Participatory Value Evaluation as a Tool for Value Extraction and Opinion Mining
Reduce Manual Data Analysis by Automated Value Extraction

Master thesis by Lionèl Kaptein

Frontpage image retrieved from ("Steel Eye," 2020)

This page is intentionally left blank

Participatory Value Evaluation as a Tool for Value Extraction and Opinion Mining

Master thesis submitted to Delft University of Technology

in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE

in **Engineering and Policy Analysis**

Faculty of Technology, Policy and Management

By

Lionèl Nathanaël Kaptein

Student number: 4251334

To be defended in public on 30 September 2020

Graduation committee

Chairperson   : Dr. Ir. B. Enserink, Multi-Actor Systems, Delft University of Technology
First Supervisor   : Dr. Mr. N. Mouter, Engineering Systems and Services, Delft University of Technology
Second Supervisor   : Dr. S. Spruit, Multi-Actor Systems, Delft University of Technology

Preface

This document is the final deliverable I will submit for the master Engineering and Policy Management at Delft University of Technology. I would like to use this opportunity to thank those people that helped me in completing this research.

First of all, I want to thank my graduation committee for helping me throughout this research. I still remember the first moment that I walked into the office of Niek Mouter and Shannon Spruit. They inspired me by their scientific and entrepreneurial mindset to let me reflect upon opportunities. Also, Bert Enserink has helped me in the process of deciding which graduation project would be most relevant. After some setbacks in trying to arrange another project at a corporate company, I am thankful for the patience and warm welcome in the team of Participatory Value Evaluation. I think I have visited all supervisors only twice due to the situation of Covid-19. However, sharing a coffee online during our meetings has been a pleasure to me and I am happy with the critics, compliments and advice to keep improving my work. I will not forget the moment at which I have visited the first physical citizen meeting after the 'lockdown' in Súdwest Fryslân. I remember Shannon mentioning that it felt like a holiday trip after being home for a long time due to the 'lockdown'. Another memorable moment has been the week of conducting the Covid-19 PVE at which everyone was working and available throughout the weekend to discuss the progress.

I also want to thank Pradeep Murukannaiah, Enrico Liscio, Michiel Meer and Luciano Cavalcante Siebert from the department of Computer Science at the TU Delft to help and reflect in the field of Natural Language Processing. It was fun to process the motivations of participants together. Also, Anatol Itten has given valuable feedback during the monthly thesis circle.

Lastly, I have enjoyed many opportunities during my student life in Delft, The Hague and Kenya. I would like to thank my parents, family, friends, colleagues at Kojac CV, housemates and Leonie for their support and distractions over the last eight years.

Delft, September 2020

Summary

Many grand challenges like climate change or health security cannot be solved by only the policymakers. Support and expertise of citizens is needed to solve these challenges (Gerton & Mitchell, 2019). Due to the demand for public participation, Participatory Value Evaluation (PVE) fulfils the needs of involving many participants in a scalable way in the process of participation and evaluation. PVE is unique in the way that it allows people to be a virtual decision-maker in a specific context. This method is not a decision-making tool in itself, but it allows participants to make their quantitative decisions and asks them to motivate their answers qualitatively. This results in qualitative and quantitative data about their choices. In this research, two case studies have been conducted, one regarding the Sustainable Energy Strategy in Súdwest Fryslân, another one for reducing Covid-19 measures of an intelligent lockdown in The Netherlands. In both cases, qualitative data is asked from the participants by asking for motivation of their choices. However, the previously used methods cannot analyse all the qualitative data which decreases legitimacy, support and quality of decision making as public expertise was often unused due to the lack of human capacity for data processing. Human annotation of motivation is limited in scale, one hundred motivations take four hours to annotate, the Covid-19 case study did involve >30.000 participants. This research focused on exploring opportunities for automating the process of qualitative data analysis to support public participation. The following research question has been proposed:

How to support public participation by using automated qualitative data processing in Participatory Value Evaluation to extract values that can be used in decision-making?

First, public participation has been explored to evaluate the added value of large amounts of qualitative data input for decision-makers. Decision-makers have stated that they are striving to use the expertise of their public in decision-making. From this data it is therefore valuable to decision-makers to extract arguments, to determine prerequisites of policies and to know about (unique) design principles that are proposed. The focus in this research has been set on human values to categorize the motivations of participants, according to stakeholders the values should be complemented by policy arguments and prerequisites as well. The main reason for policymakers to use values is that it allows discussion that is not aiming for solutions or negotiation and that there is no good or wrong in human values.

Automated data processing methods will make it possible to categorize motivations based on values. This can be done by software packages for annotation to make the human processes more efficient and to be able to quantify and filter the results of qualitative data analysis. However, Natural Language Processing (NLP) will rapidly increase the speed of this analysis as the language is automatically being processed. This can be done (semi) supervised and unsupervised NLP for labelling based on values via topic modelling. For topic modelling in the context of PVE semi supervised NLP will have to be used to give directions. Afterwards sentiment analysis can be done to find the sentiment of specific topics or to find positive and negative themes in general. The current topic modelling method is under development and needs a trained dataset of human annotations and keywords corresponding to the value in the specific contexts. Preliminary tests have been done in topic modelling and were able to identify 60% of the arguments for the PVE about relaxation of corona measures in the Netherlands and it could give an indication of their quantity. Also, a semi-supervised method Farthest First Traversal delivers promising results for topic modelling. Creating the value taxonomy saves 75% time compared to the manual creation of this taxonomy. This will help familiarizing participants and policymakers with the data by minimizing the duplicate arguments to be shown. Also, human value annotation will increase in efficiency if experts do not have to label duplicate motivations.

Knowing the value taxonomies and quantification of values helps the municipality to define their policy in its early stage and act upon the important values. This has been validated via interviews with stakeholders involved in the process of Súdwest Fryslân. The counsellor mentioned: 'Values will become prerequisites of the process as participants do not decide about a specific choice'. Values themselves are hard to translate to policy instruments, for that reason policymakers are interested in subcategories of arguments within these values. One option for visualization is to present the quantification of values/topics in a bar plot. If for example the value leadership is dominant, the municipality might use this as a mandate to take the lead as the public mentioned this to be important. A value hierarchy can be another more specific principle to subcategorize the values based on their arguments and prerequisites. Interviewees have seen high potential in using this instrument of value hierarchies for familiarization, discussion, validation and for coming up with more creative solutions to involve the expertise of the citizens. For all

manual and automated analysis transparency is mentioned to be most important. Stakeholders should know where the information comes from. Also, awareness about the uncertainties and biases of these models should be created.

Results from qualitative data analysis can support the process of policymaking, but they should never become a decision-making tool by itself. Their outcomes cannot be fully copied into decision-making, though it helps in creativity and reflection of the policy. Policymakers want to have outcomes of public evaluation and they want to be to act upon them. These value outcomes give substance to certain values and can give guidelines how to take these values into account. One way of doing this is visualising the values, norms and design principles in a value hierarchy using Value Sensitive Design. Based on interviews, there seem to be a high potential when values have been subcategorized by their corresponding norms and design principles. By doing this, knowledge of 'the public' can be used to support making policies.

It is recommended to continue doing research into the development of NLP within PVE to increase accuracy and reduce the negative influence of human biases. In this process values can be used for topic modelling, though insight in arguments and prerequisites will always be a requirement that should be taken into account.

Contents

Abbreviations

RIVM	National Institute for Public Health and the Environment
PVE	Participatory Value Evaluation
SWF	Súdwest Fryslán
NLP	Natural Language Processing
LDA	Latent Dirichlet Allocation
OECD	Organization for Economic Cooperation and Development
CBA	Cost Benefit Analysis
WTAPB	Willingness to Allocate Public Budget
WTP	Willingness to Pay
CAQDAS	Computer Assisted Qualitative Data Analysis Software
VSD	Value Sensitive Design

1. Introduction

The Climate Agreement of Paris that The Netherlands has signed in 2015 includes targets for sustainable development (United Nations, 2019). One of the topics includes electricity. Between now and 2050, limiting climate change will require a carbon-free electricity system. (Dutch Ministry of Economic Affairs and Climate, 2019). In 2030 The Netherlands targets to reduce carbon emissions by 49% and greenhouse gasses by 55%. Infrastructure has to change drastically to reach these goals, but this will come at the cost of our landscape. Another grand challenge is related to healthcare, specifically the making of policies in the context of an epidemic situation. The Dutch government has announced an intelligent lockdown that highly affect the daily life of citizens (de Haas, Faber, & Hamersma, 2020). All the measures to minimize the spreading of the virus, have led to side effects like mental risks and economic instability as well (Verbeek et al., 2020; World Health Organization, 2020).

These two examples can be seen as grand challenges and are a selection of projects in which policies affect the daily life of citizens. One of the key tasks of public organisations is designing their policies to reach their goals and improve welfare within the municipality. Most of the policies can expect social opposition, for his reason public participation is essential. In this research, two case studies have been conducted, one regarding the Sustainable Energy Strategy in Súdwest Fryslân (SWF), another one for reducing Covid-19 measures of an intelligent lockdown in The Netherlands. In both cases citizens are affected and would like to share their opinion which will increase support, quality and legitimacy of policymaking (Klijn & Koppenjan, 2000). Stakeholder engagement is essential and public participation is often obligatory when creating a policy proposal if citizens get affected by policymaking (Alemanno, 2015). Public participation is for example a requirement for environmental projects in most OECD countries (Bishop & Davis, 2002). For this reason, a participation plan needs to be established.

Based on Mostert (2003) there are different levels of participation for which a participation plan can be made. Information and consultation are two levels in which generally the most people are involved. Current methods for information and consultation are focusing on small scale participation meetings (Marissing, 2005). Different methods can be used for consultation on a larger scale, of which Participative Value Evaluation (PVE) is promising and relatively new (Mouter et al., 2019). PVE is an online evaluation method in which individuals can give preferences to allocate public budget. It is often used in the context of mobility, environment and energy transition. PVE can be seen to be an evaluation tool, though public participation can also be the main goal. PVE allows citizens to participate by informing them and letting them evaluate the proposed policies. The participation process of PVE might also include co-designing and discussing as well. This participation process can contribute to better decision-making (Koster, 2020).

The qualitative data of PVE is currently being analysed by the researchers themselves, which comes with two problems. Often, these researchers have been involved in the design of the PVE. For this reason, they might interpret values differently than the actual values of the participant of the study. Data analysts that are not involved in the process of the PVE design can be biased as well. This bias has different sources of which cultural dimensions are just one source of bias (Hofstede, 2011). The second problem of analysis done by the researchers is that it is time consuming to analyse all the qualitative data. This makes PVE less scalable if the advice should be based on all information that participants have filled in, while this is important for validity and support by participants.

Semi-automatised data analysis could partly solve this problem. A promising method is using Natural Language Processing (NLP). NLP is a subfield of artificial intelligence concerned with the interaction between computer and human natural languages. In particular programming the computer to process large amounts of natural language data. NLP is being used for many reasons, relevant examples are sentiment analysis, topic modelling and relationship extracting. The combination of using NLP to extract values for decision-making is a new research field. The algorithms for topic modelling are rapidly increasing in accuracy and often open source. Some studies have applied this method to value extraction and expect this to become more accurate in the near future (Liu et al., 2019, Wilson, 2019). If this method becomes accurate enough or if it creates new possibilities, it might contribute to a renewed time saving way of analysing the data of PVE.

Knowledge gap

Independent of the context of PVE, the output of this tool gives qualitative as well as quantitative data. Looking at the qualitative data, it is valuable to extract the useful information for policymakers. Based on previous PVEs this

useful information exists of values and arguments of the participants to be used for designing or validating policies. However, currently the process of analysing qualitative data is labour intensive and does not include the information of every single participant. This process of scanning the motivations until a point of saturation can lead to recommendations for policy makers as can be seen in the heat transition PVE in Utrecht. Still, the researchers scanning this data can be biased and will have their own perception on the values and arguments that they annotate in the motivations. This way of analysing the qualitative data did not give insight in the number of occurrences, not for the most common values, neither for the quantity of arguments or the number of conflicts between values and arguments. Currently, PVE is being used for more than 30.000 participants. These numbers of participants will make it impossible to analyse all data manually and required automated solution for PVE to be scalable. This leads to the following problem statement:

> *Participative Value Evaluation has a high potential to become a participation tool that can be used at a large scale. The current method of processing qualitative data is not scalable as is does not process all textual data and it does not quantify this qualitative data. It is also not known what information from this qualitative data is required for decisionmakers and how they will use this data in designing their policies.*

Research objective

This research will focus on finding an alternative way of analysing the qualitative data of a PVE to make public participation possible at a large scale. This includes the exploration and validation of the information from these participants that is requested by policymakers. Human values are one of information types that is demanded, an approach has to be found to select the values that will have to be extracted. Secondly, this research will focus on trying to automate the process of analysing qualitative data. The research will look at possibilities and their advantages and disadvantages compared to manual data analysis. This process needs validation to conclude if the alternative data processing is accurate, if the outcomes fulfil the goals of this analysis and if this analysis is less biased. Using NLP to extract the values can be the method, this process is the mean and not explicitly the goal of this research. Lastly, an advice will be given on how to adapt the PVE to increase the success of automated data analysis that can be used for policymaking based on reflection by experts.

Research questions

In this section, the research question is presented and explained. The objectives and sub goals have been defined and have resulted in the sub-questions. Literature has shown that there is a high potential for PVE. Also, the quality and accuracy of NLP techniques will grow in the near future. As PVE is limited in potential by the amount of qualitative data analysis, a solution can be discovered to semi-automatize these analyses. Therefore, the main question of this study is:

> *How to support public participation by using automated qualitative data processing in Participatory Value Evaluation to extract values that can be used in decision-making?*

Sub questions

To be able to answer the main question, sub questions have been defined. The objective of this study is to explore to which extend semi-automated data processing can support the analysis of qualitative data within PVE to increase the its applicability for decision-makers. Positively influence by supporting the process by means of speeding up the process, obtaining results with less biases and with the opportunity to dive deeper into the results. The objective can be translated in the following sub questions.

1. **What is the contribution of qualitative data within PVE in evaluation and participation? (Conceptualisation)**

   The use of PVE as a public participation method is growing. To be able to extract the desired information from the qualitative data it is essential to know why public participation is important. More important is how qualitative data resulting from this participation process has contributed in decision-making within previous PVEs.

2. **What are the limitations of the current data processing methods of PVE? (Conceptualisation)**

PVE seems to be a suitable method for public participation and consultation. However, a preliminary field study shows some limitations. Manual qualitative data analysis will be executed to get insight in these limitations and the time needed for analysis of qualitative data. Another common dilemma in manual analysis is bias. Cognitive bias is likely to influence the outcomes. We do not know the unknown and this could mean that there is always some risk left of untouched problems in the manual data analysis.

3. **What semi-automated data processing methods are available and what are their advantages? (Model formalisation)**

> PVE generates a lot of data. To be able to give policy recommendations, it is preferred to process as much input by participants as possible. The participants do expect that their input is taken seriously. Automatising the process means that more data can be processed and in less amount of time. However, this will also incorporate limitations in the gained results due to limitations due to a lack of human interpretation. For some types of automation, experts will have to train the algorithm by annotating manually to supervise the process of language processing. To answer the question which (semi) automated data processing methods are available, different methods and their potential application will be explored. As PVE generates a lot of qualitative data that is expected to be analysed, semi-automated data processing could give preliminary conclusions.

4. **How can (semi) automated techniques be implemented to automatize value extraction of the qualitative data of PVE? (Model implementation and data analysis)**

> This sub question should provide an insight on how automatised data analysis will influence the processing of data through NLP to get an understanding of the public values. NLP is a method to understand natural language by a computer is able to determine relations. Research has to be done if this method extracts the most important information.

5. **How do decision-makers interpret the model outcomes and is their interpretation in line with previous used methods? (Model validation)**

> In order to validate a new method, it should be compared to previous used method(s) of which the results are confirmed. This research questions will contribute in the research by getting to know which information has been seen relevant in the PVE of Súdwest Fryslân. First, this sub question will validate how value extraction can be valuable for decision-makers and what information is demanded from values. For this specific research, reflection on the use of context-specific values will be incorporated. The second part of the validation will explore the potential and the demand of automated data processing from the perspective of stakeholders in the process. This sub question will use interviews to let stakeholders reflect on the results from the qualitative data and to discover which opportunities they discover to apply this in their policy making processes. With this information the method of PVE to extract values can be evaluated and eventually the method could be adapted in order to obtain better results.

By using these five sub questions, the main question can be answered. Figure 1 shows an overview of the research questions and their approach. The full research approach is explained in chapter 2.

RQ	How to increase public participation by using automated qualitative data processing in Participatory Value Evaluation to extract values that can be used in decision-making?	
SQ1	What is the contribution of qualitative data PVE in evaluation and participation?	Literature study (exploration) Case studies (exploration) Interviews (validation)
SQ2	What are the limitations of the current data processing methods of PVE?	Literature study
SQ3	What semi-automated data processing methods are available and what are their advantages?	Literature study
SQ4	How can (semi) automated techniques be implemented to automatize value extraction of the qualitative data of PVE?	Case studies
SQ5	How do decision-makers interpret the model outcomes and it this in line with previous used methods?	Interviews

Figure 1 - Overview of the research questions and their approach

Scientific relevance

Research on value-deliberation, public participation and natural language processing is widely available. Scientifically, it is interesting to combine these methods in case studies to increase the success of each individual method. Especially automated language processing is increasing in quality and methods available. This is being used in political context for sentiment analysis (Ramteke, Godhia, Shah, & Shaikh, 2016), but not often based on personal or collective values. This technique will first be applied and validated based on qualitative data of case studies. Looking at the literature of values, there is currently not an approach to define a value taxonomy in the context of political decision-making (Fritzsche & Oz, 2007; Graham et al., 1991; Raaij & Verhallen, 1994; Schwartz, 1992; Wilson, 2019). This research will test the demand for a selected approach to define a value taxonomy that can be used for the labelling of qualitative data.
Future research might incorporate use of NLP to extract values that can be used within agent based modelling. Atkinson, Bench-Capon, Cartwright, & Wyner (2011) have created models that use sentiment analysis for deliberation. PVE itself is not a method build for deliberation, it is instead used for consultation. However, if values can be extracted automatically, this can be used to transforming individuals of agent-based models into agents that can interact and deliberate about a policy based on values. These questions will be out of scope for this research, it verifies the societal relevance.

Societal relevance

PVE is currently being used in many domains. Examples are healthcare, energy, infrastructure and water management. Besides these domains, there are many other potential applications in which PVE could increase the quality of decision-making. It has the potential to gather insights in many decision-making processes. For now, we can state that PVE is involved in decision-making processes in which decisions about grand challenges have to be made. To solve these challenges, agreements have to be made based on a high quantity of information. One of the sources that is becoming more important is the public opinion gathered by public participation methods. PVE is a tool in which a lot of qualitative data is incorporated. If the values of motivations can easily be extracted, it is likely to increase the quality of decisions made. Decision-makers can try to take into account the values and arguments that people find important. If the amount of data gathered by these public participation methods will rise in the near future, automated processing of this data will be essential. The societal relevance was confirmed during the start of this thesis. First, one case study was being used for analysis. Until I received a phone call from my supervisor who

mentioned that another PVE was being conducted. It was expected to have over 10.000 participants and the results had to be analysed within one week. My problem statement was confirmed. Automate data analysis is needed to save time and to potentially reduce the bias by data analysts or policymakers.

The scope of the research

The PVE for the municipality of Súdwest Fryslân, which deals with policymaking to reach sustainability goals, was the starting point of this research to dive into PVE and to improve the public participation process by focusing on qualitative data output. 'Relaxation of corona measures in the Netherlands', another PVE that aims at creating a policy to reduce policy-measures taken against the Coronavirus, came across during the start of this thesis and created the opportunity to apply automated data processing techniques at two cases that differ in size. The main focus of this research is the opportunity for automated data processing of qualitative data within PVE. Both case studies will be analysed based the information that is required by policymakers in this public participation process. To get to know this, public participation is within the scope to analyse the need of information and how this can be used for policymaking based on the level of participation that is required. The focus in the explained automated qualitative data analysis will be on human values. Defining a value taxonomy is an essential part for this analysis, though the philosophical approach of defining a value taxonomy is out of scope and recommended for further research. Validation of the outcomes is included in the scope and will be done via interviews. Overall, the research is meant to be an exploratory thesis that can be the start of future work.

Document Structure

This report consists of roughly four parts: 1) Orientation and conceptualization, 2) Analysis and implementation, 3) model validation and 4) Conclusion and recommendation.

The orientation and conceptualization include chapter 2, 3 and 4. This starts with the approach to be able to conduct the right literature research in chapter 3. This literature research focuses on reasons for public participation and information that policymakers gather from this participation method. Also, Participatory Value Evaluation is being proposed and the use of human values in decision-making. Research approach and background are a parallel process due to the explorative character of this research. Some background is essential to determine the research approach and the research approach influences the need for background information. The output of both chapters will be a background of public participation, PVE and value theories that can be used for further analysis. Chapter 4 analyses the case studies based on their motives and what their role is in the public participation process. This will give insights in their goals and information needs regarding the qualitative data analysis. The output of chapter 4 will be the definition of the context for each case study. This information is needed to perform relevant experiments for information extraction.

Next, chapter 5, 6 and 7 contains the analysis and implementation. Chapter 5 briefly illustrates to the reader the process of extracting information from the qualitative data output of PVE. This will give insight in the information that could be retrieved from the qualitative data of PVE. Afterwards, chapter 6 explores the possibilities to implement automated data processing techniques within PVE to extract information from its qualitative data output. The output of chapter 5 and 6 will give insight in the information from PVE and a background of techniques for extracting information. Next, chapter 7 will combine this information to explain all the automated data processing methods that have been explored in the case studies within this research.

Chapter 8 and 9 will reflect and validate upon the gathered results. First in chapter 8 by reflecting on the process of defining the value taxonomy in Súdwest Fryslân. One of the essential questions that had to be solved to perform this study was the definition of the value taxonomy, the conducted approach is reflected by stakeholders in this process in chapter 8. Chapter 9 will evaluate on the process and the possibilities for automated qualitative data analysis, this will reflect upon the potential of value extraction within PVE. This will be done based on interviews to let them reflect on the process and by asking them to think about future possibilities based on some visualisations of results.

Chapter 10 and 11 reflect upon the research question based on the results. First, the conclusion will be given in chapter 10. Afterward, chapter 11 will discuss recommendations and limitations of this study. An overview of this process can be seen in Chapter 2: research approach.

2. Research approach

This second chapter presents the methodological framework that is used throughout this research. The approach is built around the application of the PVE-method in a case study approach. The first section gives a structural overview of the approach. Next, data gathering, interview techniques and collaborations will be discussed. The setup and goals of the case studies will be discussed in chapter 3.

