherlands there are 8 organisations that are tasked with the job to deliver the electricity to the sockets. Throughout the years multiple mergers have been executed and this is the current landscape. Stedin, Liander and Enexis together have access to more than 90% of the market.

Important: They want to deliver electricity to all their connections at the lowest price posible.

Appendix C

Expert workshop

I have spoken to different experts between 18 and 20 September.

Table C-1: My caption

Role interviewee	Company	Category
Energy connection manager	Vodafone/Ziggo	Multi-site user
Energy expert	Rijkswaterstaat	Multi-site user
Energy expert	Anexo	Measuring company
Strategic Partner Manager	Eneco	Supplier
Accountmanager	Stedin	DSO
Accountmanager	Enexis	DSO
Energyteam-member	KPN	Multi-site user
Energyteam-member	KPN	Multi-site user
Energyteam-member	KPN	Multi-site user
Energyteam-member	KPN	Multi-site user

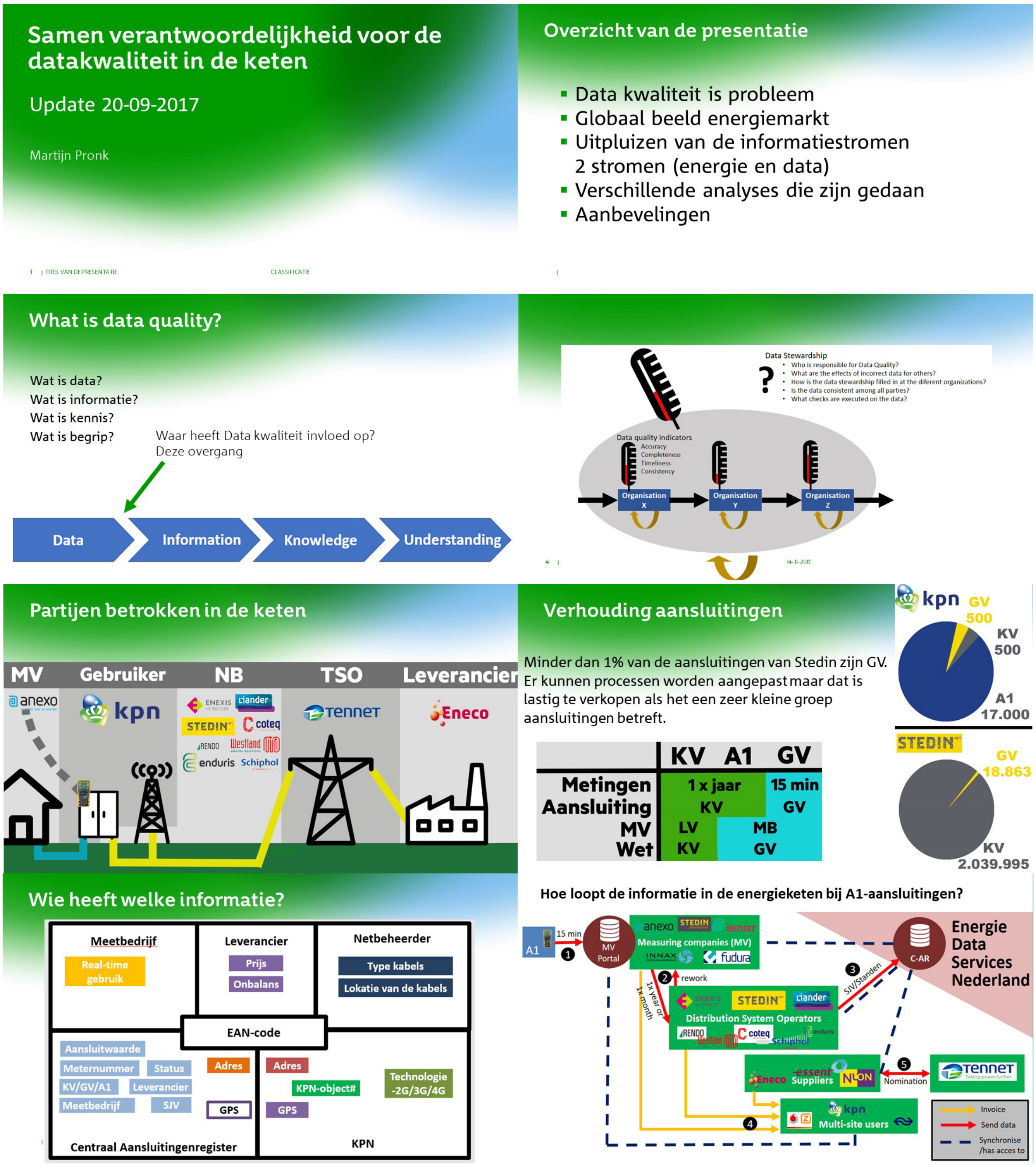


Figure C-1: Slide 1-8 feedback session

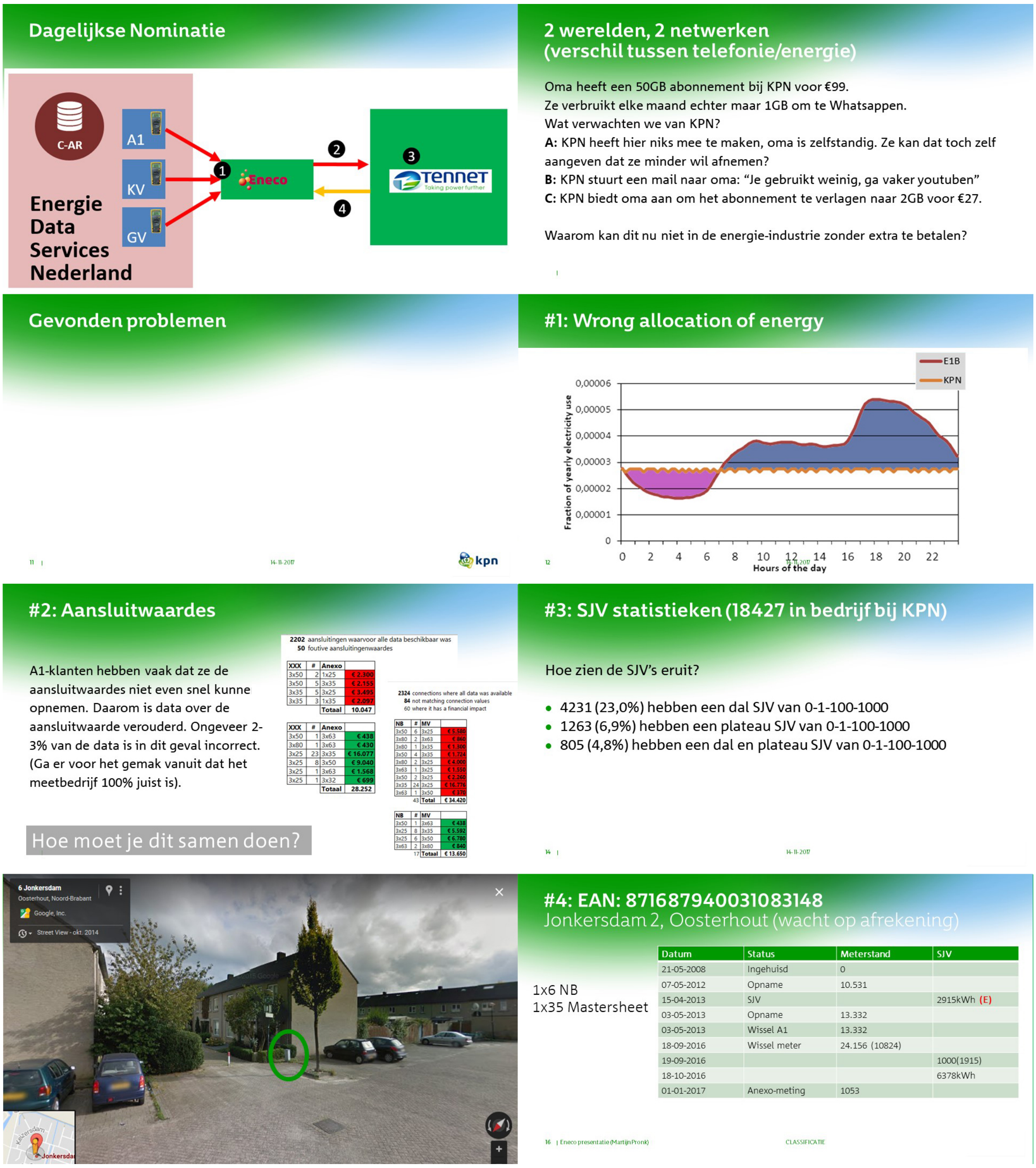


Figure C-2: Slide 9-16 feedback session

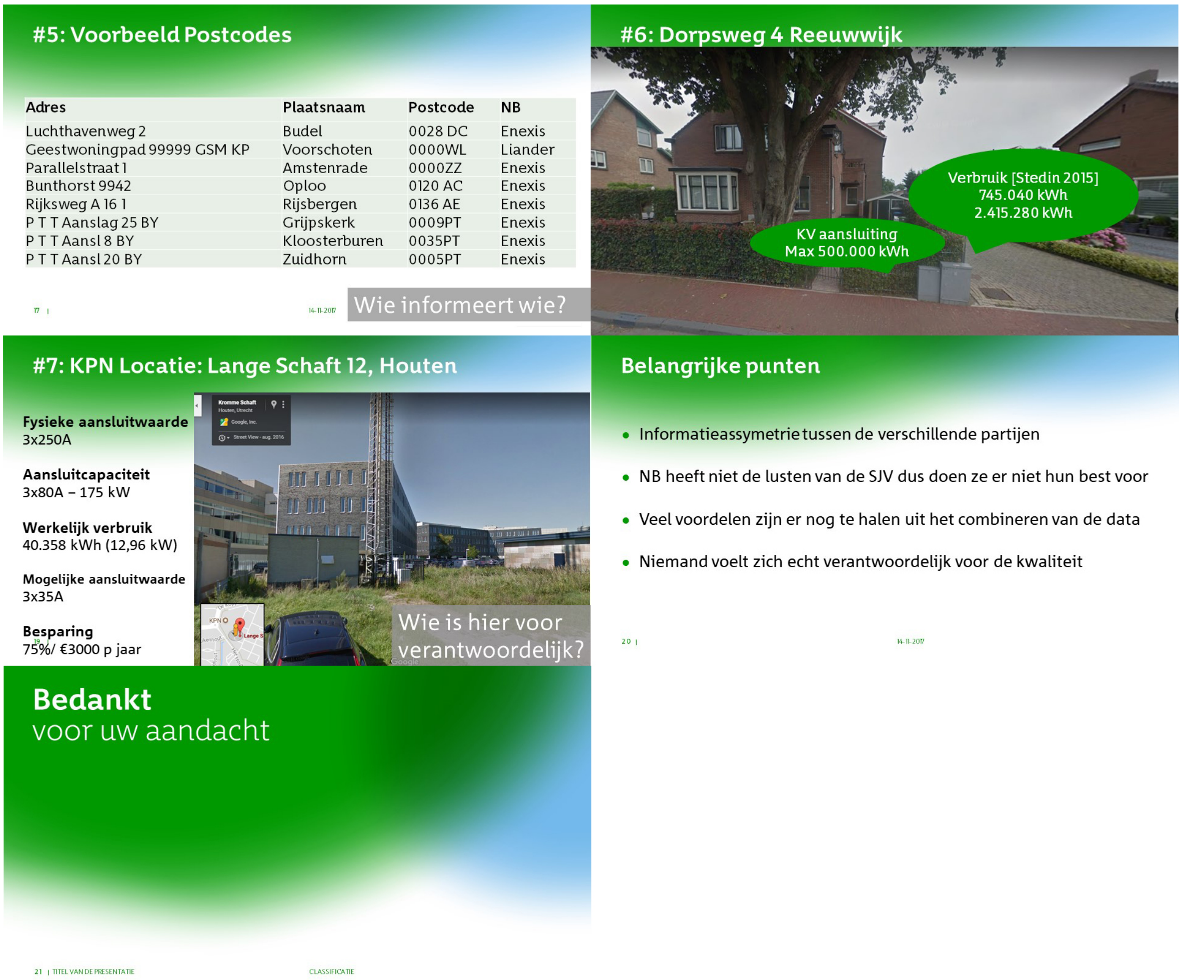


Figure C-3: Slide 17-21 feedback session

Appendix D

Findings in the electricity chain that do not influence the data quality

D-1 Findings in the electricity chain that do not influence the data quality

H: Flexibility in the chain