Overview of research approach

The research in this thesis will address a practical problem, how can the analysis of qualitative data within PVE be scalable. This research will be exploratory to find out the potential and risks of automating qualitative data analysis. The focus within this data will be on human values. To analyse the applicability, a practice-oriented research is suggested. According to the study of Verschuren & Doorewaard (2010), a case study is the appropriate research strategy. The research strategy can be perceived an embedded research strategy. Two case studies will be used to test the application of automatizing of qualitative data within the field of PVE. For these case studies it is important to make sure that they fulfil the requirements for value-based decision-making and the requirements for automation. These requirements will be discussed for both cases. The main disadvantage of experiments and case studies is external validity (Stoecker, 1991). For the PVE method to gain external validity, it should be applied on a wide variety of case studies. This study will only cover two different contexts of PVE. For that reason, the outcomes will be validated by interviewing decision-makers that have knowledge in a wide range of politics.

The methods that are being used in the case studies are focusing on the automation of qualitative data analysis, focusing on human values. These automating techniques will be selected based on literature and by co-operating together with the faculty of EEMCS of the Technical University of Delft. If the automation techniques involve machine learning, the technique of Natural Language Processing will be used. Next to the automation of data analysis, the case studies will be analysed based on the outcomes that can be gathered by manually analysing the qualitative data of PVE. The focus will be on value extraction from textual data to be able to compare PVEs of different contexts. For this value extraction, a well decided approach will be selected to create the value taxonomies for the two case studies. The grounded theory approach will be used together with value theories that will be selected during desk research.

This research is exploratory, for this reason other potentials next to value extraction will be tested to explore future applications of automated data analysis. The main purpose of this is to validate the potential of automation at decision-makers and to get to know the demand of the analysis that can be helpful for making decisions. An overview of the process can be seen in Figure 2.
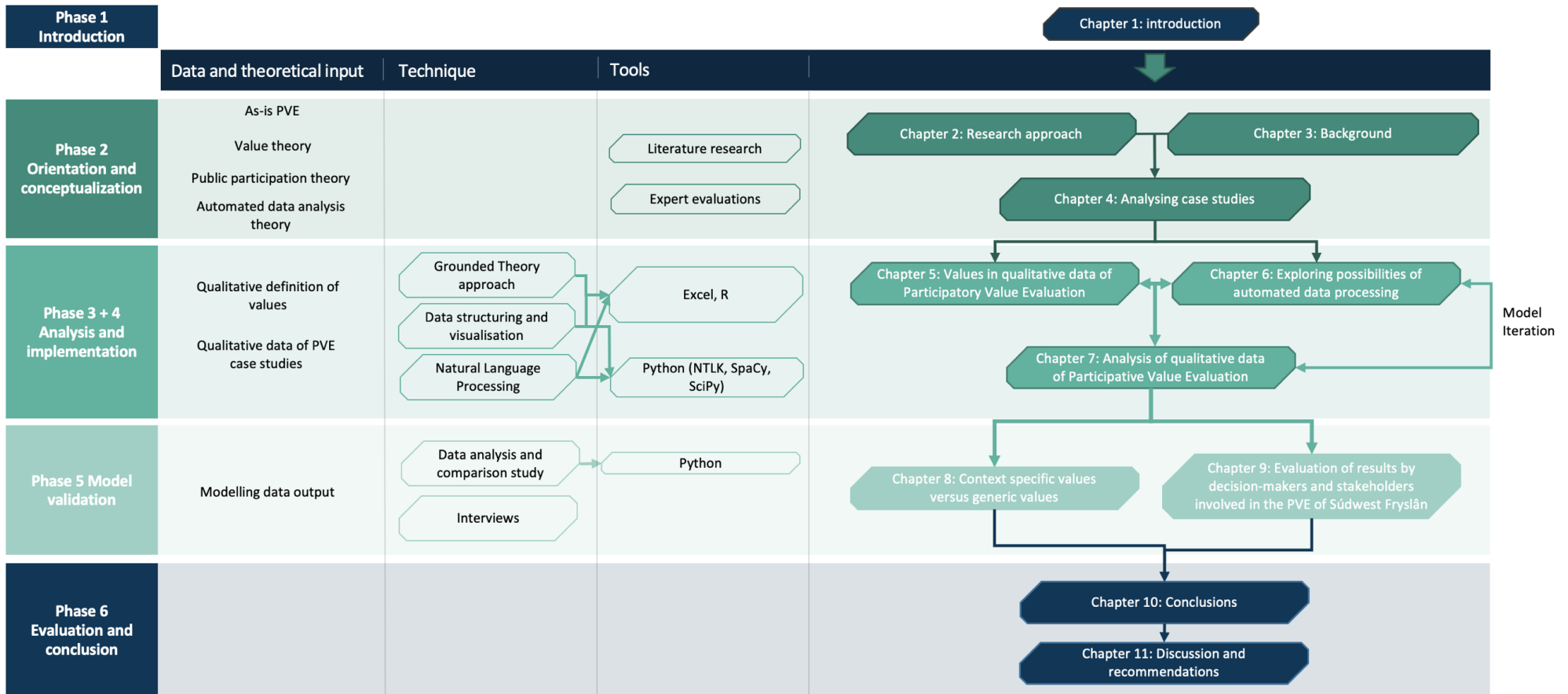
Figure 2 - Overview of the research process

Data gathering

The data for this research will be gathered via two case studies and via interviews. The first case study will be about energy transition. The municipality of Súdwest Fryslân has been conducting a field study by using this the method of participative value evaluation. The goal of this method is to let their citizens evaluate the process on who should take the lead in reaching their goals for the regional energy strategy. In total 1376 participants have completed this PVE. This data has been gathered in April and May 2020. In this case study, the municipality has been responsible for the participants and thus the data that is being used in this research.

The second case study has been conducted together with the Dutch Ministry of Health and Environment. The goal was to ask citizens of The Netherlands to select which policy measures could be reduced regarding the intelligent lockdown due to the Covid-19 virus. This data has been gathered between April 28 and May 3, 2020. This data was available for this research and consists of >30.000 participants.

Interviews

Interviews have been conducted to bridge the gap in literature regarding the usefulness of the information that could be gathered by manual and automated qualitative data analysis. The interviews take place within the context on the case study of Súdwest Fryslân. The people that have been interviewed will stay anonymous, but their relation or function to the project can be shared. The decision for this case study has been made due to three reasons.

- Súdwest Fryslân has explicitly asked to evaluate the opinions of citizens to support decision-making. The PVE for Covid-19 has many stakeholders that contribute their stakes and knowledge, the specific contribution of value gathering will be hard to determine.
- Súdwest Fryslân did set-up the PVE to be a participation process that they will use to make further decisions. It is willing to accept the values that are being advised to incorporate in the decision-making. Accepting the opinions about specific policy options cannot be done due to the practical conflicts.
- The timespan of conducting the PVE in Súdwest Fryslân is six weeks, this will give the opportunity to manually annotate the values before applying automation techniques.

The interviews were set up to be fulfilled in one hour, the questions have been a combination of general questions about the process and questions based on the visualizations that were shown to the participants. All interviews have taken place online and the audio has been recorded via Zoom.

The purpose of the interview is to let stakeholders reflect on the results from this research and to ask them which opportunities they see to apply this in their policy making processes. With this information the methods for automation can be evaluated and eventually the method could be adapted in order to obtain better results. The interviews are conducted by a semi-structured list of questions and will include five domains:

1. General experiences of PVE within the process of Súdwest Fryslân
2. Needs in decision-making from participants' opinions
3. Experiences with human values in the decision-making
4. Interpretation of results in Súdwest Fryslân
5. From values to design principles

Reflection of these five domains can lead to conclusions, discussions and recommendations for this research. It will mainly be used as a validation study of the research.

Collaborations

This research has been created in with the help of other researchers of the Technical University of Delft. Part of the programming and optimalisation of the Natural Language Processing methods have been developed by the researchers Enrico Liscio and Pradeep Murukannaiah, both members of the department Interactive Intelligence of the faculty of Electrical Engineering, Mathematics & Computer Science of the TU Delft. The method of Word2Vec, CorEx and the online annotation platform have been developed by this department. This method has been reprogrammed to be used in the context of the case studies in this thesis. The manual data annotation in chapter 6 has been done together with the student assistants Wim Gommans and Poppe Jytsma and with the help of Dr. Shannon Spruit. The method Wordify for sentiment analysis has been developed by Pietro Lesci and Dirk Hovy of the University of Bocconi in Milan, Italy. Dirk Hovy has advised on the use of this method. The interviews have been done with participants from the municipality of Súdwest Fryslân, NPBO and Public Mediation.

Background

The background of this research will give information about the state-of-the-art in public participation and Participatory Value Evaluation (PVE). It will discuss public participation in the context of PVE. Next, the concepts natural language processing and values in decision-making will be introduced. These concepts all together will illustrate the relevance of this work and they will form the base of the knowledge gap

What is public participation?

Within environmental decision-making, stakeholder participation takes a central role and public participation is often mandatory. In its most general sense, stakeholder participation refers to a process that facilitates the inclusion of those involved in, affected by, knowledgeable of, or having expertise or experience relevant to the issues at stake (Pigmans, Aldewereld, Dignum, & Doorn, 2019). Public participation is a subfield of stakeholder participation and has many different definitions in literature. (Mostert, 2003) defines it as direct participation by non-governmental actors in decision-making. In this research public participation is the involvement of the citizens to inform them and let them provide input in the ongoing decision-making. Often the local citizens are individuals that together can be perceived a collective stakeholder. These stakeholders might be affected by a process and need to be involved in the decision-making process. When applying participatory processes often people are affected by the consequences of policy measures. Also, public budget is being used to realise these projects, which will be at the cost of other projects. For that reason, it is essential to consult and inform citizens about the decision-making. Studies have also shown that public mediation contains the potential not only to reduce or solve conflicts among citizens, stakeholders, and policy-makers, but can also increase the effectiveness and the substantive quality of decision-making and foster social cohesion and urban democracy (Itten, 2017). Also (Ruiz-Villaverde & García-Rubio, 2017) has shown that more transparent decision-making processes increase public acceptance and reduce litigation, delays and inefficiencies (Ruiz-Villaverde & García-Rubio, 2017). Participation does not mean decision-making itself. Some benefits and problems of public participation can be found in Figure 3 and are based on participation in water management project. All of them are applicable for most environmental projects that is often the case for PVE.

Potential benefits	Potential problems
Better-informed and more creative decision-making	Reluctant government that gives no serious follow-up, resulting in disappointment and less public acceptance of decisions.
Greater acceptance of decisions, fewer implementation problems	Limited and unrepresentative response
Social learning of all involved	Low-quality response
More open and "integrated" government	Inconsistent decision-making
Enhanced democracy	Costs and time
Environmentally and economically sustainable management	

Figure 3 - Potential benefits and potential problems of Public Participation (Mostert, 2003)

Public participation methods

A democratic society indicates that citizens have some power to influence decision-making, formally by choosing the governing legislation. However, democratic does not mean that the public is involved in specific decision-making processes that affect them. Public participation is the active participation of the public in decision-making. Public participation improves the quality of federal agencies' decisions about the environment (Dietz & Stern, 2008). Well-managed public involvement also increases the legitimacy of decisions in the eyes of those affected by them. Public participation is a requirement for environmental projects in most OECD countries (Bishop & Davis, 2002).

In the recent years a new form of participation to accompany interactive decision-making has become more popular (Klijn & Koppenjan, 2000). However, several studies that have tried to involve citizens in the design of their technology did not succeed in reaching discernible political impact (van Oudheusden, 2014). The low impact might be explained by several problems from which real-life participation exercises suffer. One of the challenges is about representativity, when can participants be perceived to represent "the public"? If there is a need for including interested, concerned and informed citizens, this tends to require a lot of time and is most likely to attract biased participants. In policymaking these citizens are likely to gain from potential decisions, for this reason they want to spend time and think along in the process (Irvin & Stansbury, 2004). The other type of participant that has the time is typically described by Niek Mouter being the 'retired TU Delft professor that has a lot of technical knowledge' but might not be representative for the public. Another issue related to participation is that there are many questions

regarding the operationalization of citizen engagement to design these public values (van Oudheusden, 2014). Examples are: "What is the public? Should citizens be encouraged to reach for consensus?". Some of the challenges might be solved by the process in which PVE is being used as a tool in this participation process. These challenges will still play a role in the analysis of PVE as a participation method and will be discussed for the selected case studies.

To be able to determine if PVE is actually a participation tool, it is first important to know the characteristics of public participation. There are three reasons for public participation according to Klijn & Koppenjan (2000):
1. Creating support. Interactive decision making can be seen as an attempt to maximize support for policies and to minimize resistance by involving potential veto-groups in the process of policy formation.
2. Improving quality. By involving citizens in the process of policy formation, local information, experiences and preferences are mobilized in order to improve the quality of local policies.
3. Improving local democracy. The participation of citizens, consumers and local pressure groups can also be viewed as an attempt to bridge the gap between citizens and local governments. New forms of participation are part of a remedy for the existing crisis in local representative democracy.

To reach the goal of the participation process, a participation method is required. Participatory methods structure group processes in which non-experts play an active role in order to articulate their knowledge, values and preferences (Van Asselt & Rijkens-Klomp, 2002). Different participation methods can be used to involve stakeholders in the decision process. Van Asselt and Rijkens-Klomp (2002) have mapped these methods based on two scales. First, if the process is a goal or a mean. Secondly, if the participation is based on reaching a consensus or for mapping out diversity. Another way of defining the level of participation is created by The International Association of Public Participation (2019) and can be seen in Figure 4. They define this to be a sequential process in which every step is required to reach the desired state of empowerment.
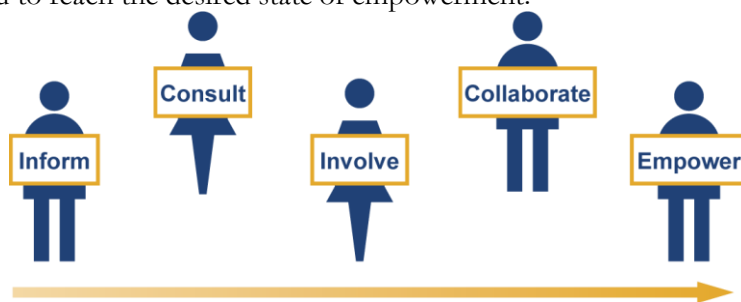
Figure 4 - Involve citizens in the planning (The International Association of Public Participation, 2019)

The last conceptt of categorizing participation in this paper is based on the concept of Mostert (2003), who has been inspired by the 'ladder of public participation' of Arnstein (1969). This concept shows a reverse correlation between the level of public participation and the number of people involved as can be seen in Figure 5. There are some exceptions for his theory as some ways of decision-making have a high level of participation, though still involve many people, in example a referendum. In general, this theory can be used to determine the number of participants needed in a public participation process. The type of public participation also influences the goals to be reached in this process. The objective and motive for participation will give insight in the information that is required by the decision-makers. The process of designing and analysing PVE is correlated with the objectives of the policymaker and can be taken into account when presenting the results of PVE.
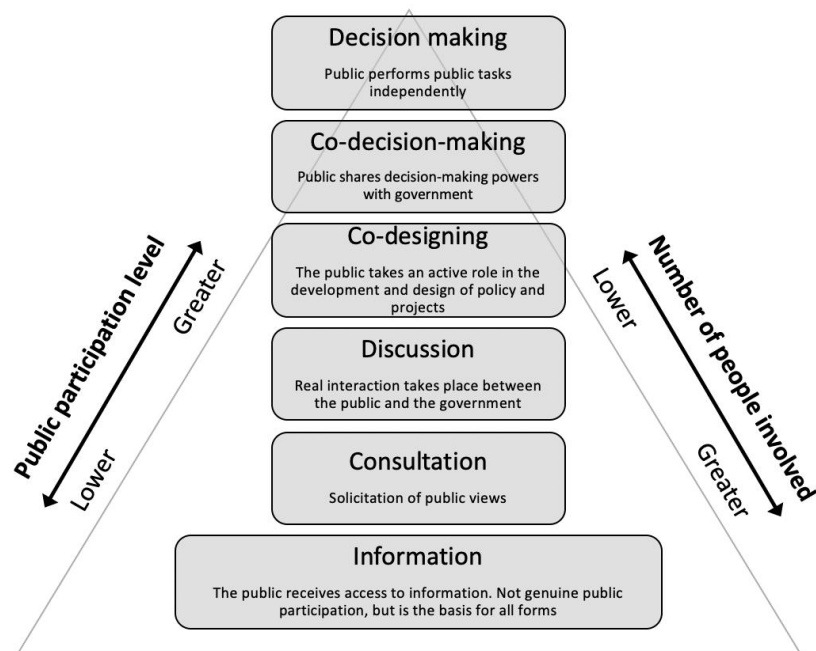
Figure 5 - Overview of public participation levels by Mostert (2003)

Participatory Value Evaluation (PVE)

This chapter will discuss the background of Participatory Value Evaluation (PVE) to understand the tool that is being used. First, the tool PVE will be shortly introduced to be able to familiarize with the concept. Secondly, PVE will be introduced based on the literature as a policy-appraisal tool or an evaluation tool. Lastly, PVE will be introduced being a participation tool next to its function of evaluation. By knowing the objectives of the evaluation and participations, the analysis of PVE can be done more specifically to fulfil the requirements of policymakers.

Process and overview of a PVE

A PVE is conducted via an online tool, an example of this tool can be seen in Figure 6. The tool is easily distributable and can be done from every location using a device connected to the internet. This allows for affordable citizen participation at a large scale because of the low entry barrier (Mouter, Koster, & Dekker, 2019). The tool places the citizen in the seat of the decision-maker. Participants can spend the public budget on a selected set of projects. The projects are defined by attributes, the participant can see these attributes and pick projects according to their preferences.

Afterwards the participants are asked to motivate their choices and to answer some questions, this can be demographical or socioeconomical depending on the context of the PVE. The generated qualitative data can give more insight into the quantitative data and it can include other types of qualitative information as well. The qualitative data can be codified to give quantitative results. In some PVE's, the decision can be delegated to an expert or other stakeholder. Additional qualitative questions might be asked depending on the context, an example can be to express potential concerns in the process or the level of trust in experts.
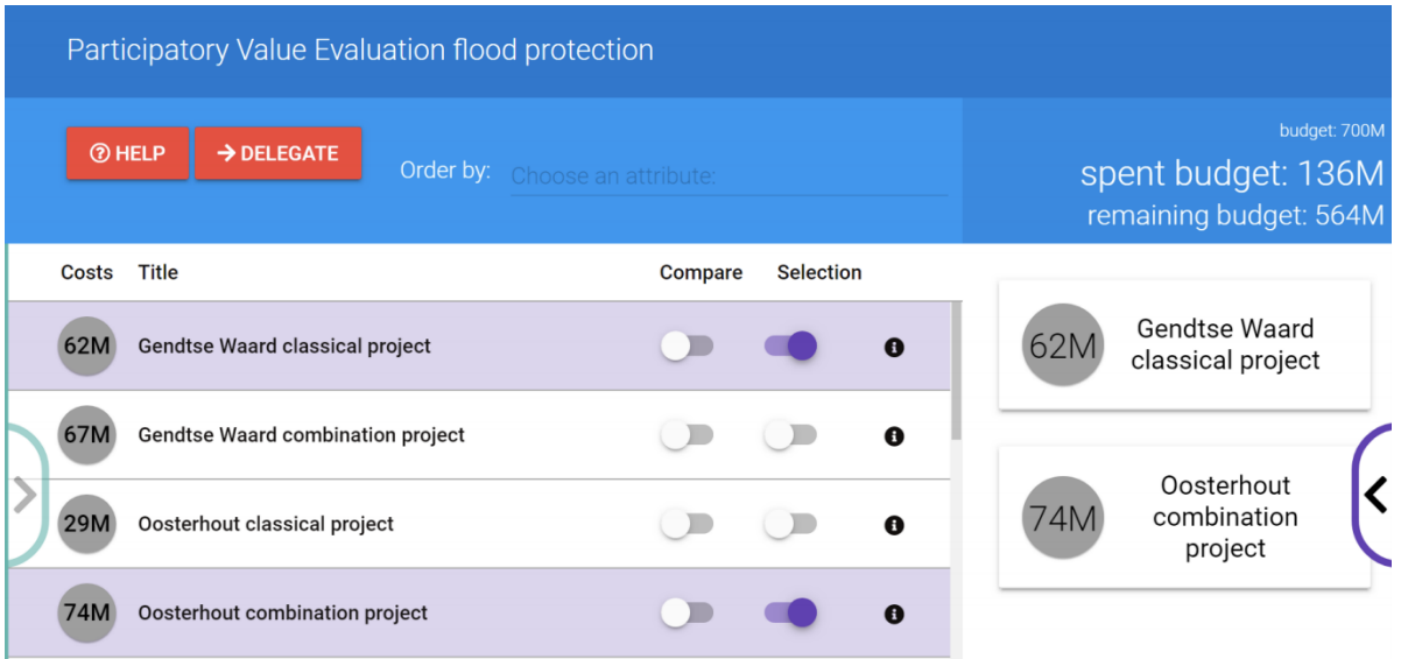
Figure 6 - Screenshot of a part of the page of the fixed-budget PVE on which respondents could select projects (Mouter, Koster, & Dekker, 2019)

Participatory value evaluation as an evaluation method

This first paragraph will introduce the initial purpose of PVE, which is a policy-appraisal tool. In the next paragraph, PVE will be introduced as a participation tool. Participatory Value Evaluation as a participation tool provides input for decision-makers and does not give official limitations in the decision-making. PVE is an online evaluation method in which individuals can give preferences to allocate public budget (Mouter et al., 2019). This might affect their own income and welfare. This method has been developed in response to the limitations of Cost-Benefit Analysis (CBA). CBA is a widely applied appraisal tool to support decision-making processes (Thomopoulos, Grant-Muller, & Tight, 2009). CBA determines the societal desirability and economic efficiency of a proposed policy or project (van Wee, 2012). This method is most known for infrastructural projects, but is also being used as a method in the policy domain of energy management (Sofia, Gioiella, Lotrecchiano, & Giuliano, 2020). By using CBA, the willingness to pay (WTP) can be determined and translated into monetary value. This can be defined as the amount of money one must take away from her in order to leave her with exactly the same utility level as she would have had without the environmental change (Nyborg, 2014). This can also be aggregated to achieve the net willingness to pay. CBA itself is in itself not a method for participation. It competes with PVE is such a way that both methods analyse and compare policy proposals. PVE can be seen a public appraisal tool as participants are often able to allocate public budget (Mouter et al., 2019). Nyborg mentions two concerns: It is not clear to what extend monetary value helps as there are many examples in which CBA is not influencing policy-making that much. Secondly, if the outcome of CBA is given in monetary terms it is essential that policymakers agree the value judgements and assumptions. Otherwise, they might disregard the analysis (Nyborg, 2014). Also, Mouter, Annema, & van Wee (2015) have analysed limitations on CBA and how to reduce its effects. Most stating problems are the incompleteness of CBA, uncertainty of effect estimation and the validity of effects that are hard to estimate. Going from willingness to pay to willingness to allocate public budget (WTAPB) let people decide on a set of policy options. This can solve some concerns as citizens tend to focus more on marginal utility and values instead of maximizing individual utilization. PVE combines WTP and WTAPB integrated in one valuation framework in which people can both value based on their individual utility as well as the collective utility. Validation that PVE solves some of the concerns of CBA can be seen in a comparison study in the context of urban mobility investments (Mouter, Koster, & Dekker, 2020). Many issues regarding sustainable energy supply are related to infrastructure, it has to be validated if the assumptions also hold in other sectors. For example, in the selected case study, safety tend to be an important value for individuals, this value will also be relevant in other projects. The use of PVE as a participation tool has already been validated in different sectors.

The current public participation processes are often based on information evening and participation meetings (Marissing, 2005). The average age at these meetings is relatively high and according to policymakers in Sûdwest Fryslân often people with a strong interest or dominant attitude are representing 'the public'. These policymakers also mention the influence that people have on each other when discussing. PVE is an online participation method that allows the municipality to talk with individual citizens, while allowing them to share all their thoughts and knowledge. The number of participants in an online platform is practically unlimited and increases the possibility of a more diverse group of participants. Knowing if and how PVE is being used in the participation process will help to extract the desired information for creating an advice to policymakers.

To know when PVE is being used as a participation method. The following questions can be asked according to the goals defined by Klijn & Koppenjan (2000).

- Is PVE being used to create support for the policymaking?
- Will the involvement of participants via PVE increase the quality of policymaking?
- Is the goal of using PVE to improve local democracy?

Afterward, the level of participation can be decided. Knowing and being able to communicate the levels of participation can help to define the value of PVE to participants and policymakers. If PVE has to be placed in the pyramid of Mostert (2003), it will always be consultation as inhabitants can share their opinions about a specific project. In this process of consultation, participants are able to share their opinion about the decisions that have to be made.

Next to consultation, PVE is an important attribute to inform participants. PVE informs people about the consequences of different measures, participants can compare different strategies and read the background. In the design of PVE, at least a minimal amount of background should be present. PVE can also be used as a helpful tool for discussing and co-designing policies, though they are both optional. It is possible to share your opinion and ask for more details based on the PVE. Also, an optional evaluation meeting can be organized to discuss the results with the participants. Co-designing is not always incorporated within PVE, it can be that inhabitants are invited on beforehand to help in creating the options of the PVE. Both case studies will be analysed based on the pyramid of Mostert (2003).

Natural language processing

This section introduces the practical application of natural language processing (NLP) and the state of the art of the application for value extraction. NLP is a subfield within machine learning and machine learning is a subset within artificial intelligence. NLP is a subfield of linguistics, computer science, and artificial intelligence concerned with the interactions between computers and human language, in particular how to program computers to process and analyse large amounts of natural language data.

Atkinson et al. (2011) has created a method to extract and define social values based on reasoning and incorporate them in the decision-making process. This method is based on manual data analysis and consist of a mathematical method to draw conclusions based on this data. NLP is a process in which language processing can be done automatically. NLP has many sub fields of which relationship extraction; topic modelling and sentiment analysis are most relevant in this research. Topic modelling can be used to define different values as being topics. The relations of words or sentence that are related to specific values can be extracted. An example of how correlations can be visualized can be seen in the study of Boyd et al. (2015). His research uses Amazon's Mechanical Turk that include a data set containing languages, values and behaviours. This trained data set can be used to automatically analyse other data to determine the relation between personal thoughts and values. This principle has been applied in different case studies. Liu et al. (2019) tried to determine values based on Amazon reviews to find out if behaviour results from values or personality. These values have been determined using NLP. Wilson (2019) has explored the degree to which we can employ computational models to infer people's values from the text that they write and the everyday activities that they perform. This is done by using topic modelling within NLP. Wilson created a model that determined the relationship between the textual behaviour of people and their values. His research also included predictions about the choices people will make based on their previous behaviour and values. The research of Wilson highly correlates with the goals of this research. However, the application within policy-making and public participation is not yet discovered.

Values in decision-making

Qualitative information can be used to extract other information like topics, arguments or practical design principles. The research question mentions the extraction of values, which is another category for extracting information. These values can be used as an argument in itself, though, often the value is incorporated in the argument (Maio et al., 2014). Also, other qualitative information can insist the favour or disfavour of certain values. This paragraph will explain what is meant with values and why insight in the human values can support decision-making.

Value Frameworks

A value is defined as "a concept explicit or implicit, distinctive of an individual or characteristics of a group of those desirable traits which influence the selection from available modes and ends of action" (Hills, 2002). The concept of values is defined in multiple ways and there are several points of views to interpret values. Values are linked to effect, so understanding values will help in understanding the public in public participation. Values motivate people's choices and guide the selection of action. Values are used to clarify the motivational principle of attitudes and behaviour (Schwartz, 2012). In daily life, personal values often form the basis to make a decision. The origin of these values are for example related to culture, education and environment (Hofstede, 2011). There can be significant variation between values within decision-making. There is no uniform set of values on which decisions can be made.

Schwartz is an example of a philosopher has defined ten core values (Schwartz, 2012) that are described in Table 1. Also, Figure 7 shows the dynamics of these values related to persons (Schwartz, 2012). In decision-making it is important that people have more tendency to a social focus.

Table 1 - Ten universal values by Schwartz (2012)

Construct	Description/Items: Individuals who value this believe in the importance of . . .
Power	. . . being in charge of people and resources and having money (social power, wealth, authority)
Achievement	. . . socially recognized successes (ambition, competence)
Hedonism	. . .sensual pleasure (fun, enjoying life)
Stimulation	. . .having stimulating experiences (daring, exciting life)
Self-direction	. . .independence of thought and action (creativity, freedom, independent, curious)
Universalism	. . . promoting the welfare of all people and nature (equality, social justice, protecting the environment)
Benevolence	. . .promoting the welfare of people you are close to (helpfulness, loyalty, honesty, forgiving)
Conformity	. . .controlling impulses to fulfill others' expectations (self-discipline, obedience)
Tradition	. . .maintaining traditions (moderation, respect for tradition, devout)
Security	. . .safety and security of self, family, and nation (family security, social order, clean)
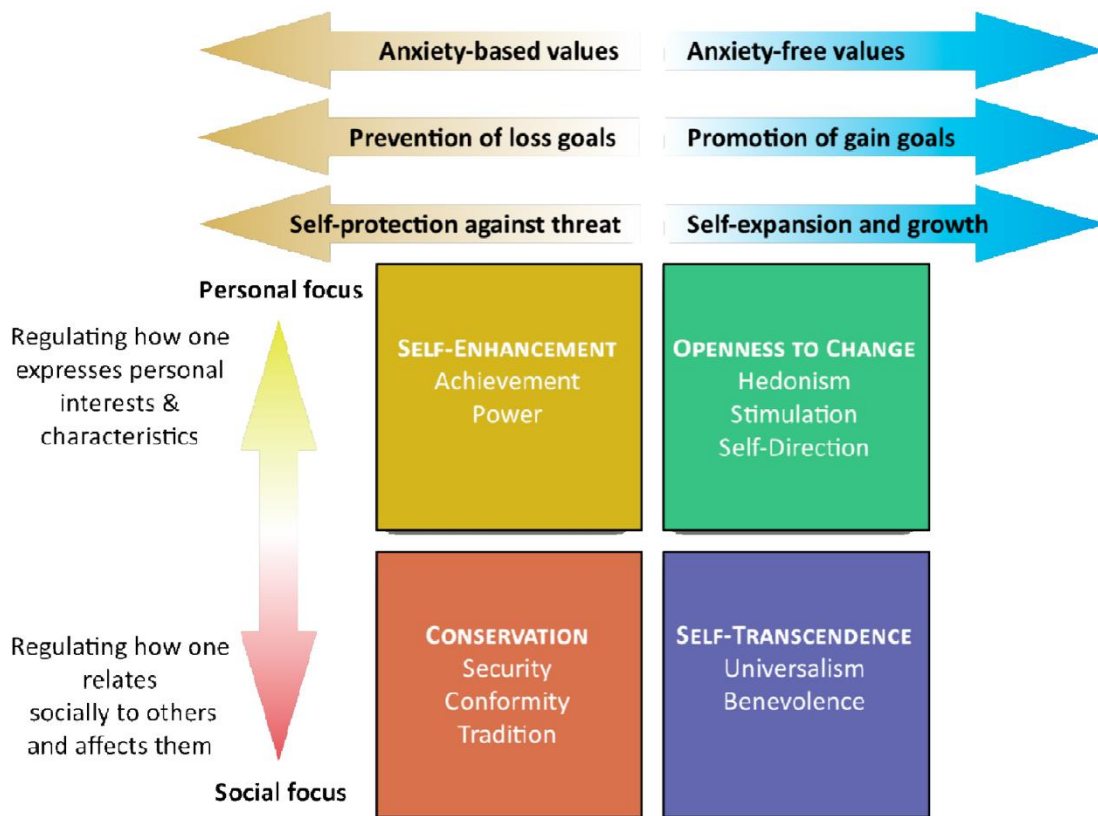
Figure 7 - Dynamic underpinnings of the universal value structure by (Schwartz, 2012)

Schwartz (2012) has defined ten broad personal values which form an important basis for values in general. However, using only a specific set of ten values might not fulfil the needs for policymakers. Kenter et al. (2016) have done research into value-based deliberation and how discussion based on values could contribute in group decision-making. A set of values had to be chosen in this study to evaluate the arguments based on values. For this reason, they have expanded the framework of Schwartz by categorising values based on the ten broad personal values Schwartz. If needed, Kenter et al. defined a subcategory. This brings this hierarchy to a total of 56 sub values to evaluate the discussion and argumentation. These sub values have been defined based on an online so called 'value compass' experiment.

Defining subcategories has not only been done by Kenter et al.. Fritzsche & Oz (2007) have observed that personal values have long been associated with individual decision behaviour. Their study tries examines personal values as they relate to five types of ethical dilemmas in organizational decision-making. Important note is that they are focussing on ethical dilemmas. To be able to valuate their participants, a vale model has been created based on four categories out of the ten values of Schwartz that are mostly related to ethical behaviour. As can be seen these categories are based on the structure by Schwartz in Figure 7. This shows that often values are being chosen based on the context in which there are being used. In this research the 'self-transcendence' by Schwartz has been replaced by altruistic values. It is interesting that this research has defined subcategories as well based on the input of the participants. An overview of such a values model can be seen in Figure 8. It can be seen that different values have been defined to be a subcategory, this makes it easier to label values and the process of labelling is more transparent.
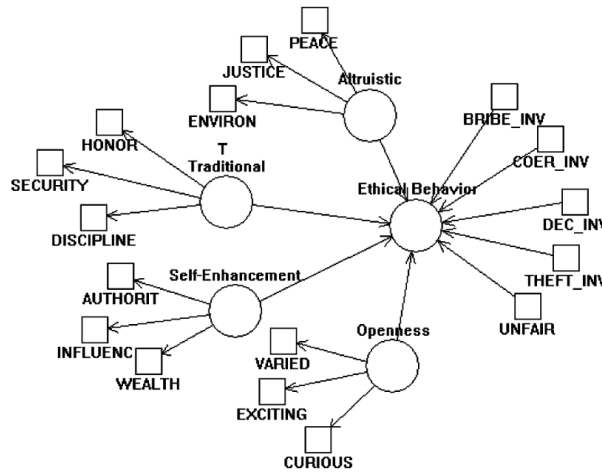
Figure 8 - Overview of a values model (Fritzsche & Oz, 2007)

The studies of Kenter et al. and Fritzsche & Oz are related to the study of Raaij & Verhallen (1994), that uses values to reflect upon products or activities. Raaij explains the theory in which some general values have a weak link with the product, activity or policy, but this value does have a stronger link with the subcategory. This can be useful when detecting values based on relationships between a general value, domain specific value and the 'product'. A Schematic overview can be seen in Figure 9. This method could be a practical way to define a set of domain specific values based on a set of general values.
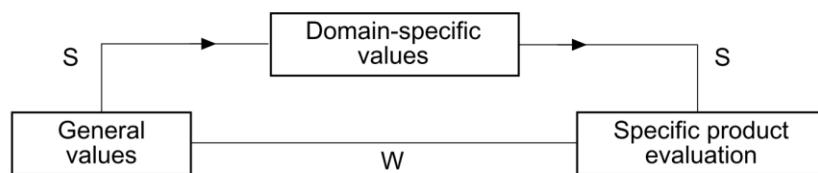
Key:
S: strong relationship
W: weak relationship

Figure 9 - The intervening role of domain specific variables (Raaij & Verhallen, 1994)

Another different value hierarchy has been created by Jesse Graham and Jonathan Haidt. It has categorised the general values in five contrasting values that can be used in the moral domain (Graham et al., 2013). These values have been placed in a moral matrix that has been described in the book 'The Righteous Mind' by Jonathan Haidt (2013). His work focusses on how different political preferences are related to values. Haidt argues that humans have six "Moral Foundations" through which we view politics and policy: The Care/Harm, Fairness/Cheating, Loyalty/Betrayal, Authority/Subversion, Sanctity/Degradation, and Liberty/Oppression foundations. The book does not mention specific project or process preferences, it is about how humans perceive their position related to politics. An example of two political streams and their corresponding values can be seen in Figure 10.
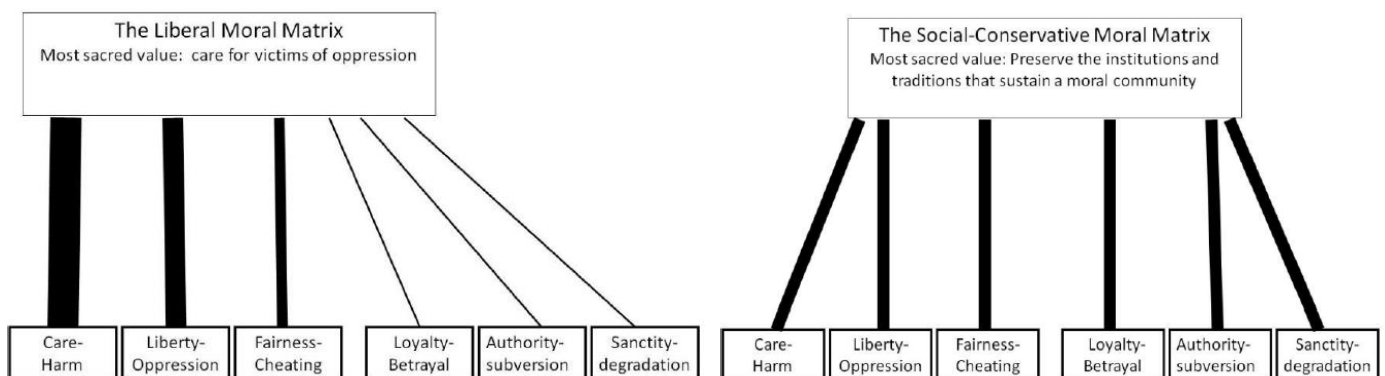
Figure 10 - The moral matrix of American liberals (left) and the social conservative moral matrix (right)

Application of values in decision-making

As can be seen before, different value taxonomies have been defined, often based on the ten core values of Schwartz (Fritzsche & Oz, 2007; Kenter et al., 2016; Raaij & Verhallen, 1994). These value taxonomies are often defined for a specific context. Knowing these values and acting based on them is an important attribute in decision-making. Within the context of PVE, participants can be biased. They might lack in knowledge about constraints of the projects, technological capabilities, institutional procedures and interests of other stakeholders. This lack of knowledge has not been defined. Having a lack of this knowledge could be a reason to adapt the process of participation or evaluation. One way of doing this is by valuating the motivations of participations based on values. The circumstances can be examined as a third party in the light of personal and organizational values. This has proven to help people reach objective outcomes (Sabini & Silver, 1983). One of the reasons is that values can be perceived valuable principles in human's nature on which people create their opinion and choices (Stålhammar & Thorén, 2019). If decision-makers can get to know the values at stake in a specific context, they can take care of these essential values in their choices. The process of consultation and discussion based on values helps bringing policymakers and participants on the same level. Adapting policies to these values will lead to less resistance and value-based evaluation reduces participation based on personal interests, at least in the context of water governance (Pigmans et al., 2019).

A practical method to incorporate values in a design, the Value Sensitive Design (VSD) principle by Van der Poel (2016) can be used to transform the determined values into practical guidelines to help in decision-making. Specification of these design requirements or design principles may be defined as the translation of a general value into one or more specific design requirements. This translation may be broken down in two steps:
1. The translation of a general value into one or more general norms
2. The translation of these general norms into more specific design requirements.
Values are relevant for evaluating the worth or goodness of certain options or objects. However, they do not directly imply certain prescriptions or restrictions for action. The translation from values to reasons is to be made first, which can be done by defining the corresponding norms. Subsequently, norms can be translated into different actions, that are defined as the design requirements. A value hierarchy based on VSD can be seen in Figure 11. In the context of PVE, these design principles might be used as practical guidelines for policy makers. The theory of Van der Poel (2016) is focusing on engineering design based on values. There are some of the general characteristics of the translation of values into requirements in design that makes VSD also applicable in other domains. This value hierarchy can be built top-down as well as bottom-up, this can be interesting when combining this method with PVE as design requirements can be made top-down by decisionmakers, as well as bottom-up by participants.
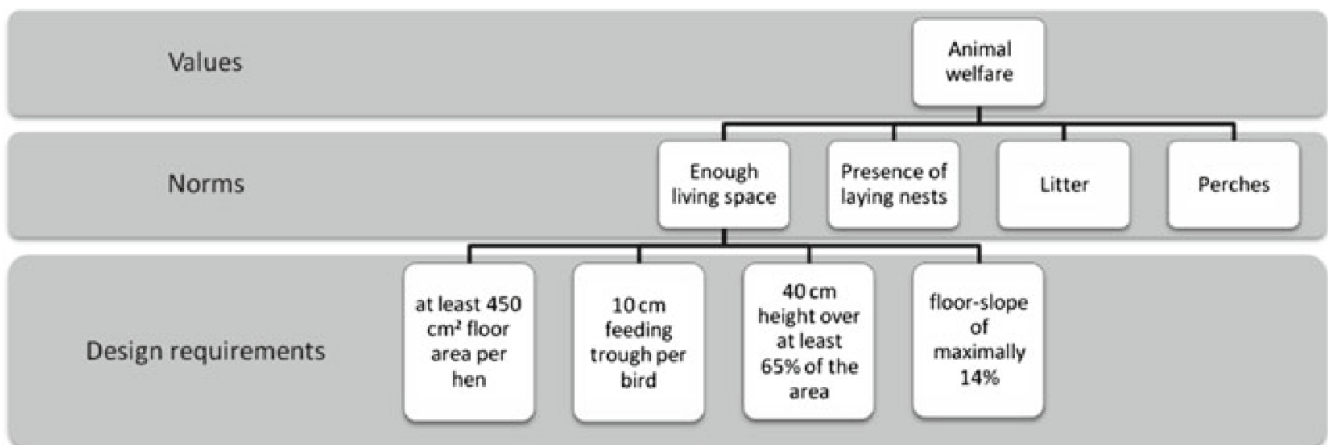
Figure 11 - Example of Value Sensitive Design (VSD) by Van der Poel (2016)

Also Kenter et al. (2016) addresses the contribution of values in discussion and decision-making. In group-based deliberation, participants were confident about values and felt that deliberated values should be used in decision-making. The results of this study (albeit with a limited sample size) suggest that shared values may be a better reflection of welfare implications than non-deliberated individual values, while at the same time more reflective of participants' transcendental values: their broader life goals and principles. This principal is focusing on group-based deliberation, which will be of relevance in the process of Súdwest Fryslân.

Summary

Participatory has been explained as being an evaluation and participation tool. Literature research shows that evaluation has been the main focus and origin of PVE. If PVE is being used as a participation tool, the process should be evaluated to include the desired levels of participation (Mostert, 2003). A background of values has been given. It seems that many different value taxonomies can be created, often based on the ten core values of Schwartz (2012). The values of Schwartz have been defined to be personal values. Many studies that have been performed in different contexts, do use different value models. However, they are often based on the personal value theory of Schwartz (Fritzsche & Oz, 2007; Kenter et al., 2016; Raaij & Verhallen, 1994). Most studies define subcategories of values based on the context in which they are being performed. When conducting this research, it will be essential to select a set of values that has his background based on value theories. To compare and apply the results based on values in different contexts, a replicable process for creating a value taxonomy is recommended. Most literature mentions the need for a context specific set of values for practical use, this can be taken into account when analysing the experiment results based on the values that they incorporate. Evaluating the set of values based on its contribution in decision-making can help in increasing the practical application for the policymakers. The desired goals of the policymakers can be taken into account to select values based on a moral, political or personal context. One practical implementation on how to use the values as a policymaker can be suggested based on Van der Poel (2016).

Analysis

First, the two case studies will be introduced in chapter 3. The reason for the project, the goals and the process will be discussed of the cases in general. Secondly, the goals, process and data of the PVE will be discussed within the context of the specific case. After the introduction of the case studies, some general research will be proposed in chapter 5 and is about automation of qualitative data analysis that can be tested on the case studies. To be able to use automation, it must be validated that values are relevant in both cases. An introduction will be given of the values within the motivations of both PVE's in chapter 4. This part forms the start of the analysis and research, based on this further testing can be done on the case studies.

The experiments will be done for both case studies in parallel in chapter 6. Both case studies have used manual analysis of qualitative data as well as automated data analysis. To be able to propose the advantages and disadvantages of both methods, the manual analysis and automated analysis will be compared based on uncertainties and purposes. Some analysis can be done only manually, some only automated and some analysis can be done in both ways. These analyses will be discussed based on statistical validity as well as validity based on interviews with the decision-makers.

3. Analysing case studies

To be able to detect and evaluate the values within a specific context, the case studies will need to be examined. It is important to determine the original reason for development as well as the goals that the PVE is aiming for. This chapter analyses two case studies that have involved PVE as an appraisal tool. Appendix A until D will give more detailed information about the online PVE that has been developed.

PVE 'Súdwest Fryslân'

The Climate Agreement of Paris that The Netherlands has signed in 2015 includes targets for sustainable development (United Nations, 2019). One of the topics includes electricity. Between now and 2050, limiting climate change will require a carbon-free electricity system. (Dutch Ministry of Economic Affairs and Climate, 2019). In 2030 The Netherlands targets to reduce carbon emissions by 49% and greenhouse gasses by 55%. To achieve this goal, provinces have created a regional energy strategy (RES). Municipalities must propose a policy to define their contribution for the RES (Rietveld, 2020). This also holds for the municipality of Súdwest Fryslân (SWF). In this policy proposal the support of their inhabitants as a source of information is important in defining the strategy for zero emission electricity consumption (Straver, 2018).

The province of Súdwest Fryslân is preparing their proposal for a policy to provide to reach their energy goals for 2050. This means that the province will become energy neutral in 2050, households will have to be energy neutral in 2030 (Gemeente Súdwest-Fryslân, 2018). The municipality of Súdwest Fryslân wants to cocreate their policy together with their inhabitants. One of the reasons is the resistance in previous projects that are related to sustainable energy. The municipality of Súdwest Fryslân has the largest area of all municipalities in The Netherlands. For this reason, it is inevitable for them to play a role in the wind energy. Of course, this leads to resistance of the citizens. The collaborative group of citizens is defined as Mienskip, which means a group of people having the same interest that is their province and its landscape. In 2019 a wind energy project in Hiddum-Houw has not been consulting the Mienskip enough according to the citizens. A documentary has been made to show the anger and frustration as a result of a lack of participation (Hoogendoorn, 2017). This process is still leading to a lot of criticism (Hoving, 2020). One of the lessons that the municipality has learned is that evaluation and participation is essential, starting at an early stage, this has been validated in the interviews with the government (Appendix L). The second example is 'wind park Fryslân' which is located at the south of the Afsluitdijk. The newspaper article of the Frysian daily paper illustrates the conflict between energy interests and the villagers (Jansons, 2020). This has also been validated via interviews as one of the reasons to change this procedure for the Regional Energy Strategy.