Something that could happen is that names change over time. KPN back in the old days has had different names for their connections. Unfortunately it is not so easy to change names without making new contracts. Flexibility in the chain is therefore almost impossible to achieve. The different parties all have their own processes which makes it almost impossible to synchronize and be flexible in the chain.

After an analysis of the different connections in the chain it can be seen that DSOs have their own way of working. They work with codes that make sense within the own company but for the other parties in the chain are hard to be understood. DSO Coteq uses special codes like T1500 to talk about a GV connection. Enduris accidentally added a code NSL100 to 80 connections which means that they should be deleted, however this is just a code used that is used by Enduris only. When Stedin provided the data it was hard to interpret since they had different codes for the connection value.

DSOs have been faced with a lot of mergers over the past 20 years. Data had to be brought together where companies used to have entirely different systems. Migrations therefore can be tough to manage when it comes down to data management. [17]. The data will enter the new organisation and the previous organisations may have performed changes that are not always seen. Therefore a registry of the actions on the connection are not always complete. By standardizing the rules it will lead to a uniform entry of the data. Currently the values are only understood by the DSO for instance. Other DSOs might not understand what is written down since a lot of DSOs have their own codes.

I: Inconsistent ways of working between DSOs

DSOs have different rules for when connections are being switched from KV to A1. In the case of Liander you keep your EAN since you will stay within the same system. Stedin will create a new code and make the switch when the first day of the month arrives, the KV leaves the system and the A1 starts functioning. At Enexis you have to take care of it youself and you have to dismantle the A1 connection yourself. This shows already that even though the job of DSO is the same they have a different way of treating it. Eneco also told us that they have special SJVs that override the C-AR ones. Eneco has tried to contact the DSOs but they will stick. So Eneco has their own SJVs.

J: Incorrect administration at DSO

Medio 2017 Stedin has asked KPN about connections that KPN has left in 2013. Back in those days Stedin never deleted the connections from their system properly. KPN was asked to provide more information while it had left the place already for 4 years.