Súdwest Fryslân has tried to let public participation and public evaluation be the main focus in the process of creating a policy that contains their contribution to the Regional Energy Strategy. Every subregion has to place their bid, though, the province of Fryslân is divided in two regions, Súdwest Fryslân and all other municipalities together. The reason for their separation is that Súdwest Fryslân is providing 70% of the wind energy of the province. Due to the conflicts in the past, the municipality of Súdwest Fryslân is the only municipality that decided to create their own

policy and to direct their own energy transition. As they are part of an energy region, they have to negotiate their final plans with the other subregions within their energy region.

The councillor of sustainability Erik Faber has said in the paper Binnenlands Bestuur (Hendriksma, 2020):

"First, we want to know how our inhabitants want to shape the energy transition"

For this reason, they have decided to ask the organization NPBO (Dutch Platform for citizen participation and public policy to think about the process that ensures to let citizens play a central role. NPBO involved the TU Delft and the company Public Mediation to take the lead in this process in which the councillors play a minor role. The municipality explicitly did not involve themselves and they did not set any restrictions to make sure the public or Mienskip will interpret the PVE results. To be able to know the interests of a significant group of people, PVE has been chosen to evaluate the view of their citizens on the regional energy strategy. (This PVE can be reached via http://pve.splicedgene.com/sudwest-fryslan.)

PVE in this case has partly been used an evaluation method, but mainly as a participation tool. The last one is quite new as being the main purpose of PVE. The necessity of PVE as a participation can give insight in the relevant information and in the number of participants that are expected to participate in this process. To test the level of participation of PVE as a participation tool, the following questions will be answered based on participation goals by Klijn & Koppenjan (2000).

- Is PVE being used to create support for the policymaking?
  The municipality has decided to start the participation process before making any plans or decisions. They do not ask to evaluate their policy, instead they ask their citizens to think of the process, potential options and design principles. The decisionmakers will continue their process of policymaking about the RES after the consultation of the public.

- Will the involvement of participants via PVE increase the quality of policymaking?
  The goal of the decisionmakers is to involve the Mienskip (their citizens) in the process. Based on interviews a councillor has mentioned: "We know that the experts are part of our community, we do not want to rely on external experts or only our own expertise. We hope that the participants will share their expertise during this participation process and afterwards". This is a signal that public participation has a goal to improve the quality of decision-making.

- Is the goal of using PVE to improve local democracy?
  One of the interviewees has mentioned that to involve the Mienskip in this decision as it is important that the energy transition is a process that is being done together. He mentioned it would be great if the RES will be a way to bring the Mienskip together. The role of the specific tool PVE as a way to improve local democracy has not explicitly mentioned. It is important to mention that PVE has been chosen as the municipality was not able to come up with an alternative method that involves this number of citizens. This is essential as previous projects have faced a lack of local democracy (Hendriksma, 2020; Hoogendoorn, 2017).

As can be seen there is a demand for PVE as a participation tool to involve as many people as possible. An invitation has been sent to 40.000 participants of Súdwest Fryslân. As participation is the main purpose, this case will be interesting for research on automated data processing as a higher quantity of data increases the results of Natural Language Processing. Involving more citizens will increase the support, quality of decisions and local democracy. The municipality wanted an innovative way to inform the most citizens, to consult their view on this process on a large scale. They want to discuss the advice and concerns with a selected group of participants, and they want to partly co-design their policies with this group. An overview of the participation process in Súdwest Fryslân based on Mostert (2003) can be found in Figure 12.
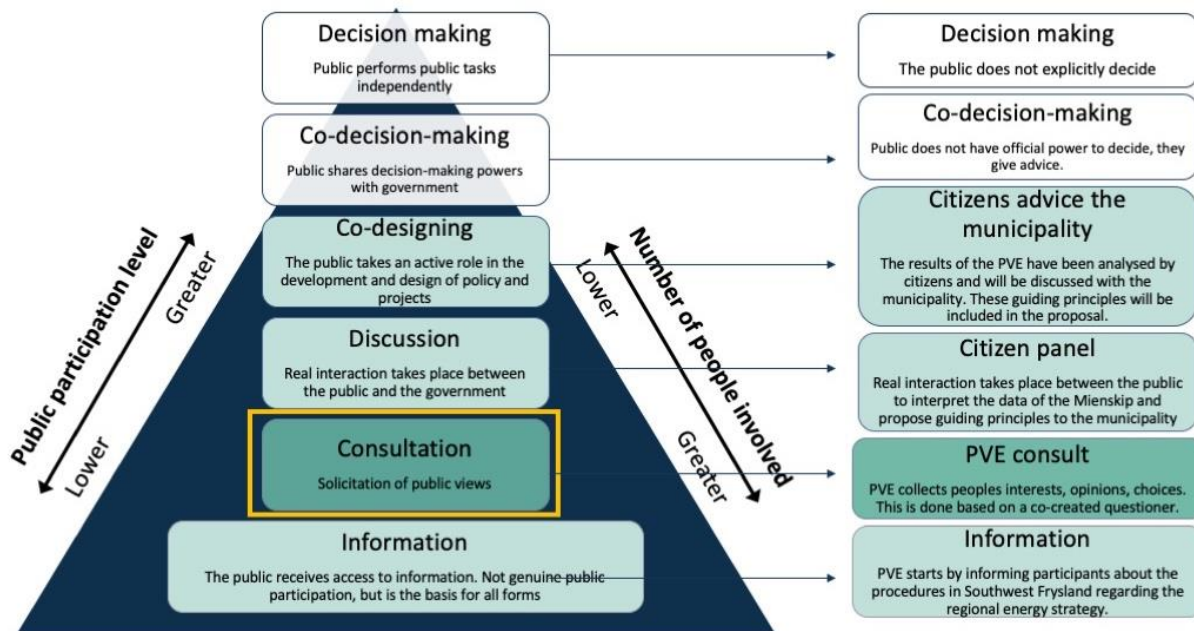
Figure 12 - Overview of public participation process for Súdwest Fryslân based on Mostert (2003)

Participative Value Evaluation in Súdwest Fryslân

As Súdwest Fryslân wants their citizens to think along in the process of defining their regional energy strategy, a lot of technological knowledge is involved. By the use of PVE, policymakers can gain insight in the technological questions of citizens as well as other constraints or worries. Most important in this PVE, is getting to know the core values of citizens related to this new energy strategy. To support legitimacy, this process has been guided by external process managers that have involved citizens in all stages of setting up and reflecting upon the PVE. First a hackathon has been organized to co-create the PVE together with a selected group of citizens. Secondly, over 40.000 participants have received an invitation to join the PVE. In this questionnaire, participants have divided one hundred points among six policy options, these policy options can be found in Appendix A and a visualisation of the PVE can be seen in Appendix B. The qualitative data is gathered by asking participants why they have chosen a specific policy and what concerns they might have. Extracting the human values has been done by manually analysing and annotating all the qualitative data input of the 1356 participants next to the statistical analysis of the quantitative data. Afterwards, the results have been analysed by a citizen panel that co-created an advice and guiding principles for the municipality. Members of this citizen panel have given the final recommendations to their local government. To get an overview of the most important participation methods used for Súdwest Fryslân, figure 12 has been adjusted for the PVE in Súdwest Fryslân. This overview can be seen in figure 4.1. The central participation method is PVE for which an invitation has be sent to 40.000 inhabitants of the municipality.

PVE 'relaxation of corona measures in the Netherlands'

The year 2020 can be perceived the year of Covid-19. This epidemy has impacted the daily life of the Dutch people. Employees had to work at home, some professions were temporarily not allowed, the catering industry had been closed, universities transformed to online education. Also, many people have been infected by the virus, that resulted in fatality for some of them. The Dutch healthcare system had to shift their focus on the virus to increase the amount of beds on the intensive care available for people detected with Covid-19. All the measures have decreased the spreading of the virus, but have led to mental risks as well (Verbeek et al., 2020; World Health Organization, 2020).

Participative Value Evaluation in PVE 'relaxation of corona measures in the Netherlands'

The National Institute for Public Health and the Environment is one of the clients that has asked for a PVE regarding a strategy on how to reduce the Covid-19 measures. This PVE can be found via http://pve.splicedgene.com/raadpleging-versoepeling-coronamaatregelen, also an overview of the online PVE can be found in Appendix D. By using PVE, the participants were able to take the seat of the policymakers in a topic that highly influences their daily life. In this case study, the participants had to choose between policy measures that influence the pressure on the healthcare system. The selected set of different policy measures can be found in

Appendix C. These measures result in a percentual increase of healthcare capacity due to the increase in number of patients. The total pressure on the healthcare system is not allowed to increase more than 50 %. However, 0-25% has been specified as no extra pressure, 25-40% as overpressure of the healthcare system and 40-50% as heavily overpressure of the healthcare system. After selecting a specific set of measures, participants were asked to elaborate on their choice. The question asked for their projects was the following: 'Please motivate your selection'. In this process, 30.000 participants have taken part, an overview of the percentual selection made by participants can be found in Figure 13.
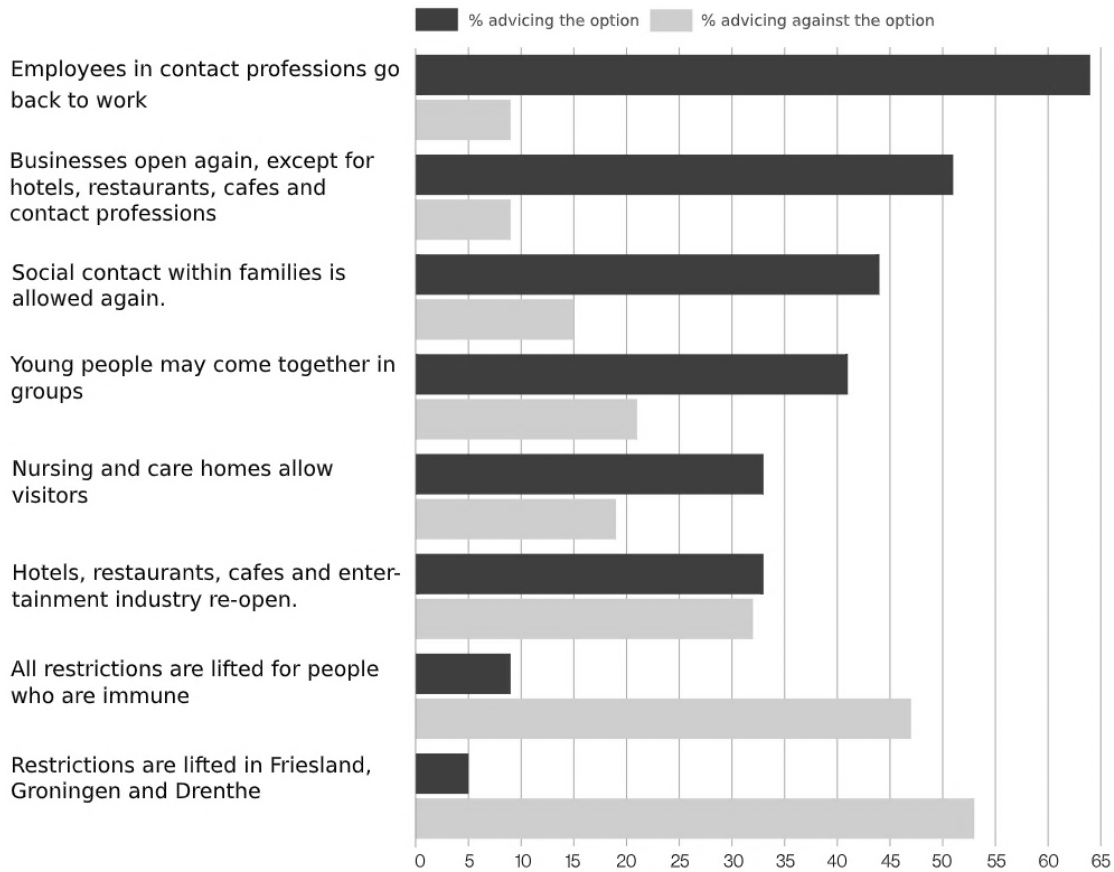
Figure 13 - Percentage of participants that have selected the policy measure

It is important to mention the different use of PVE compared to Súdwest Fryslân. In Súdwest Fryslân, the focus of the PVE was the highest possible participation of citizens. No decisions about the Regional Energy Strategy would be made without involving citizens and a co-created decision. To validate the use of PVE as a participation tool, the following questions will be answered based on participation goals by Klijn & Koppenjan (2000). The clients for this PVE are the National Institute for Health and Environment and the House of Representatives. These decisionmakers have not been interviewed like the policymakers for Súdwest Fryslân.

- Is PVE being used to create support for the policymaking?
  The impact of the PVE results on policymaking is not exactly clear. The results have been presented just before the national announcement to reduce the Covid-19 measures. The RIVM is expected to use them for their advice to the House of Representatives, the contribution of PVE in this advice is not defined. If the results of this PVE correspond with the decisions made, it can create support for the decision-making.

- Will the involvement of participants via PVE increase the quality of policymaking?
  The quality of policymaking can be improved as 30.000 citizens have shared their opinion. Many specific design principles have been mentioned by the participants that can be used for the RIVM. It is not defined which principles have been used by the policymakers and therefore it is hard to say if the quality has been increased based on this PVE.

- Is the goal of using PVE to improve local democracy?
  The government often used the term 'together' in their communication. PVE can help people to have the feeling of being together. However, this is not officially defined as one of the goals for this PVE. Also, PVE is being used to evaluate the public opinion, it might bring citizens together. It does not have any official influence in decision-making, so it might not contribute to democracy. On the other hand, political parties have been chosen via national elections, this process involves the people in the process of the political leaders that they have chosen.

To conclude, it is not yet possible to say that public participation is a goal of the PVE for relaxation of corona measures in the Netherlands. The main purpose of PVE is evaluation, as can be seen by evaluating the two case studies, public participation does not have to be a goal of PVE. For this thesis the level of participation is mainly relevant to know if the aim of this PVE is to involve a significant number of participants to be able to apply automated data processing. In this consultation process of this PVE there has not been set an aim on the percentage of citizens that will participate, the target group exist of almost 18 million Dutch citizens. It is likely that the number of participants will be high enough to be able to apply NLP in the analysis. An overview of this participation process and the role of PVE has been visualised in Figure 14.
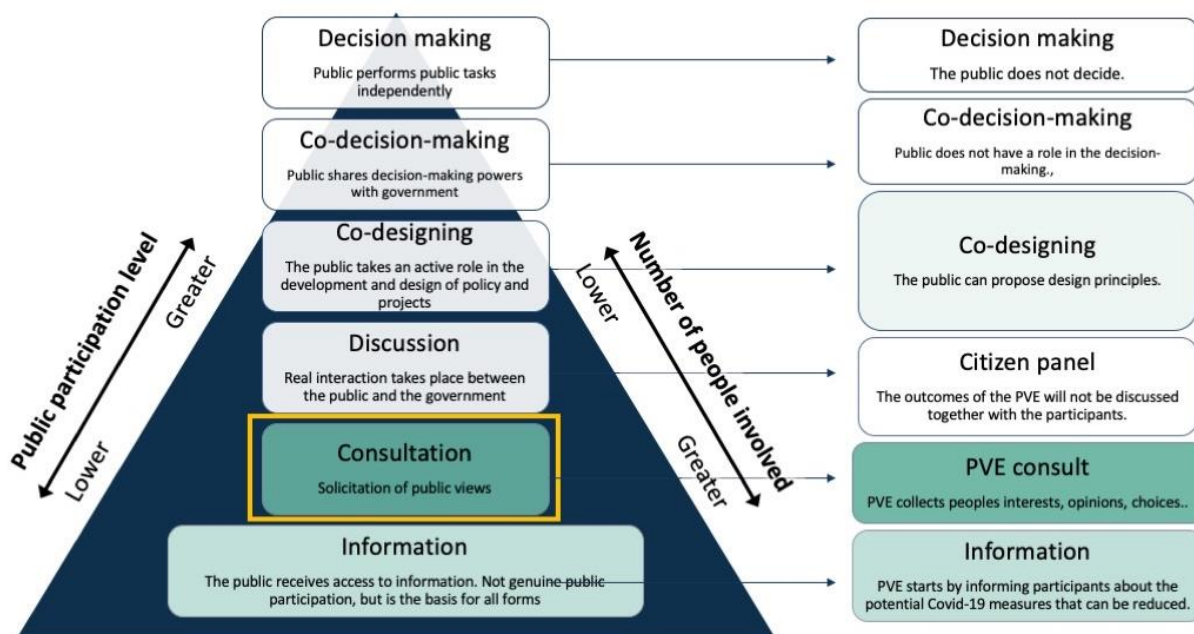
Figure 14 - Overview of participation process for the PVE 'relaxation of corona measures in the Netherlands' based on Mostert (2003)

Using natural language processing for these amounts of data needs less supervision compared to Súdwest Fryslân. The reason is that it is more likely that topics occur more often. If only 1356 people participate, it might be that topics occur less than 20 times, this does not make it representative. Unsupervised or semi-supervised learning can be used to define the size of different topics or values. It can also give insight in conflicting values about which a strategy has to be chosen, an example is safety versus autonomy. New technologies are currently being applied by PhD candidates of computer science to discover more efficient and precise methods to determine the values. This means less human annotation with more precise results.

4. Values in qualitative data of Participatory Value Evaluation

Motivations within PVE do contain a lot of information. The question that is being asked to participants is: 'please motivate your selection'. Based on the case studies, these motivations can be divided in arguments, prerequisites and design ideas. Design ideas which will be defined as design principles. This chapter will summarize the type of information that qualitative data can contain within PVE. It will also motivate why manual or automated annotation based on human values can be of added value for methodological purpose. The goal is to prove the types of information that can occur within qualitative data. The usefulness of this information will be discussed in chapter 8 and 9.

Types of information within motivations

When analysing previous PVEs, it can be seen that qualitative motivations of the selection of policy options includes different types of data. People can give arguments, insights, complaints or they could even just it to relief their heart and say something philosophical and off topic. To get to know which information is expected to be relevant for automation and analysis, previous reports of PVEs will be checked to analyse their use of qualitative data.

In the advice to the policy makers, the quotes of participants can be used as an illustrative statement (Mouter et al., 2019). It can also be used as a quote to represent an argument in favour of or against a policy (Mouter, Spruit, Itten, Shortall, et al., 2020). In the PVE of the heat transition in Utrecht, the arguments have been shown in the report to give an overview positive and negative arguments. Afterwards quotes have been presented to validate the arguments based on the motivations of the participants. The PVE commissioned by the Transport Authority Amsterdam does mainly include quotes as an illustrative statement to verify and validate the quantitative results (Mouter et al., 2019). Both PVEs do present some quotes of design principles and prerequisites.

First, some random examples from both case studies have been selected to try and categorize the motivations based on how they have been used in previous reports, being an argument (+/-), prerequisite or design principle. It can sometimes occur that an argument does not contain any information at all. These arguments will be skipped in the analysis. Next to the annotation based on the demanded information in previous reports, it can also be possible to annotate based on values. The concept of value-based decision-making has been explained in the background of this study. The methodological reason for choosing values is that they can be compared between different PVEs, this is harder for context-specific arguments, prerequisites or design principles. Independent of the type of motivation, values can be extracted. Based on the interviews with the people involved in the Regional Energy Strategy of Súdwest Fryslân, it was valuable and maybe even essential; to focus on the values. Two quotes from my interview are listed below to verify this:

> *"You can discuss and negotiate about arguments, you can talk about values, but you cannot judge them."*

> *"You need to think about where you are in the process, if this is an early stage, it is essential to think about values that can be taken into account when designing your policies. Values will become prerequisites of the process as participants do not decide about a specific choice."*

To perform the research of automated data processing, a decision should be made what topics will be modelled. An example of the different types and values can be seen in Table 2. This can be done based on the type of information to find i.e. arguments. However, the focus in this research has been made on values. This is due to the relevance within the case studies and the fact that values can be compared across multiple contexts. This will be relevant for further as training topic modelling on its topics across different cases is likely to increase accuracy, this concept is called transfer learning (Ruder, Peters, Swayamdipta, & Wolf, 2019).

Table 2 - examples of given argument and underlying values

Case study	Motivation	Type	Values
Súdwest Fryslân	De gemeente is een publiek orgaan en zal redelijkerwijze beleid moeten formuleren over de energiehuishouding binnen de grenzen van de gemeente.	Argument (+)	Responsibility,

	Typisch een taak voor de gemeente zo'n groot project	Argument (+)	Responsibility, leadership
	Gemeente mag de inwoners stimuleren om energiezuinige maatregelen te nemen, door te subsidiëren, maar niet door te verplichten.	Prerequisite	Self-determination, support
	Als we het aan het individu overlaten wordt het een inefficiënte toestand. Veel kosten en twijfelachtige baten. De gemeente kan bijv. op waterstof via het bestaande gasnet inzetten, wat voor een gezin te kostbaar zou zijn. De infrastructuur is er grotendeels al!	Design principle	(cost)-efficiency, technology
	Belangrijk	No information	No value
Relaxation of corona measures in the Netherlands	Zodat economische schade wordt beperkt	Argument (+)	Economic stability,
	Indien er een 1,5m kan worden aangehouden moeten bedrijven weer open.	Prerequisite	Safety
	Net als de twee hierboven valt ook deze onder de versoepelingen met de minste druk op de gezondheidszorg. Tevens is het economisch gezien van grote waarde wanneer bedrijven weer open kunnen gaan, mits aan alle voorzorgsmaatregelen wordt voldaan (thuisblijven bij klachten, hygiënemaatregelen, 1,5 meter afstand, spatschermen bij de kassa, thuisblijven bij klachten, testen personeel e.d.)	Design principle	Economic stability, safety, responsibility
	Slecht idee	No information	No value

5. Exploring possibilities of automated data processing

This chapter will explore possibilities of automated data processing. Automated data processing is the processing in which part of the work is being replaced by computers. This is a wide definition that makes it possible to explore relatively simple solutions for data structuring as well as advanced Natural Language Processing solutions, which is within the field of Artificial Intelligence. An exploration has been done in this chapter to support in making a selection of methods that can be applied in the case studies. It also investigates the possible applications of the techniques.

Automatization of qualitative data analysis

The current procedure of analysing qualitative data of the PVE is mainly based on manual analysis. It is based on analysing data until the point of saturation at which no new information is being extracted from the data that is being analysed (Mouter, 2020). This is an iterative process in which different values and arguments are grouped together so that analysts and policymakers can draw their conclusions. However, automatization can take place and as PVE is being used up to 30.000 participants, automatization has to take place to analyse all data. (Wiedemann, 2016) gives two developments that show the need for automatization. First, the availability of digital text worth to investigate is growing rapidly. Second, the improvement of algorithmic information extraction approaches allows for further bridging the gap between qualitative and quantitative text analysis. An overview of the possibilities that have been discovered to automate data analysis can be seen in Table 3.

Table 3 - Overview of possibilities for automating qualitative data analysis

	Possible application	Pros	Cons
Excel	Excel is able to manually annotate data or import data that has been annotated automatically. Excel can filter, count and visualize the data.	Widely available software that most people are skilled in.	Excel in itself cannot automatically label arguments. Manual labelling faces cognitive biases.
CAQDAS	CAQDAS can manually structure and label qualitative data by using existing software. It can give quantitative results of the qualitative data.	CAQDAS makes it possible to cut the textual data in different parts to label them separately. Different types of labels and dimensions of labels can be applied to text, i.e. values, arguments, norms, design principles.	CAQDAS requires training to structurally annotate the data. Also, cognitive bias by manual analysis should be taken into account.
Total Correlation Explanation (Corex)	An unsupervised model to detect different factors or topics in textual data.	There is no cognitive bias in detecting the topics	Detected topics might not be the relevant information for the researcher
Anchored Correlation Explanation (Corex)	A semi supervised model to detect topics based on different anchors	The anchors can be chosen based on prior analysis. The method is a quick way to quantify known topics	Anchoring bias can occur
Word2Vec	This semi supervised method is trained based on annotations in the specific context. It can predict the non-human annotated qualitative data. It can compare this prediction to the human annotated data.	Word2Vec can be trained in a specific context.	It is essential to annotate a part of the dataset. This can be time intensive if the dataset is large and diverse.

Wordify	Wordify can try to find the words that are most positive or negative related to a specific label.	Wordify is an online platform that can be used by everyone that has knowledge of Excel. It quickly gives insight in the pros and cons within a topic.	Wordify can process data, but it cannot label the data itself. Pre-processing is needed. Wordify does not give insight in the occurrences of the negative or positive words. This should be done afterwards if requested.
FTT	FTT is being able to find textual data that has the least relation with the selected data.	This method can give insight in other's opinions. It can quickly help to understand the variety of opinions. The method itself does not need supervision.	The method itself does not give information about the number of occurrences of different argument. It does neither apply a label or topic to the textual data.

Excel

One of the goals of the qualitative data of PVE is to analyse the arguments that have been mentioned. Projects show that qualitative data can be used to validate the quantitative data. Also, qualitative data can provide new insights in why people chose a set of projects (Mouter et al., 2019). If CAQDAS is too advanced, Excel can be used to structure qualitative data (Ose, 2016).

Excel as a tool can easily structure qualitative data. This type of automation is a manual process itself. Annotating the data can be done in Excel to be able to filter and visualize the data. A major advantage is that Excel is widely being used and it does not need extra skills to be able to annotate or filter the data. Structuring qualitative data in Excel can be easy as the software is widely available. However, a well thought method or framework should be chosen to structure this data (Braman & Jones, 2002). Braman has proposed a structure that exist out of five steps.

1. Familiarization, to list key ideas and recurrent themes.
2. Identify the thematic framework
3. Index or label the text
4. Charting or visualizing the annotated data.
5. Map and interpret the data.

This process has been applied for the PVE in Súdwest Fryslân. The description of the process can be found in chapter 6.

By switching the annotation process from Word to Excel, a separate sheet can be made to count the occurrences of annotations for every list of data input. It can also show visualizations and the user can intuitively filter the information based on the labels. One example that shows how Excel has been used to annotate values can be seen in Table 4.

Table 4 - Overview of manual annotation process in Excel (Dutch)

De Gemeente neemt de leiding en ontzorgt	Waardes
Gemeente kan grootschaliger projecten opzetten, de gemiddelde burger heeft er te weinig verstand van, geen goede overtuiging, of geen middelen dit te realiseren	vertrouwen; daadkracht
Het lijkt mij goed dat de Gemeente in samenspraak met de inwoners het voortouw neemt. Duidelijk en democratisch proces. Het moet alleen niet zo zijn dat de Gemeente keuzes en maatregelen gaat opleggen (niet dwingen)	inspraak; samenwerking
Om de belangen van eenieder goed te beschermen, dient de gemeente haar coördinerende en controlerende rol te behouden. Via de gemeentelijke verkiezingen heeft de burger invloed op de koers van de gemeente.	eerlijkheid; leiding; inspraak

Excel can be used if the PVE requires all qualitative data to be processed by humans. This does not necessarily mean that the quality or accuracy of annotations will improve. It does make sure that most unique arguments, design

principles and prerequisites have been detected by this manual processing of the qualitative data. Also, Excel can be used to support and evaluate other tools, it is advised if the stakeholders are not proficient in Natural Language Processing and its corresponding software. To give insight in this data, Excel can be considered to act as an interactive database.

Computer assisted qualitative data analysis software (CAQDAS)

Dedicated programs known as CAQDAS (Computer Aided Qualitative Data Analysis Software) were initially developed to enhance the accessibility of qualitative data by overcoming the physical limitations of paper data records (Wickham & Woods, 2005). CAQDAS tools manage the traditional processes of manual coding and support retrieval of coded segments (Richards, 2002). This traditional process is characterized by assigning a label to different parts of textual data without using a program to structure the data. CAQDAS is a method to structure qualitative data and if often used in social sciences.

In the context of PVE it could be used to analyse the motivations of participants. Examples of CAQDAS applications are NVivo or Atlas.TI. These software methods have been created to structure textual data and to specifically label part of the text. If several arguments are being mentioned within a motivation, this method can subdivide the labels among the different arguments. This method is being able to process the quantitative data about the labelling, including filtering and the creation of tables and graphs that show an overview of the labels. Woods, Macklin, & Lewis (2016) have done intensive research into the impact of CAQDAS applications. They have compared many applications of CAQDAS, including research that compares this method to manual data analysis. CAQDAS is mainly used for larger amounts of text or for example in interviews. If text has to be split for labelling, or if text contains a lot of different types of information, this method is advised. In this research, there is no need for labelling different types of information within textual data.

CAQDAS can be used if the research requires human reflection on every annotation in a context that requires different types of labelling. This method will be costly due to the time-consuming human research that is needed. If a PVE has a low number of participants that generate a high amount of data, CAQDAS can be used to structure the annotations.

(Anchored) Correlation Explanation (CorEx)

Correlation Explanation (CorEx) will be distinguished in total CorEx and Anchored CorEx. Total CorEx van be perceived unsupervised, where anchored CorEx is directed by human input and thus semi-supervised (Gallagher, Reing, Kale, & Steeg, 2017).

Total Correlation Explanation

Correlation Explanation can be used for semi guided topic modelling. Generative models like Latent Dirichlet Allocation (LDA) can be used for topic modelling, but they are not directed by anchors or human input and does not need labelled data or prior knowledge (Ver Steeg & Galstyan, 2014). In generative models, assumptions and hyperparameters are specified, though this generative model also reduced its applicability for different contexts. The principle of total Correlation Explanation (CorEx) is to learn representations of data that "explain" as much dependence in the data as possible (Steeg, 2017). Correlation Explanation (CorEx) is an unsupervised model that can predict the topics within text and is in itself not influenced by anchors given by humans. CorEx tries to determine a number of topics within the qualitative data input. CorEx does not assume a particular data generating model, and instead searches for topics that are "maximally informative" about a set of documents (Gallagher et al., 2017). This relies on a mathematical method to determine the correlation between words in text, the mathematical description have been described by Gallagher et al. CorEx itself speeds up the process of topic modelling compared to generative models, though adding anchors can add information for this model.

Total CorEx can be used when analysing topics within i.e. news articles, motivations and reviews. Another example related to this research in social science is about the latent factor model of 'Big 5' personality factors (Steeg, 2017). Based on the data from an online survey with the big CorEx was able to perfectly reconstruct the Big 5 factors while standard methods failed (Ver Steeg & Galstyan, 2014, 2015). CorEx can also show the words that are most related to the detected topic. These words can be used to validate the conclusion about a specific correlation or topic. If these words do actually correspond to the detected topic, i.e. to a specific personality factor, then this can be used to validate the model.

Anchored Correlation Explanation

Anchored Correlation Explanation (CorEx) is a semi supervised model that can predict the labels of text based on the given anchors (Gallagher et al., 2017). By doing this, word-level domain knowledge can be flexibly incorporated within CorEx though anchor words. This requires minimal human intervention and can result in topic separability and topic representation. The anchors can be chosen to quantify specific topics within qualitative data or within motivations. Anchored CorEx will try to label the textual data based on the anchors. The method is looking for the words that are most relating to these anchors and tries to detect them in the textual data. It is possible to create extra unsupervised topics that includes the words that are most likely to form a topic together and do not result from the anchors. Based on this unsupervised information, the analyst can decide to add another anchor. This iterative process will decrease the anchoring bias as the unsupervised part of the process might lead to new information. Compared to total CorEx, to anchored outcome can be influences by cognitive biases, such as the anchoring bias. The anchoring bias is a principle that researchers should be aware of. The anchoring bias may reflect the optimal speed-accuracy trade-off when errors are benign. If accuracy is important and speed is not crucial, then people perform more adjustments and the anchoring bias decreases (Lieder, Griffiths, Quentin, & Goodman, 2018). So, when applying this method within PVE, one should be aware of the speed-accuracy trade-off.

Conclusion

(Anchored) Correlation Explanation can be used to explain correlations within the data and it can quantify this correlation. This might be useful when insight in the size of a topic or an argument is required. The correlated words found by CorEx can be used to explain the argument and it can be used to validate the quantification. If the words do not correspond the argument, CorEx itself does not give a sentiment about the mentioned topic. Also, CorEx itself is looking for correlations to detect factors or topics within the textual data. When CorEx is being used for values, it will need anchors to prevent the method to detect different types of arguments. Also, when drawing the conclusions of CorEx, the researcher should be aware of the cognitive bias, or anchoring bias. This bias does influence the outcome. An important limitation of anchored CorEx is the fact that is detects frequent occurring topics, not the unique topics that might be of high interest for the policymaker.

Word2vec

Word2Vec is a method of converting words to numbers and representing them in a multi-dimensional space. This model has been developed by Google and their initial paper has been cited over 17.000 times (Mikolov, Chen, Corrado, & Dean, 2013). It is one of the most popular technique to learn word embeddings using shallow neural network. This word embedding model predicts the current word based on the context, and predicts surrounding words given the current word. Unlike the generative models explained, Word2vec is a discriminative model. This indicates that it is analysis conditional relations. In this context this means that that Word2vec is able to make predictions on different topics or classify different motivations based on the given conditions. These conditions can be a pretrained training based on other data sources, or it can be trained based on an annotated part of the data, i.e. a manually annotated set of data that exist out of a certain percentage of the total amount of data. This discriminative model will analyse the textual data based on the conditions that the analyst has given. Word2vec can be used together with a generative model, i.e. Latent Dirichlet Allocation. It uses LDA to assign topic for each word, and then employs Word2Vec to learn word embeddings based on both words and their topics. An interesting application of to illustrate the use of Word2Vec is the analysis of the classification of twitter streams during epidemic outbreaks (Khatua, Khatua, & Cambria, 2019). They have used Word2Vec model that have been trained in a different context to analyse the twitter data based on the labels: health risks, prevention, symptoms, transmission and treatment. It resulted in an overview of arguments for all labels, but it could be used to predict the stage of the virus and prevent proliferation. To conclude, in the context of PVE, Word2Vec could be used to label the data based on a pretrained set of data. It is essential to train Word2Vec in the right context or to try different pretrained types of Word2Vec, i.e. trained based on Google News or PubMed data. Unfortunately, this data is not available in Dutch. It could be trained based on Dutch data or the qualitative data can automatically be translated to be used as the input of an English Word2Vec model.

Some challenges for Word2Vec have been mentioned by Bakarov (2018):
1. Semantics can be a vague idea. Which relations do words have?
2. There is often a lack of proper training data, especially in languages other than Dutch.
3. It is hard to deal with ambiguity, the vector of a word is averaged based on the different meanings of a word, which influences the outcome.

4. Word2Vec can't handle out-of-vocabulary words, it is using a random vector if the word is not known.

Testing Word2Vec in this research has only been done based on already annotated data. Having annotated data makes it possible to compare the label prediction of Word2Vec with the labelling of the data source. Word2Vec uses in this context a specific percentage of the annotated data to train itself and analyse the word embeddings. Based on this training, Word2Vec makes it possible to predict the data that has not been used for training. This way of using this method makes it possible to benchmark its own results. The predicted labels can be compared with the already annotated data that had been used as an input. This reliability will be presented based on two indicators: accuracy and F1 score. The accuracy focusses on the percentage of correct predicted annotations. The F1 score is a bit more difficult and is based on precision and recall.

$$F1 = 2 * \frac{precision * recall}{precision + recall}$$

To conclude, Word2Vec can be used if part of the data has been manually annotated. This annotated data will be used to further train the word embedding of Word2Vec next to the pretrained model. It is essential to validate the applicability of the pretrained model in the specific context. There is never a perfect pretrained model, evaluating based on human values makes the process of selecting this model even harder. For this reason, different pretrained models can be compared. At this moment the Dutch word embedding models are not yet applicable for specific policy contexts that are being trained based on human values. Word2Vec gives feedback on the human annotation and it can validate the accuracy and precision of its model.

Wordify

Wordify is an online tool developed by the University of Bocconi in Milan, Italy. Wordify is an online tool that allows researchers to find the most descriptive terms for the object of their studies in text. It uses a variant of Stability Selection (Meinshausen & Bühlmann, 2010) to fit hundreds of logistic regression models on random subsets of the data, using different L1 penalties to drive as many of the term coefficients to 0. Any terms that receive a non-zero coefficient in at least 30% of all model runs can be seen as stable indicators. Wordify is using a trained dictionary based on BERTje (de Vries et al., 2019).

The output of Wordify gives an overview of the most positive and negative words for each topic. The output of this method for the case of Súdwest Fryslân can be found in Appendix F. This tool can be interesting to use when validating insights of other methods. An example is the negative sentiment about the landscape within the motivations for the option 'large scale energy will occur on a few places'. This validates the concerns about this policy option. The process takes just a few minutes and is a contribution to already labelled data as it familiarized the analyst with the positive and negative words within labels. This label can be based on values or it can just be the project for which a motivation has been given.

Farthest First Traversal (FFT)

Farthest First Traversal is not an automation technique itself. It is a practical mathematical method that can be used to find motivations in the vector space of an existing NLP technique. It will look for the farthest first event in the vector space (Shuyang, Heittola, & Virtanen, 2018). Shuyang et al. have used this method in their annotation procedure for sound event classification. Although, sound event classification differs a lot from linguistic interpretation, the example can perfectly be used to show the method of FTT. Shuyang et al. have done test in which they did train their algorithm based on partial human annotation. They have used different methods to find out the accuracy for a different number of annotations. It can be concluded that farthest first search increases the efficiency of annotation. Their proposed method can be tried to save labelling
budget for other classification tasks.
The reason why FTT is mentioned within the list of automation tasks is not to increase efficiency of other tasks. It is true that annotation based on FTT decreases the number of annotations to train the data.(Basu, Bilenko, & Mooney, 2004; Shuyang et al., 2018). Though, another application is interesting on its own. The first step for a data analyst is to familiarize yourself with the data (Braman & Jones, 2002). One example of finding a new argument based on FTT can be seen in Table 5. This example on its own cannot be used to draw conclusions, though, it shows the diversity of arguments.

Table 5 - Overview of contrasting motivations based on Farthest First Traversal

	Argument 1		Argument 2
English	More space between tables and placing order via mobile app or paper form submission and pick up order via a window.	→ **Farthest First Traversal**	I'm keeping my 15-year-old son at home right now! He'd love to go after a friend again
Dutch	Meer ruimte tussen tafels en bestelling plaatsen via een mobiele app of papieren formulier indienen en ophalen via een loket		Ik houd mij zoon van 15 thuis op dit moment! Hij zou graag weer eens een keertje na een vriend gaan!
Motivated policy option	Catering industry and entertainment re-open		Young people do not need to maintain 1.5-meter distance among each other.

FTT can be used to be able to familiarize based on a wide range of motivations. Motivation that occur quite often will be filtered as they are close the motivation that the analyst has already seen. This might lead to the discovery of unique arguments.

Uncertainties specific for Natural Language Processing

Uncertainties within manual qualitative data analysis have been analysed. The same will be done for automated data processing to be aware of the certainty of outcomes. Knowing the uncertainties will already help to reduce them during the process and will help in the interpretation of the results. Shah, Schwartz, & Hovy (2020) have done research into the predictive biases that influences the outcomes and uncertainties of Natural Language Processing. This will be explained based on the process overview of Figure 15. Shah, Schwartz, & Hovy (2020) have based this framework on an extensive survey of the relevant NLP literature. Four distinct sources of bias have been identified: selection bias, label bias, model overamplification, and semantic bias.
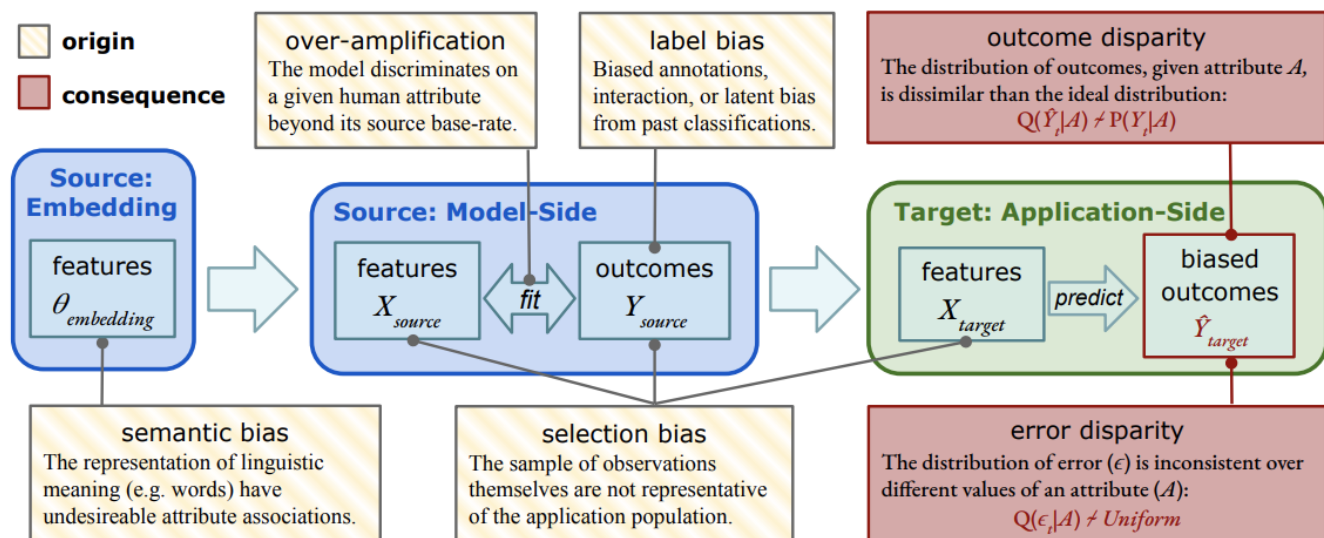
Figure 15 - Predictive biases within Natural Language processing (Shah et al., 2020).

1. Selection bias

Selection bias emerges due to non-representative observations. This can be due to two reasons. The first reason is that the total space of observations is not representative. This is the same kind of bias that could occur in manual data analysis, the outcomes within PVE can be biased if the group of participants is not a realistic representation of the public. The second reason can be that the NLP method uses a non-representative to i.e. train the algorithm. One of the reasons can be that a certain percentage of the participants are being used to predict the resulting percentage of participants. If this selection is not a good representation of the full group of participants, then this will create

selection bias that decreases legitimacy of the outcomes. This last reason for selection bias will decrease if the group of participants increases.

2. Label bias
Label bias occurs if the set of labelled values differs from the ideal distribution. This can be the case when the group of annotators is not representative or when annotators lack in knowledge (Joseph, Friedland, Hobbs, Lazer, & Tsur, 2017). Interestingly this is a bias that occurs in manual analysis as well as automated analysis. Label bias occurs more often in discriminative models that are semi supervised. This means that human annotation is involved.

3. Model overamplification
This type of bias is often the result of label bias or selection bias, though it could also occur independent of the others. If a model relies on a small difference between human attributes with respect to the objective, this difference can be amplified to be much more pronounced in the predicted outcomes (Shah et al., 2020). So, the source training distribution is different from the human annotated distribution and is different from the ideal distribution.

4. Semantic bias
Semantics are about word embeddings and their interpretations and correlation in text. This type of bias occurs in the manual labelling that is being used to train an NLP model. This type of bias is even more challenging for NLP itself. Training the model should be done by using semantically related data. This can be hard to find, especially for languages rather than English. A Dutch word embedding model that predicts semantics is often less accurate compared to the English model.

Awareness of these biases is necessary to minimize the impact of bias. Therefore, it is important to thoroughly think about the process of labelling and the representation of the participants and experts. The participants have to be selected to be representative for the desired group of people. The experts on the other hand should be able to label constantly so the difference from the ideal distribution can be minimized. Also, semantic bias can be reduced by choosing an appropriate trained model for the right language and the right application. The parameters of the embedding model can also be adjusted to get closer to the ideal distribution (Shah et al., 2020).

6. Analysis of qualitative data of Participative Value Evaluation

One of the proposed solutions to manage the incorrect use of Cost Benefit Analysis is to provide high quality quantitative information with regard to the intangible effects (Mouter et al., 2015). Participatory Value Evaluation is offering a method to overcome this limitation by incorporating questions that ask for a qualitative answering. An example is the PVE experiment on public decisions for investments in flood risk protection in the Netherlands (Mouter, Koster, Dekker, & Borst, 2018). The combination of quantitative and qualitative data makes it possible to not only decide based on willingness to pay. The qualitative data can give a clear evidence about the acceptance of public sector policies, it will give insight in the motivations of people's choices (Dekker, Koster, & Mouter, 2019). Another example for the need of qualitative data is a PVE conducted in Utrecht (Mouter, Spruit, Itten, Shortall, et al., 2020), analysis of qualitative data gave insight in the mentioned arguments. It can be used for further research on why people have selected projects. Next to that it can be used to discover unique insights and policy recommendations from the 'public experts. This chapter will explain the manual and automated analysis of the two case studies that have been introduced before.

PVE Súdwest Fryslân

Manual analysis of qualitative data

To be able to detect the type of information and to be able to extract the right information for policymakers, it is needed to familiarize yourself with the data before automatically label the data (Braman & Jones, 2002). Some machine learning techniques, i.e. Correlation Explanation or unsupervised topic modelling, do not need the step of familiarization. It can be advised to familiarize if the analyst is not aware of the topics at stake. Doing this first can help to interpret the outcomes of the machine learning techniques. One of the ways to familiarize with the data is to just read through a selection of the data. Another way that will be explained in this section is to manually annotate the data to compare this with the automatized way of annotating. This has been done for the PVE in Súdwest Fryslân to be able to validate the set of topics. In this example the topics have been chosen to be values.

The data is structured in an Excel file based on the output of the PVE software. This include the quantitative data as well as the qualitative input of participants. This annotation process will be based on the qualitative data that includes the argumentation for a specific policy-option that the participants have given points. The question that participants have to answer is simply: "motivate your choice". One other textbox has been analysed containing the question "motivate your concerns about the selected policy options". The output Excel has been filtered based on the people that have filled in one or more of the motivation boxes.

Defining values based on the motivations

The method that has been used for the manual annotation of textual data in the case study of Súdwest Fryslân is closely related to the grounded theory approach (Heath & Cowley, 2004). This approach is based on a cycle of data gathering, data analysis and reflection (Charmaz, 2003). The data gathering gathers the data as well as the theory for the analysis to develop i.e. categories for annotation. The next step is data analysis which is a process of constant reflection. How does the textual data compare to the previous one, does it contain a new category or value? These reflections will be annotated for every part of the text, in this case the motivation of a participant. Finally, the annotations should be compared and related to scientific literature.

To be able to define values within motivations of participants of a PVE, it is first needed to define this set of values. The grounded theory approach can be used to define these values. An overview of this approach can be seen in Figure 16, the steps are the following:

1. **Free value annotation**
   The basic theory of values has been discussed with the experts, but the decision has been made to have no restrictions for annotating keywords and values. Four annotators started by having complete freedom to define the values that they recognize in the motivations that participants have given. The annotators could also annotate keywords from the text that did not include values but are closely related to values. Results from qualitative analysis are susceptible for bias due to the coder's perspective. We minimized this potential for bias by involving three authors in the coded process. In this process it is important to stay close to the participants' own words (Jansen, Claassen, van Kamp, & Timmermans, 2019). That is, when participants talked about leader, we coded 'leader'; when participants talked about uncertainty, we coded 'uncertainty', etc.

   This has resulted in 589 different value related annotations of which 121 have occurred at least three times.

Figure 16 - Overview of the manual annotation process for the PVE in Súdwest Fryslân

2. **Selection of main values**
   These 121 values have been grouped together in different main values. This is done based on different value frameworks of i.e. De Argumentenfabriek (2012), Schwartz (2012) and Mouter, De Geest, & Doorn (2018). Until the point of satisfaction 17 main values have been defined. This list of values can be found in table Table 6.

3. **Annotation based on a value taxonomy**
   An iterative process for annotating values starts after defining the main values. After one batch of hundred annotations evaluation takes place to discuss if the main values still satisfy the annotation. During this process, one extra value has been added to the value taxonomy as this value was missing during annotation. If no essential new values have been discovered, the value taxonomy has been created and should reflect the context of the qualitative data.

This iterative process has resulted in a context specific value taxonomy that can be seen in Table 6. In the background difference between context specific values and generic values have been discussed. Interviews have been done afterwards to ask policymakers how they perceive these differences, these results will be presented in chapter 7.

Table 6- Overview of values to be used for annotation and their description

Value	Description
(Anti) market segment	Used both positively and negatively. On the one hand, market forces are positive, but there is also strong commentary on market forces and their negative effects. This is not really a value, but it is common so therefore added as a separate category.
Ambition	Striving and being driven to be the most innovative, sustainable or largest solution, for example.
Cooperation	Working together on a goal. Residents can work together, but also groups and organizations.
Cost-effectiveness	Money must be well spent, and the project must be profitable. No waste. Costs should not be too high.
Equality	All people are equal and have equal rights. Renewable energy must be accessible to all.

Honesty (of distribution)	Lusts and burdens are distributed fairly among inhabitants. Possible returns are distributed fairly.
Knowledge	Availability and sharing knowledge about renewable energy.
Landscape	Nature and environment are important. Horizon pollution is often seen as negative. Preserving the Frisian landscape is central.
Leadership	Clarity and control over the sustainability of the energy system. Often it is about an organization (the municipality) that has to take charge.
Liveability	Liveability of the region is important. Often used in a negative sense: The quality of the living environment should not be affected by (nuisance of) projects carried out.
Local	Participants value local projects. This is about generating local energy, implementation as well as project locations can be local and take place in their own environment.
Necessity	Certain (technical) steps are seen as necessary for sustainable energy generation.
Participation	The opportunity for residents and stakeholders to give their opinion and to think and do along.
Responsibility	Sustainable energy generation as a task or obligation. Both organisations and individual inhabitants are expected to contribute to this (and even have a moral responsibility to make it more sustainable).
Self-determination	The opportunity for residents to make their own decision on renewable energy and to be able to implement it.
Support	Unburden, realize that you can/will be helped in the right way.
Technology	Innovation is important. Technological solutions and innovations are being dedicated to achieving the objectives.
Trust	The importance that organizations (governments, companies) are honest and you can count on them.

The grounded theory approach states that one case study is being used to develop the theory, which can be validated for another case study afterwards. The other case studies might contribute to the theory that has been designed to annotate the values. In this context a value taxonomy has been created. This has been done based on the context of the energy strategy in Súdwest Fryslân. The decision has been made not to use only a set of generic values, the reason for this context specific set of values is validated via interviews and described in chapter 7.

The grounded theory approach has led to a set of context specific values. To reach the final set of values, three iterations of this process have been made, which included 300 motivations to annotate. By annotating the values based on the value taxonomy of Table 6, the unlabelled motivations can be annotated. This takes on average 5 hours per 100 participations in the case of seven open questions. This can also vary based on the amount of project that a participant can select. In the case of Súdwest Fryslân, participants were able to select and motivate all six policy options and their one textbox for their concerns. value overview can be made to visualize the number of value occurrences. An example of such an overview can be found in Figure 17. This overview has been chosen as it has the highest number of votes from citizens and is expected to have a higher influence on the decision-making process. The total annotation time for 1356 participants by one annotator takes around 65 hours. The results after 700 motivations were quite similar. However, these results do not contain all motivations and participants expect them to be read or at least to process this information to be used.
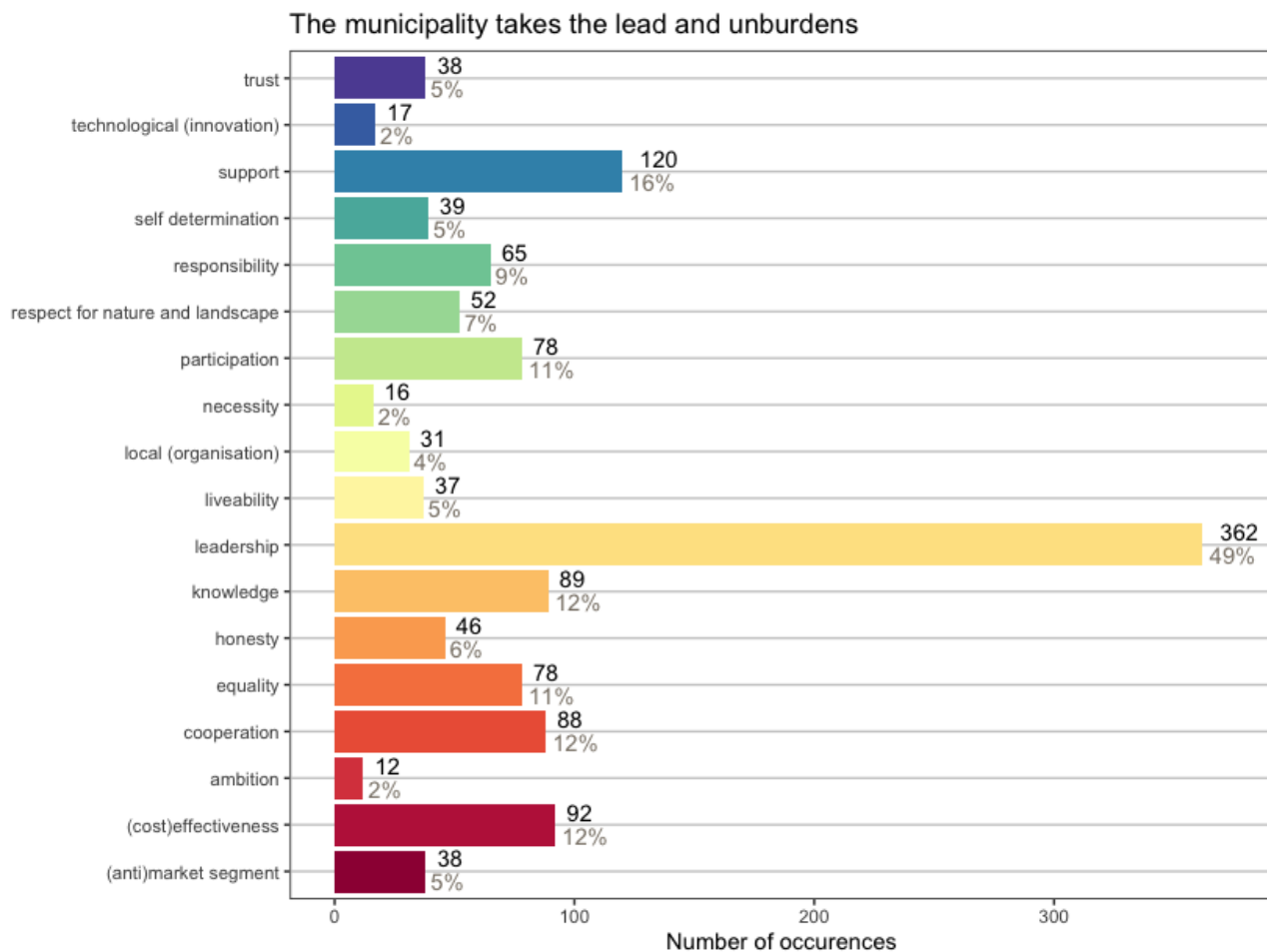
The municipality takes the lead and unburdens

Category	Number	Percentage
trust	38	5%
technological (innovation)	17	2%
support	120	16%
self determination	39	5%
responsibility	65	9%
respect for nature and landscape	52	7%
participation	78	11%
necessity	16	2%
local (organisation)	31	4%
liveability	37	5%
leadership	362	49%
knowledge	89	12%
honesty	46	6%
equality	78	11%
cooperation	88	12%
ambition	12	2%
(cost)effectiveness	92	12%
(anti)market segment	38	5%

Figure 17 - Overview of policy option for Súdwest Fryslân: The municipality takes the lead and unburdens

Uncertainties in manual analysis and intercoder reliability

The manual annotation process in Súdwest Fryslân has been done by two student assistants, by Shannon Spruit and by Lionel Kaptein. One of the questions often asked about automation of data analysis is about reliability. Can an algorithm predict better results than human annotators? To be able to give an answer on this question, the value extraction analysis has been done manually first. Manual value annotation faces some types of uncertainties.

1. Level of available knowledge. One out of three annotators has a background in social sciences. To validate the procedure for value annotation and the defined value taxonomy, an expert Nienke Van Schie has been asked to give her expert judgement on the value taxonomy. The other three annotators have done some research in types of values, though, their level of available knowledge about values might be lower.
2. Semantics can be confusing, during the discussion all annotators perceived a motivation differently. As values are not always explicitly mentioned, annotators can have a different perception in annotation. This type of uncertainty can be the result of cognitive bias. Our brain might not have any doubts about a certain value perceived in that specific context. Still, another annotator might have a different cognitive bias. This is one of the reasons to have multiple annotators.
3. There is a difference in number of annotations per motivation. Some annotators will annotate a value if they are i.e. 70% sure that the value is hidden in the context of the motivation. Another annotator might prescribe the value if he is 90% sure based on the specific words that are related to the value. This uncertainty has been tried to minimize by annotating in an iterative process in which some difficult motivations can be discussed together.
4. Biased by your own personal values. According to Hanel et al. (2018) people are more likely to interpret according to their own personal values. Their results show that people systematically misperceive comparison groups' values. This has not been validated in the context of this case study, but it is tried to be minimized due to the number of annotators.
5. Values in general are hard to assess and hard to quantify. They are hard to be assessed by the public as the 'magnitude' of a human value is hard to describe.

In this research there has been explicitly chosen to choose a diverse team of annotators. The goal of this PVE is to deliver understandable results to the citizen panel and the municipality. The citizen panel will create a report that includes an advice which shows some (contrasting) values that the decisionmakers should take into account. For this reason, it might not be a problem that three out of four annotators are normal citizens as well. If these persons perceive a specific value in a motivation, it will certainly be in there, even if it is just a spark. This can make the process more understandable for the citizen panel and the municipality as well. And if some motivations are unclear, it will give interesting insights to discuss them with the citizen panel. They might perceive even different values, but the process of discussing about different values increases the understanding of the context and increases the quality of the advice.

Some statistical conclusions can be drawn based on the manual annotation. This report will not give a conclusion on if you should aim for the highest rate of annotators' agreement. However, it is essential to find out about this agreement to be able to compare this with automated data analysis. To be able to say something about intercoder reliability, the Kappa Cohen score is being used. This function computes Cohen's kappa (Cohen, 1960), a score that expresses the level of agreement between two annotators on a classification problem. It is defined as

$$\kappa = \frac{p_0 - p_e}{1 - p_e}$$

where p_0 is the empirical probability of agreement on the label assigned to any sample (the observed agreement ratio), and p_e) is the expected agreement when both annotators assign labels randomly. is estimated using a per-annotator empirical prior over the class labels (Artstein & Poesio, 2008). Table 7 gives an overview of how to perceive this Kappa score (Landis & Koch, 1977).

Table 7 - Kappa Statistic Strength of Agreement (Landis & Koch, 1977)

Kappa score	Strength of agreement
< 0.00	Poor
0.00-0.20	Slight
0.21-0.40	Fair
0.41-0.60	Moderate
0.61-0.80	Substantial
0.81-1.00	Almost Perfect

In this case study the Kappa score is on average 0.42, an overview can be seen in Appendix E. Looking at the six most mentioned values, this average is 0.46, but this does not confirm that more annotations significantly increase the overall agreement of annotation. These kappa scores can be perceived to be moderate according to the statistic strength of agreement by Landis & Koch. The top three of values upon which the annotators agree most are: knowledge (k=0,73), (anti) market (k=0,63) and leadership (k=0,58). We can state that these values were easier to recognize in the motivations as the annotators often agree on them. Both knowledge and (anti) market do have a substantial level of agreement. The reason for disagreement and a lower kappa score can be due to the uncertainties that have been defined in this chapter. Values can be vague, and values allow different perceptions based on the context where they are being used. After optimising automated data analysis by using NLP techniques, this kappa score can be compared to validate the algorithm.

Automated qualitative data analysis
This section will describe a few automatized methods that have been applied for the case of Súdwest Fryslân. The context and theoretical background of these methods have been described in chapter 4. These tests are basic tests to validate the usefulness of these methods, the methods can be optimized. This can be done by testing different regression models in the language processing, or by using different trained datasets that include the vectors of every word in the trained contexts.

Automated data structuring
Data structuring is the start for further automation of data analysis. This structuring itself can already be seen as an improvement to automate the data processing. This starts by creating the right output of the results from the PVE.

These outputs, if properly structured, can be processed afterwards to extract information. If the (automated) data analyses has been done, this output should also be integrated in the defined data structure. This creates a kind of iterative process that can add analyses to the data.

Structuring of annotations can help in visualization

One of the simple examples is to annotate in a way that it can be used for further data processing. Previous PVE's (i.e. infrastructure in Amsterdam and the heat transition in Utrecht (Mouter et al., 2019; Mouter, Spruit, Itten, Shortall, et al., 2020)) often used qualitative data for validation and as inspiration for the reader. Reading through quotes for different policy-options gives insight in the topics at stake. As mentioned before this structuring has been done in Microsoft Word and cannot be used as a data resource afterwards. The (human) annotation for Súdwest Fryslân has been done in Excel, this makes it possible to perform analysis based on the occurrences of different annotations. It also prepares the data to be used in other software, Python and R have been used for analyses. Figure 17 shows a result of such an overview that can be used in presenting and reporting. The structured data also allowed the researchers to share quotes for every value within every policy option.

Getting to know the relevant topics

Another simple examples that can be used in automatization is creating overviews of how often words are mentioned or topics discussed. This resulted from the demand of citizens in the citizen panel. They asked for the technologies that are being mentioned and which technologies are discussed most in the motivations from the PVE. The script of which a screenshot can be seen in Figure 18 has been made during a one of the meetings with the citizen panel. The reason for this was that often relevance of specific topics were being asked i.e. the lack of information in the PVE about biogas. It was easy to respond to this candidate that biogas is considered to be less relevant by participants as biogas related topics have only been mentioned five times. As can be seen in Figure 18 the topic related to wind energy has been mentioned far more often. The goal of the citizen panel is to represent 'the public', if the public perceives biogas to be less relevant, the citizen panel can shift their focus.

```
In [*]:   1  print("\nLoading...\n")
          2
          3
          4  %run 'processingdata.ipynb'


Loading...

What do you want to do?
        A. Show overview matrix of values
        B. Show data of specific value
        C. Show motivations of specific value for specific question
        D. How often do people mention:
        E.
Enter the letter of your choice
D
About which word do you want to know more information? You can choose abbreviations
wind
Number of occurences for a word that contains wind
```

index		0	1	2	3	4	5	6	7	8	sum
152	windmolens	21.0	18.0	8.0	88.0	9.0	14.0	55.0	34.0	5.0	252.0
427	wind	7.0	3.0	9.0	24.0	56.0	7.0	13.0	15.0	6.0	140.0
631	windmolen	4.0	5.0	3.0	24.0	1.0	1.0	8.0	4.0	3.0	53.0
1403	windpark	2.0	1.0	1.0	4.0	nan	nan	2.0	nan	nan	10.0
1583	windmolen-	1.0	nan	nan	nan	nan	1.0	1.0	nan	nan	3.0
1589	windmolenparken	1.0	nan	nan	15.0	nan	nan	9.0	5.0	nan	30.0
1603	wind-	1.0	1.0	2.0	4.0	4.0	nan	3.0	3.0	2.0	20.0
1605	windpark	1.0	3.0	nan	5.0	1.0	1.0	6.0	3.0	1.0	21.0
2802	windparken	1.0	2.0	1.0	12.0	1.0	1.0	9.0	nan	1.0	28.0

Figure 18- Screenshot from Python of a method to give insight in the occurrences of technologies

Discovering correlations between annotations

Having structured data sources makes it possible to perform statistical analysis based on the quantification of the qualitative data. An example can be seen in figure 15, this overview includes the values for the policy option 'the municipality takes the lead and unburdens'. This visualization is an overview created in Python to show the occurrences of each annotation and how a specific label correlates to other labels. This overview includes only the

values that have been mentioned at least twenty times. The reason for this is the that uncertainty grows if the amount of annotations decreases. Increasing the size of the data will also increase the reliability of the results. An overview like this can be used to validate assumptions about relations between values. An example can be that 38% of the people that have given a motivation that includes the value support, also mentioned the value leadership. Other strong relations are leadership when mentioning knowledge (40%), leadership when mentioning fairness (42%), leadership when mentioning cooperation (38%) and support when mentioning equality (22%). The last one might validate or indicate that equality plays a role in the support by the municipality.

These overviews might be relevant to show stakeholders the correlations between different values visually. The larger the circle, the more participants have mentioned this value. The larger the connecting line, the higher the number of participants that have mentioned both values. Such an overview can be used to validate consumptions about i.e. the link between values or the uniqueness of a value in a specific context.
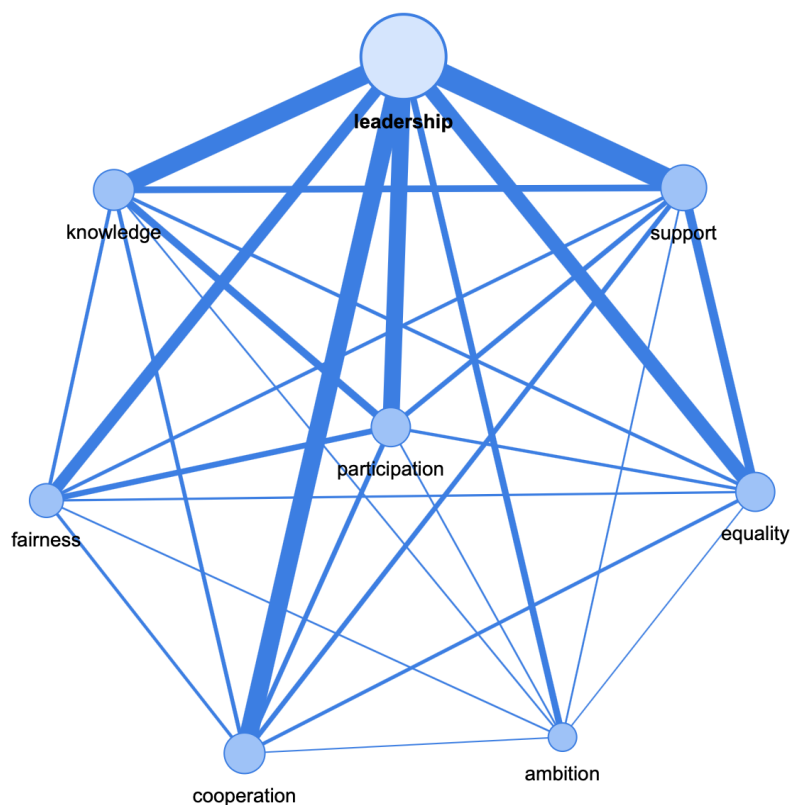
Figure 19 - Visualization of the values and their correlations with other values

Automated topic modelling
The methods of automated topic modelling have been defined in chapter 5. In the case of Súdwest Fryslân three methods have been tested. Unsupervised Correlation Explanation, the unsupervised method Wordify and a semi supervised word embedding model Word2Vec has been used. All three techniques will create another kind of result and can be used for different purposes.

Correlation explanation
Correlation has been used as a trial to find out if different arguments can be found automatically using the technique of Correlation Explanation. In the case of Súdwest Fryslân, CorEx has been used in two ways. First to determine if automatic topic detection can be used to familiarize yourselves with the data. Secondly, CorEx has been used to find out if is it can detect motivations for different values. In this context the dataset on which CorEx has been performed is filtered based on a specific value within the motivations of a specific policy-option.

Familiarize by using CorEx to determine topics
CorEx can be used unsupervised to discover the topics that are being discussed. This method creates a predefined number of topics based on the configuration. A small test has been done to try and find arguments in the full dataset for a specific policy option. This is in itself not related to values and the word embedding method might find other topics that are often discussed or words that are used together. For simplicity, an overview including the results can

be seen in Figure 20, this investigates why inhabitants want to do it their selves. This policy option has been chosen to prevent the bias based on Figure 17 - Overview of policy option for Súdwest Fryslân: The municipality takes the lead and unburdens. In this visualization the words are scaled based on their relationship with the topic. This method might be used for familiarization with the data, though it is an abstract process to define random topics and try to interpret them. It is essential to read the motivation or to have a significant background in the specific context to be able to interpret this data. In this case 'inhabitants do it their selves' has something to do with 'energy generation', 'collective', 'initiative', 'support base' and 'being involved in choices'. Still this concept can be quite vague, and the information of the policy option is required to be able to interpret the data.

Figure 20 - CorEx for Policy option: 'inhabitants do it their selves'.

Using CorEx for finding arguments of specific values
Another approach can be to try to use word embeddings to subdivide specific values. This needs an annotation process for these values first. This annotation can be done manually or automated. Based on the annotated dataset, CorEx can be applied on the qualitative data for one specific value. In Figure 21 an example has been given for the value leadership, that is the most mentioned value within this policy option. The values can be relevant for decisionmakers if the participants propose guidelines on how to use these values in the decision-making for the specific context. It seems that leadership is related to 'direction', 'facilitating', 'stimulating', 'making choices' and 'keeping track' on the process. This can help to validate if the most mentioned arguments for a specific value are part of the results and recommendations of the PVE. Often the report of a PVE includes quotes to explain the concept. If the concept is based on values, this method can be used to define the main arguments for each value, and it is able to quantify the occurrences of these arguments. Also, it can be possible to click trough the data and quotes that are within each topic. An important limitation is that this method will not find one-of-a-kind arguments as it filters data based on the frequency of words occurring in a specific context with other words.

Figure 21 - CorEx for value leadership in 'the municipality takes the lead and unburdens'

Sentiment analysis in Wordify

Wordify has been advised by Dirk Hovy and Pietro Lesci from the University of Bocconi, Italy. This method can find the sentiment of words that are being used within a topic. In this analysis, we could have used Wordify to find positive and negative terms related to a value mentioned by participants. For now, it has only been used to gain insight in the positive and negative words from the motivations for the six policy options. The results can be found in Appendix F. This can be useful, for example if the stakeholder that is related to a policy options wants to know how people perceive them. A few examples of positive and negative words can be seen in Table 8. Interesting insight is that participants voting for the municipality are positive about leadership, direction and the government. This does validate assumptions that have been made based on manual analysis. Creating the Wordify output takes less than 10 minutes and briefly shows the concerns and enthusiasm about policy options. The more qualitative data input, the better this sentiment analysis works.

In this example sentiment analysis has been applied based on the labels of different projects that participants have motivated. Another option could have been us the labels of values. This will quickly give insight in the positive and negative relations according to a specific value. Policymakers might use that information to incorporate the positive or negative topics in their strategy.

Table 8 - positive and negative words defined by Wordify

Policy Option	Most positive words	Coefficient	Most negative words	Coefficient
The municipality takes the lead	gemeente	0,558	opslag	0,396
	leiding	0,438	plekken	0,396
	regie	0,414	eigen	0,394
	overheid	0,396	markt	0,394
The inhabitants do it their selves	inwoners	0,514	gemeente	0,396
	eigen	0,486	landschap	0,396
	draagvlak	0,422	markt	0,396
	zonnepanelen	0,42	opslag	0,396
	initiatief	0,404	plekken	0,396
The market decides	markt	0,558	opslag	0,396
	marktwerking	0,4	energie	0,392
	bedrijven	0,394	inwoners	0,392

	marktpartijen	0,384	landschap	0,392
	komen	0,336	zonnepanelen	0,35

Semi supervised topic modelling Word2Vec

Word embedding has been tested in the case of Súdwest Fryslân to find if NLP can predict the labels for unlabelled motivations. Word2Vec can be programmed to choose the percentage of trained data and the percentage of data to predict. If Word2Vec is being used for a fully annotated dataset, the method is able to benchmark its own prediction with the annotated dataset. The risk of this benchmarking is that the annotated data includes uncertainties as well, this is mentioned in the section 'Uncertainties in manual analysis and intercoder reliability'. Also, this word embedding method has been trained by twitter data. The twitter data is likely not to be trained on human values. This will decrease the quality of the correlation vectors that have been created. The data can be trained based on human values, though this will need many PVEs that are related to values as an input.

It is still interesting to try this method to compare the accuracy or F1-score with the reliability of manual annotated data. If the manual analysis is not fully reliable, but if it helps to draw conclusions or to familiarize with the data, then the automation might not need to be reliable as well. This depends on the output and the type of information that is being used by decision-makers. An example of the accuracy and F1-score of Word2Vec can be seen in Figure 22 and Figure 23. The training size does influence the accuracy and F1_score. The percentage of train size has been set in the range of 10% - 80% in which 80% indicates that 80% of the manually annotated data has been used to predict the 20%. There are some limitations of this method in the context of Súdwest Fryslân.

1. Is the train size is relatively low, it can even occur that a value has not been annotated in the trained data. If the value has not been annotated enough, the method won't be able to analyse the word embeddings. Every value can be used in different arguments. If the value occurs just a few times in the train data, it will cannot be used accurately to predict the values in the test data.
2. The training dataset is based on the manual value annotation. This indicates that all uncertainties for the manual value annotation influence this automated process as well.
3. The technique of word embeddings analysis text that is mostly split into pieces. The best results will be gained if every part of the text contains one label. In that way, the label having the highest correlation can be chosen. In this research, the motivations of participants often include several labels. In that way the algorithm will apply a label if the correlation reaches a certain threshold. There will always be some kind of relations between the text and the topic. It can be advised to split the text in different parts to apply more specific labels.

First, this process has been done for all 18 values that have been annotated in the motivations of the PVE of Súdwest Fryslân. An overview of the accuracy and F1_score can be seen in Figure 22. Between 11-19% of the annotations have been predicted correct, this accuracy is a lot lower than the reliability of human annotations, so this contains quite some uncertainty. Further research is needed to improve this accuracy, for which recommendations will be done. One of the reasons is the low occurrence of some values. The algorithm has to decide for 18 values if they can be matched to the motivations. Especially values that often occur together with other values are hard to predict, i.e. support and equality or leadership and support. This also decreases the F1_score, which is around 0.3. A value of 1 would be perfect, 0.3 indicates that precision or recall is not yet accurate. It can be stated that this method should be improved to be used for this number of different labels and with this number of participants. The method will be trained when having more participants and motivations.
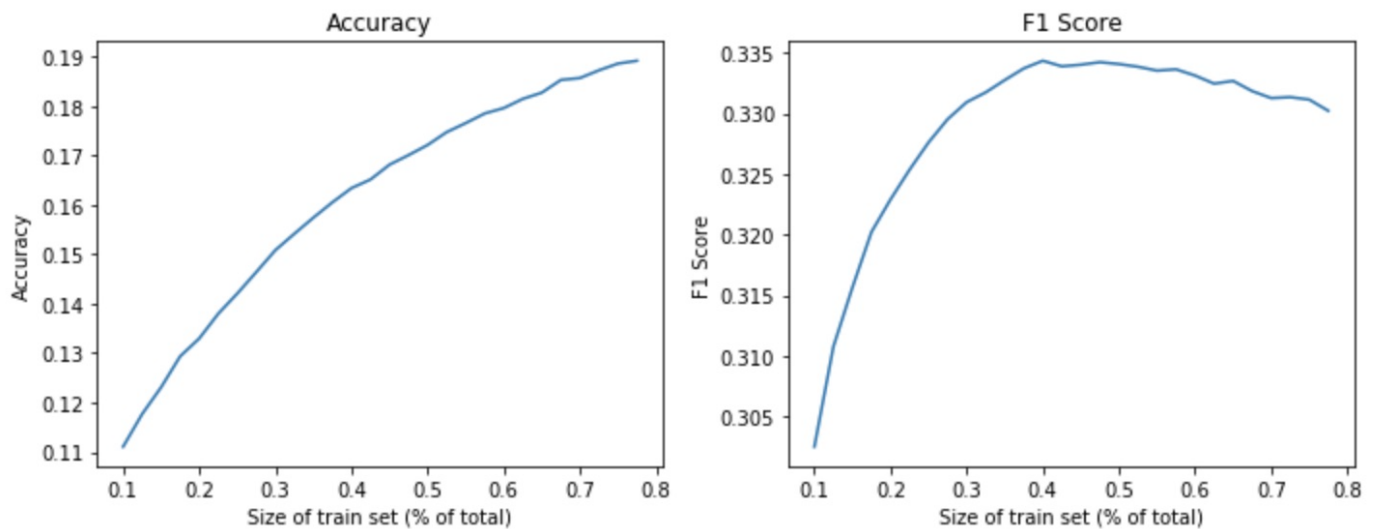
Figure 22 - Overview of accuracy and f1 score for Word2Vec value prediction (all values)

To minimize the limitations of the first analysis, other analyses can be done based on the six most mentioned values. This can give an indication if reliability can be increased when more trained data is available. If i.e. 10% of the train set is being used to train 90% of the data, then every value should be mentioned at least 20 times in the trained subset. An overview for Word2Vec of the six most mentioned values can be found in Figure 23. As you can see the accuracy and F1 score has been increased. Still only 1/3 of the motivations have a correct annotation. One way to overcome this limitation is to train the data based on the context, which will be explained in the next section. The F1 score of around 0.55 is actually pretty good. It is also interesting to see that is does not increase that much based on a higher train set. One of the reasons can be that the keywords for specific values have a high overlap for different motivations.
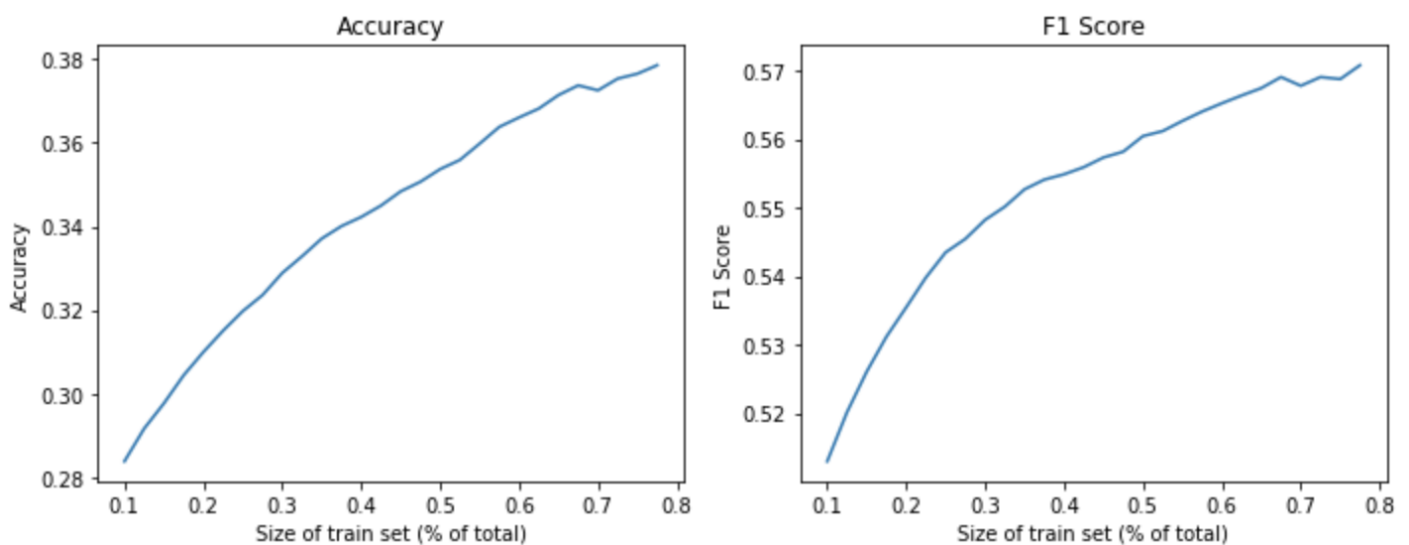
Figure 23 - Overview of accuracy and f1 score for Word2Vec value prediction (six values)

Farthest First Traversal(FFT) Mode to create a value taxonomy

FTT can be used to try to make the value taxonomy more efficient. This method has been embedded in a web application that has been created by Enrico Liscio, he is working on his PhD in Computer Science at the Technical University of Delft. At the time of writing, this method has only been developed to create a value taxonomy. The system proposed the motivation of a participant. It gives the Dutch motivation, the English motivation and the selected policy option. A screenshot such a random chosen motivation can be seen in Figure 24. This is the motivation that can be annotated. Using FTT will reduce the number of duplicate arguments and saves time in annotation. One explorative session of annotation takes one hour for every annotator.

Motivation (English)

If each energy generator has its own battery to feed the converter, we only need the grid at peaks in demand and yields.

Motivation (Dutch)

Als elke energieopwekker een eigen batterij heeft om de omvormer te voeden, hebben we het net alleen maar nodig bij pieken in de vraag en opbrengsten.

Choice

Betting on storage

Figure 24 - Screenshot of motivations shown based on FTT (by E. Liscio, 2020)

The annotation that is shown will be annotated based on the context the researcher asks for. The application for Súdwest Fryslân is to define a value taxonomy based on all the values that are incorporated in the qualitative data. The second difference compare to the fully manual annotation is the definition of keywords that are corresponding to a value. These keywords are the words that the researcher needs to make his decision about the occurrence of a value. These keywords can be used for semi-supervised topic modelling in the next stage to train the algorithm which words indicate a specific value. The keywords that relate to the value do not have to be in the specific motivation. An example is the keyword 'direction' when annotating the value 'leadership'. Another keyword that can be added manually is 'director' or 'control'.

leadership ✖

director ✖   leader ✖   control ✖   supervision ✖   protection ✖   wild grow ✖

direction ✖   conditions ✖   central point ✖   democratic ✖   necessary ✖
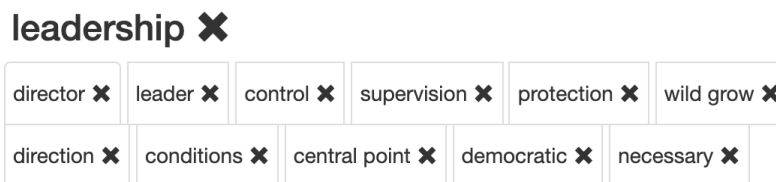
Figure 25 - Screenshot of the value and corresponding keywords

This method is still under construction, for now it can be interesting to show the list of values based on the manual annotation compared to the semi-supervised annotation by using FTT. This overview can be seen in Table 9. Note that this is the annotation by the writer of this thesis. This will be verified by other annotators. It is interesting that a one-hour annotation process already comes up with most of the values mentioned. 15/18 motivations have been found using this method. Also, some new values have been defined and might be added if the other annotators also identify the relevance of these values. In the previous manual process, this took four up to eight hours. During the manual analysis, the data was already being annotated, but only partly on the final value taxonomy, so most of this process had to be redone. This shows that FFT is a promising way to define the value taxonomy.

Table 9 - Overview of values by manual annotation compared to semi-automated annotation by using FTT

MANUAL ANNOTATION	GUIDED ANNOTATION BY FTT
Responsibility	Responsibility
Leadership	Leadership
Knowledge	Knowledge
Equality	Equality
Cooperation	Togetherness
Support	Support
Trustworhtiness	Trust
(Cost) effectiveness	Efficiency
Ambition	
Self determination	Freedom
Landscape	Nature
Participation	Participation
Fairness	Local diversity
Technology	
Liveability	Liveability
Necessity	Necessity

Local (organisation)	Local benefit for citizens
	Certainty
	Initiative
	Innovation
	Economic welfare

Presentation of values to decisionmakers and usefulness for decisionmakers

The theoretical explanation of the values in decision-making have been given. One of the used case studies has been done in request of the municipality of Súdwest Fryslân. To be able to validate the usefulness of values in decision-making, this case study has been analysed. This is done based on interviews and observations in cooperation with the municipality. It is hard to exactly indicate how values contribute in the decision-making. All stakeholders are allowed to offer the municipality consultation documents and reports. Some of them have been written from a business perspective, most of them have been written on behalf on the municipality.

Relaxation of corona measures in the Netherlands

Manual analysis of qualitative data

The time frame from the start of the PVE about Covid-19 until the delivery of the report has only been two weeks. This is a challenge as the number of participants for PVE has never been this high, over 30.000 participants have filled in the PVE. Also, the time frame has never been this short from the start of the PVE until the moment of policy recommendation. This also resulted in a limited possibility of analysing the qualitative data manually. There has been chosen to focus the manual analysis on three topics when randomly reading through the data.

1. Defining the arguments for every option and copying the arguments to be used in the report.
2. Defining the prerequisites when the participant has chosen a policy option and copying the arguments to be used in the report.
3. Selection motivations to be used for quotes. This motivation can be selected if it is a good representation, if it is a unique and new argument, if it proposes an interesting design principle or just because it seems interesting.

This manual data analysis has been done by several researchers and students from the Technical University of Delft. All having their own expertise, not necessarily in the field of this research.

Automated qualitative data analysis

This section will describe a few automatized methods that have been applied for the case study of Covid-19. The context and theoretical background of these methods have been described in chapter 4. These tests are basic tests to validate the usefulness of these methods, the methods can be optimized in the future. This can be done by optimizing the parameters of these methods, or by using different trained datasets that include the vectors of every word in the trained contexts.

Sentiment analysis in Wordify

Wordify has been advised by Dirk Hovy and Pietro Lesci from the University of Bocconi, Italy. This method can find the sentiment of words that are being used within a topic. In this analysis, Wordify has been used to explore the positive and negative aspects of a policy option. The results can be found in Table 10. This can be useful, for example if the government wants to know about the issues that are at stake. Wordify does not give context of these words, this should be validated by manual analysis or it can be compared to other methods for automated data analysis that include sentiment. An example of positive and negative words can be seen in Table 10.

Table 10 - Example of Wordify analysis

Policy Option	Most positive words	Coefficient	Most negative words	Coefficient
1 Nursing and care homes allow visitors.	bezoek	0,74	bedrijven	0,536
	ouderen	0,706	economie	0,536
	eenzaamheid	0,634	economische	0,536
	leven	0,634	jongeren	0,536

			laatste	0,632		weer	0,536
			bewoners	0,536		elkaar	0,526
			eenzaam	0,536		horeca	0,502
			kwaliteit	0,536		minder	0,496
			leed	0,536		houden	0,49
			mensen	0,536		kapper	0,472
			onmenselijk	0,536		werken	0,472
			sterven	0,536		gaan	0,47
			verzorgingshuizen	0,536		beroepen	0,466
			verzorgingstehuizen	0,536		contactberoepen	0,464
			oudere	0,534		economisch	0,458
			bezoekers	0,53		blijven	0,456

Creating the Wordify output takes less than 10 minutes and briefly shows the concerns and enthusiasm about policy options. Looking at the example of Table 10 for the policy option in which nursing and care home allow visitors, it is interesting if the sentiment of the words gives an indication of the arguments. This is actually the case comparing results with the final report (Mouter, Spruit, Itten, Hernandez, et al., 2020). The five arguments that have been mentioned in this report are presented in Table 11. In the second column, the comparison has been made with the Wordify results. It can be seen that 4/5 arguments can somehow be related to one or more of the most positive words. It is not validated yet if this can contribute in the decision-making. However, it does validate that sentiment can be found within data, which might be of interesting for later use when combining this to topic modelling.

Table 11 - Comparing arguments with Wordify based on positive words

Argument in favour of allowing nursing homes to allow visitors (Dutch)	Corresponding positive words (Dutch)
Deze ouderen hebben bezoek en contact heel hard nodig	Bezoekers
Het risico op corona weegt niet op tegen de eenzaamheid	Eenzaamheid
Leed van eenzaam sterven	Sterven, eenzaam, laatste, leed
Geen bezoek is voor familie ook traumatisch	Onmenselijk,
Deze maatregel is veel prettiger voor zorgpersoneel	-

Using Anchored CorEx for validating the size of arguments

CorEx can be used unsupervised as has been done for the case study in Súdwest Fryslân. Another approach is the supervised method of Anchored Correlation Explanation. The anchors can be defined based on the topics that are either being found by using unsupervised CorEx or by manually chosen anchors. For the project of Covid-19, the choice has been made to find the topics based on unsupervised CorEx and have been performed in three steps:

1. The first step in this process is to try and recognize themes in the topics that have been created automatically. These topics can be adjusted by adding anchors to the topics, this anchor is a kind of magnet that the algorithm is using to define correlations for that topic. This can be done with or without familiarizing with the qualitative data. In this process the first step of CorEx has been performed without reading the qualitative data. Basic knowledge about the context of the PVE is required.
2. Secondly, the CorEx results have been shown to the experts that have annotated the qualitative data. They were asked to interpret the topics and to add anchors to these topics. In this process, the experts can supervise how CorEx is being performed.
3. Lastly, the final set of topics will be compared to the arguments in the final report (Mouter, Spruit, Itten, Hernandez, et al., 2020). This can give an indication of the number of arguments that have been found by automated data processing compared to manual data analysis. Is can also quantify the arguments based on the percentage of motivations that can be categorized in the specific topic by CorEx.

One example can be shown in Appendix G, this has been done for all policy options. By using the method of Anchored CorEx 60% of the arguments have been found and defined as a major argument. Also a few arguments could be added that were not defined yet. An example is the argument in which people are afraid that contact

professions will operate illegally. This argument could be determined by CorEx but was not yet determined by the human approach of selecting arguments.

Farthest First Traversal Mode to create a value taxonomy

The same approach as Súdwest Fryslân has been used to create a value taxonomy for the PVE about relaxation of corona measures in the Netherlands. This will be the start of semi-supervised topic modelling that can be used to categorize the qualitative data based on different topics. The topics will be the values to find out which values are considered to be relevant and should play a role in the decision-making process. Using FTT will reduce the number of duplicate arguments and saves time in annotation. Again, one explorative session of annotation takes one hour for every annotator. The annotation process has been programmed in a web application by Enrico Liscio and allows the user of the platform to define values based on the motivations that are being shown. Keywords can be added to these values to give guidance in the context of these values. One example of a value that differs from the values in Súdwest Fryslân can be seen in Figure 26.

psychological welfare ✖

| psychiatry ✖ | loneliness ✖ | walk around ✖ | visit ill people in nursery home ✖ |

| cheer up ✖ | last days ✖ | empathy ✖ | compassion ✖ | last phase of life ✖ |

| unite ✖ |

Figure 26 - Screenshot the value and corresponding keywords

The process of creating the value taxonomy based on this method is not finished yet. The values of one annotator can be given, this list of values will be adapted based on the values of other annotators. Also, some values might be combined in the same category, an example is psychological welfare and fun. It has to be decided if mentioning both values gives more information compare to the combined values. This deliberation about the value taxonomy will take place in the next step of this method. The values that the author of this thesis came up with are safety, economy, vitality, health, friendship, psychological welfare, freedom, responsibility, equality, intimacy, trustworthiness, fairness, fun. Six out of the fourteen values have been mentioned in the motivations of the case study of Súdwest Fryslân as well. This indicates that the relevant values are depending on the context of the PVE.

7. Context specific values versus generic values

Literature does not agree upon one value hierarchy (Graham et al., 1991; Schwartz, 2012). Based on the information of values in the background, a knowledge gap has been defined in how to create a value taxonomy. The approach that has been applied in this study has been discussed in chapter 6. This iterative process has included the background knowledge about values to define context specific values that have a stronger relation to the policy in the specific context. In the case of Súdwest Fryslân, the first annotation process without boundaries gave 589 unique keywords or short sentences related to values in that motivation. The theory of basic values by Schwartz mentions a universal set of ten values(Schwartz, 2012). This is already an indication that values can be diverse. If automated data processing is being used to extract values, it is essential to select which values to use. First the context specific value taxonomies will be compared to gain insight in the influence of the context on the values. Afterwards, the values of the case studies will be compared to general value theories. Lastly, the relevance of context specific values or generic values has been discussed via an interview with the stakeholders in the PVE of Súdwest Fryslân. This chapter will reflect upon the value taxonomies and it will try to bridge the gap by evaluating the sets of values together with the stakeholders in the process.

Values comparison of case studies

Table 12 shows an overview of the values that have been defined for the case studies of Súdwest Fryslân and Covid-19. The value taxonomy of the PVE of Covid-19 is still under construction and is waiting for consultation, but this is not necessarily relevant for the comparison of context specific values with generic value theories. As can be seen in Table 12, most values are not the same in both cases. The context of making a policy for an epidemy is different than making a policy for sustainable energy. Only six values have been annotated that overlap in both contexts. An advantage of having the same values in different PVEs is that value preferences can be compared in different contexts. It can be useful to know for the municipality that the value freedom is highly important in an economical context, but less important within the context of health. This is a non-validated example, but it illustrates that if many PVEs are being performed, it could be interesting to compare value preferences based on the context. This can only be done if the value taxonomies are overlapping or if they are even identical, by using generic values.

Table 12 - Comparing context specific value taxonomies of the PVE's of Súdwest Fryslân and Covid-19

SÚDWEST FRYSLÂN		COVID-19
Responsibility	=	Responsibility
Leadership		Safety
Knowledge		Economy
Equality	=	Equality
Cooperation		Fun
Support		Health
Trustworhtiness	=	Trustworthiness
(Cost) effectiveness		Friendship
Ambition		Psychological welfare
Self determination	=	Freedom
Landscape		Intimacy
Participation		
Fairness	=	Fairness
Technological innovation		
Liveability	=	Vitality
Necessity		
Local (organisation)		

No literature has been found that favours context specific values to be used in Natural Language Processing. Wilson (2019) has used context specific values and he compared them with the Schwartz theory of basic values (Schwartz, 2012). By doing this he did not find an appropriate way to categorize his values based on the value taxonomy of Schwartz. He used NLP to try and find the relationships between his values and the one of Schwartz. This resulted in a correlation table in which some of his values are correlated to several values of Schwartz and some have no relation at all. An example is empathy which is related to achievement, universalism and benevolence. The question remains if it is more valuable for decisionmakers to know these values of Schwartz or the context specific values.

To place the context specific values in a generic value theory, the theory of Haidt has been used (Graham et al., 1991). By categorizing these context specific values, it can be discovered if all generic categories have been quantified. This overview can be seen in Table 13. It seems that the values within Súdwest Fryslân are a diverse representation as they are spread out over the generic values of Haidt. For Covid-Exit the respect for authority seems not to be a relevant value. This is interesting in the context of decision-makers by the government, but the reason for this is not defined. Still, the question has not been answered what set of values is most valuable for decision-makers. For this reason, interviews have been conducted with stakeholders of the PVE in Súdwest Fryslân.

Table 13 - Categorize context specific values in the value theory of Haidt for Súdwest Fryslân (left), Covid-19 (right) (Graham et al., 1991)

Care and empathy	
Responsibility	Responsibility
Support	Trustworthiness
Trust	Safety
Misgiving (to the market)	Health
	Psychological welfare
	Intimacy
Loyalty and pride	
Collaboration	Economy
Technology and innovation	Friendship
Local (responsibility)	
Respect for the authority	
Leadership	
Participation	
Necessity	
Purity and beauty/fairness	
Respect for nature and landscape	Fairness
Liveability	Vitality
Fairness	Fun
	Psychological welfare
Justice and reciprocity	
Equality	Freedom
Ambition	Equality
Self-determination	
fairness of distribution	

Evaluation of value taxonomies by stakeholders in the process of Súdwest Fryslân

There is a conflict between the use of generic values and context specific values. Using generic values saves time as the value taxonomy is already known. Is also gives the possibility to compare different contexts or even region with each other. Contrastingly, context specific values can incorporate the needs of the specific context or the needs defined by the policymakers. It can be that policymakers are interested in how people perceive the impact on specific values like their liveability or the effect on their freedom. In Súdwest Fryslân Mienskip or 'local cooperation' and 'anti-market sentiment' seemed to be important. To get to know the reflection on this debate by participants and policymakers, it has been decided to conduct interviews and ask for an external opinion. As explained in the method section, this evaluation will only be done based on the PVE of Súdwest Fryslân. The process of consolidating and defining the value taxonomy for the PVE of Covid-19 is still in progress in the research of Enrico Liscio.

The question that has been asked to the interviewees was based on the theory of basic values by Schwartz (Schwartz, 2012) and compared to the context specific values determined by the approach described in chapter 6. Literature did not agree upon the best way to create a value taxonomy or which generic set of values could be used. For this reason, a semi guided interview had been conducted among eight people involved in the process of Súdwest Fryslân to reflect upon the selected set of values compared to generic values. One of the reasons of this interview was to find out which values can be seen as relevant to decision-making. Most of the participants have seen the results of PVE in

which context specific values have been used. In the interview the interviewee has been asked to take his time to look at two sets of values. One set of values defined in the case of Súdwest Fryslân. The other set of values was based on the theory of basic values by Schwartz. Figure 28 and Figure 27 have been shown to the interviewees. Afterwards the question was being asked: If we show a graph similar to the above, but with general values rather than context-specific values, how would that affect decision-making?
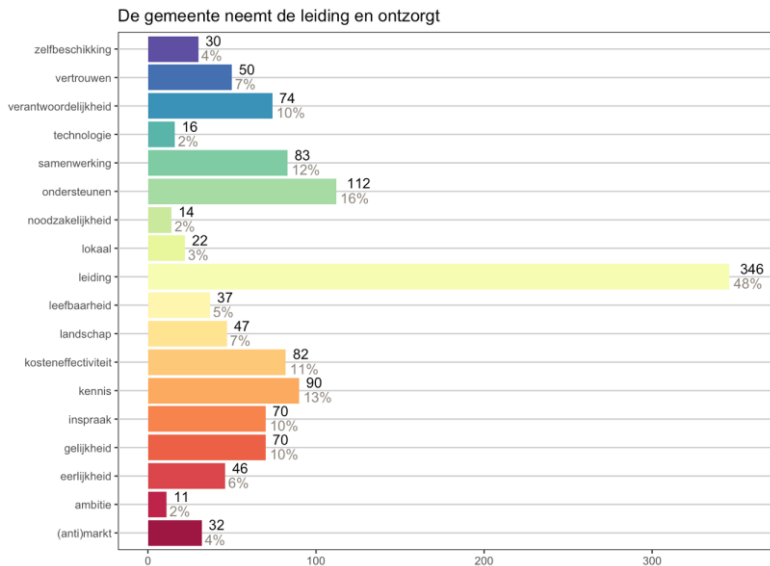

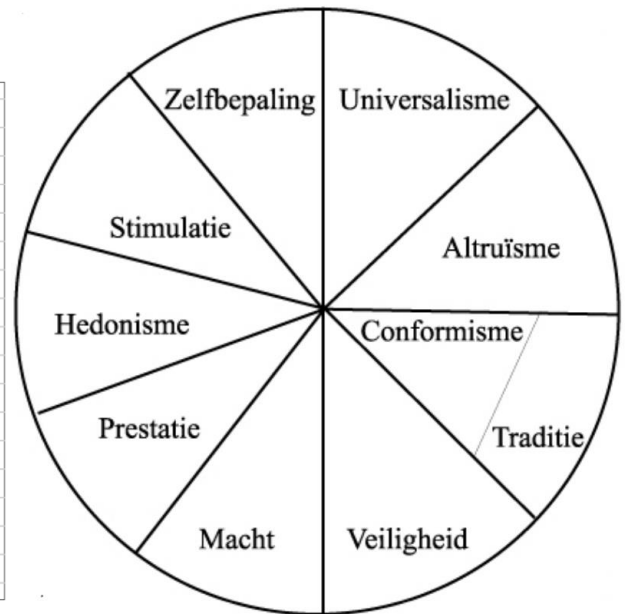Figure 28 - Figure from the interview showing an overview of the context specific values.

Figure 27 - Figure from the interview containing the basic values of Schwartz to compare with the context specific values in Súdwest Fryslân (Kuylen, 2007) .

The participants have shared their thoughts and preferences which can be seen in Figure 18. As can be seen all participants favour context specific values in the context the RES in Súdwest Fryslân. The main reason is that the results should be understandable by the participants themselves, they should be able to recognize their own values. Context specific values seem to be easier to translate into action. It is also mentioned to be important that familiarizing with the values is essential and that the list of values is never finished to give participants the feeling that they could also express values that are not in the list. Some of the interviewees have mentioned generic values to be an option based on some prerequisites. First, the decision-making should take place in the moral or behavioural domain to be able to use generic values. It could also be used if the topics is closely related to personal values. Secondly, the people involved in the processing of these values should be highly educated and understand value theories. It might not be wise to show these generic values to the participants. The people using the values in policymaking should be able to turn generic values into action.

Table 14 - Responses from the interview by participants on the question if they prefer context specific values or generic values.

No.	Motivation	Value preference
1	- It is very nice to be philosophical about values, but generic values are far from this research and are very abstract. I think it is better to select the values based on the study that has to be done.	Context specific
2	- Generic values are too complicated to interpret, the words should be simpler. I don't think generics ever work properly. Something that works well in a certain area doesn't have to work elsewhere, it's about breeding ground. Humanly speaking Schwartz is encompassing, left list of values contains everything.	Context specific
3	- The generic part has a high scientific content. Can you still connect with the participants and can participants still find these values in their own motivation? - Generic could work at a higher level of abstraction and with highly educated people. - Since PVE is a mass instrument, you need to work with context specific values to let people connect to their values.	Context specific / (Generic value for an abstract topic)

4	- Generic values do not appeal to my imagination, hard to translate to actions. You can explain some generic values in multiple ways - Advantage of generic, you can fill it out very broadly and I like it that there are less values. - If you show the generic values to the city council, it's far from what you want. - If you communicate with generic values, you should not tell it to participants.	Context specific
5	When I look at the generic values, it appeals less than the generic ones. All-encompassing values remain too vague and top down, you have to make things clear and concrete. - You have to make things concrete and clear, in principle the list is never complete, and you should use it dynamically.	Context specific
6	- What I think speaks for a specific list, it indicates that it is not finished. Values should never be presented as a list that is finished otherwise you don't seem to be right if your value in not in the generic list. - Generic values are vague and do not overlapping in daily life.	Context specific
7	- Context specific is about what we think is important as a society. What do you want the government to do for you and on your behalf? - Generic values are more often personal values, it could also be valuable, this depends on the purpose. For example, policies related to behaviour change can work with generic values. What policy is in line with what the goal is. When you talk about how public money is spent, context specifically seems more interesting to me.	Context specific / (generic values in behavioural subjects)
8	- Generic values are very applicable to moral subjects. For example, wind farms are not possible and generate a new value set. - Topics like euthanasia or death penalty can use the generic set of values. It might also be valuable to use generic values in i.e. climate justice, but that should be discovered.	Context specific / (generic values in moral subjects)

8. Evaluation of results by decision-makers and stakeholders involved in the process of Súdwest Fryslân

There are numerous examples of how automatized processing of qualitative data can be used in decision-making. Some of these methods have been tested in the case studies. The aim of testing these methods has not been to find the best algorithmic performance as this research is exploratory. Still, it did give some insight in the result that could be gained when improving this way of data analysis. In this chapter, some methods and results will be proposed to stakeholders in the process of the Regional Energy Strategy in Súdwest Fryslân. It is always mentioned if the output was gathered by manual analysis or automated analysis. The goal of the interviews is to find out how stakeholders interpret the outcomes of this research and if this is in line with the desired goals that they define for this PVE. Even if it is possible to reach the best algorithmic results using high advanced and accurate NLP techniques, if the outcome is not valuable for the stakeholders, the PVE can still be perceives to fail in its purpose. For that reason, the reflection by stakeholders is essential on how values have been selected and how they might be retrieved by using automated data analysis. In short, the goal is to find out what information the stakeholders find important based on previous analysis, that can contribute in the process of decision-making.

Information about the participants

The people that have been interviewed were from the citizen panel, the municipality, the Technical University Delft, Public Mediation and NPBO. The municipality can be perceived the decisionmaker in this process. Their goal is to incorporate the values of their citizens, represented by a citizen panel. All other stakeholders are involved in the organization of the participation process and are expected not to benefit from the outcome of this process. Interviewing this group of people will result in a diverse point of view on this participation process in which PVE has been used as a decision-making tool. NPBO and Public Mediation are two companies that are guiding participation processes and did this for the RES in Súdwest Fryslân. The citizens of the citizens panel could give insight if this process and can reflect the thoughts of the public. The policymakers are within the municipality, who is the client of this process. They will be able to reflect on how they perceive the outcomes and how they will use this in the further process. The Technical University of Delft is involved to share their expertise in the field of social sciences, statistics and the method PVE itself. Public Mediation has a lot of knowledge about participation and the use of values in decision making. Lastly, NPBO is guiding the process from beginning until the end. They will know the origin of this research and the requirements set by the government about the process.

Setup semi-structured interview

A semi-structured interview will be conducted of which the lay-out will be the same for all interviewees. The structure can be split up in five main parts. The full interview can be found in Appendix K and contains additional questions if there is time left to discuss them. Also, the interview ended with the question if the interviewee wanted anything to say, ask or advice. The grey questions can be asked only if it is relevant based on the previous question.

1) **General experiences of PVE within the process of Súdwest Fryslân**
   To familiarize the interviewee with the process, he is being asked to share some general thoughts about the participation process in Súdwest Fryslân. This will also give them the opportunity to tell something about the role of PVE in this process. By asking this some general insights can be gained about PVE is a tool in this decision-making process.

   a. Looking back at the PVE process of Súdwest Fryslân, what would you perceive the output?
   b. Were you surprised by certain results: What did you not expect and what was it that you did expect from the results?
   c. What is for you the added value of PVE?
   d. What is in your opinion disadvantages of PVE?

2) **Needs in decision-making from participants' opinions**
   In this section we ask the interviewees why they think it's important to gather opinions from the public. What they would like to know from the citizens and how can this information have value in the decision-making process.

   a. What information do you need from the opinions of participants regarding the RES in SWF?

b. What is the value in having this information?
c. How would you like to receive those opinions to be able to use those in making decisions?

3) Experiences with human values in the decision-making
We want to know what's the perception of the interviewees is regarding using human failures in decision making. Can they come up with decision making processes in which they have made their decision based on public values? By asking for these examples we can validate the way in which these values could be used in decision making. If the participants did not have any experience with using values decision making, it will be asked with they think about it and why this could be of importance.

a. What do you perceive to be important from values related to decision-making?
b. Have you worked with values in decision-making before? (what was your experience with that?)
c. What is another method that you used to take in the opinion of inhabitants and how did that work for you? (was that accurate and easy to work with?)

4) Interpretation of results in Súdwest Fryslân
The 4th part of the interview it will be about the interpretation of the results. by sharing a presentation some visualizations of manual and automated data processing will be shown. These visualisations might have been shown before in a presentation, though there are also social new abstract and really concrete examples. By doing this it can be validated if the purpose of the researcher for creating this graph was clear to the participant of the interview. For this reason, it is first being asked to share thoughts about what they see in visualisations without being able to ask questions. They might have some advice; they can see opportunities do use this in a better way or they might think that some of the information is missing to interpret the visualisation correctly. The full interview including the introductions and the visualisations that have been shown Can be found in Appendix K.

A graph including a value overview for a specific policy option is shown (Figure 28).
a. What do you see in this graph?
   What does it mean?
   Could you give an example?
b. How could this be of value to underpin decision-making for you?
c. There are almost 700 motivations behind this graph, what information would you want to see to get more information out of this graph?

These figures are so called word clouds. An algorithm can look for different topics in all motivations, that means a group of words correlated to each other. This corresponds to the topics of the policy option 'Inhabitants do it their selves'.
d. For the topic above these word clouds are relevant. What would you perceive valuable from this? How do you interpret this information?
e. Which arguments would you take out of this word cloud?
f. 400 motivations are behind this graph, what information would you want to see to get more information out of this graph?

5) From values to design principles
Lastly, a more tangible visualisation will be shown. This value hierarchy can be seen as future possibilities when improving the automated data processing. By improving automated data processing and a significant amount of data, it could be possible to extract values out of the motivations. and to look for different argumentations for this specific value. This will always be a combination manual supervision together with automated data processing to quantify arguments and values. Figure 29 contains the value hierarchy that has been shown.

a. This figure is a value hierarchy. What would you perceive valuable from this? How do you interpret this information?
b. How could this overview that has been derived from values influence decision-making?
c. This figure is a value hierarchy. What would you perceive valuable from this? How do you interpret this information?
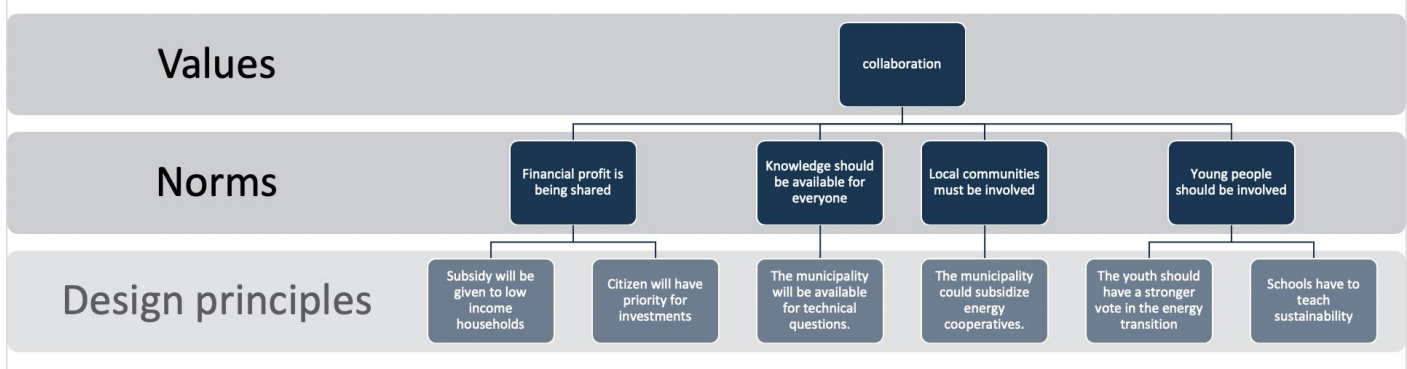
Figure 29 - Value hierarchy for 'collaboration' used in the interview

Results of the interviews

The results of the interviews will be discussed based on the different topics that can be seen in the semi-structured interview of Appendix K.

1) **General experiences of PVE within the process of Súdwest Fryslân**
   The municipality thinks that they know the position of their citizens by involving the tool PVE in this process. Their previous participation approach was an information or discussion evening. They gave the (Dutch) statement: 'There's a lot of dredging coming over you' in the physical meetings. Most interviewees explicitly mention that they think PVE is not the method for this process, but I tool that can be used to guide this process and can be used as input for discussion and policymaking. It is important to communicate the position of PVE to prevent people to think that it could be a referendum.

   Secondly, it has been mentioned that a digital approach could help due to a few reasons.
   a. A digital procedure can guide emotions as you can structure the information and questions. During a physical meeting it is often hard to prevent to discuss about the specifics.
   b. Many more people are involved, and you can speak with them indirectly, not only with the 'retired' participants of physical meetings.
   c. Participants do not influence each other. People of a physical meeting are often affected in their opinion by the people that have a strong opinion.
   d. The municipality hopes that a digital participation project also results in output for other projects. I.e. in the department of environment management when using this geospatial information of PVE.
   e. Being able to add qualitative data was important to enrichen the choice made.

   Also, some disadvantages have been mentioned. One of the often-mentioned disadvantages is about representativity. Even if there has not been a questionnaire of this scale with having 1.5-2.0% of the citizens participating, there are some doubts and citizens think that the promotion could have been done better. The municipality has insight in the website clicks to the survey, which is a lot higher than the number of participants. Many interviewees are afraid of overrepresentation of people with interest in sustainability and underrepresentation of the youth and the elderly. The last one has been mentioned as they think a digital participation process might exclude elderly. Another disadvantage is the level of education that might be needed to fill in the survey. Lower educated citizens were underrepresented, some interviewees think that it might be even a hard survey for higher educated people. There is quite some background and it takes 20 minutes, the municipality thinks is too long for a survey. Quotes given by the interviewees about the general advantages and disadvantages for this context have been listed in Table 15.

Table 15 - Quotes by interviewees about advantages and disadvantages of PVE in the process of Súdwest Fryslân

	Dutch	English
Advantages of PVE	"Wij weten hoe inwoners staan t.o.v. RES (energietransitie). Wij zijn hierin denk ik de enige in heel Nederland die weten hoe dit zit."	"We know how residents relate to RES (energy transition). I think we are the only one in the Netherlands who knows how this works. "

	"Wij komen uit een wereld van informatieavonden, wat werkt er niet aan. Vooral mensen die voor of tegen een deel van de avond zijn. Je wilt de achterhoede betrokken hebben. Dit brengt lof of bagger over je heen. Je gaat vanuit daar je beleid ontwikkelen, je kunt de vinger niet leggen op wat je precies zoekt. Men wil iets niet ivm persoonlijke mening. **Het grote voordeel van PWE, indirect met mensen spreken, veel meer mensen**. Dit kan eventueel met een congres, hier komen echter vooral gepensioneerde, niet de representatieve bevolkingsopbouw."	"We come from a world of information evenings, what is not working. Especially people who are for or against part of the policy proposals. You want to get the citizens involved. They give praise or shame about you. You start developing your policy from there, you cannot put your finger on what exactly you are looking for. People don't want something because of their personal opinion. The great advantage of PWE, speaking to people indirectly, much more people. This is possible with a congress, but the public is mostly retirees, not a representative population. "
	"De uitkomst van de PWE geeft een soort mandaat voor ons als bestuurder, dit kunnen de andere 17 gemeentes niet zeggen".	"The outcome of the PWE gives us a kind of mandate as a director, the other 17 municipalities cannot say this".
	"Iedereen kan zijn stem geven en het gevoel krijgen dat deze betrokken wordt en het standpunt wordt meegenomen. Een questionnaire zie ik niet als alternatief, deze wordt gestroomlijnd in een richting."	"Everyone can give their voice and feel that it is being involved and the point of view is taken into account. I don't see a questionnaire as an alternative, it's streamlined in one direction."
Disadvantages of PVE	"Een nadeel kan zijn leesbaar en gebruiksvriendelijk, niet iedereen gaat goed op tekst."	"A disadvantage can be readability and user-friendly, not everyone goes well on text."
	"Een PWE kan niet iedereen bereiken, de 'digibeten' zijn niet geïnteresseerd."	"A PWE can't reach everyone, the 'Digital illiterates' are not interested."
	"Het risico is dat de uitkomst een soort referendum is. PWE is niet bindend, mensen hebben niet de achtergrond/kennis en maken op basis daarvan beslissingen. De uitkomst vanuit emoties vertalen, de kennis in het advies is nodig."	"The risk is that the outcome is some kind of referendum. PWE is not binding, people do not have the background/knowledge and make decisions based on it. Translating the outcome from emotions, the knowledge in the advice is necessary."
	"Het is ontzettend lastig door te laten dringen dat alles fictief is. De scenario's zijn verzonnen zijn om te ontdekken wat belangrijk is. Ze willen meteen discusseren om te kijken wat geschikt is. Deze mensen kunnen dit niet scherp houden. Discussie ging snel over het scenario."	"It is extremely difficult to make it possible to see that everything is fictitious. The scenarios are concocted to discover what is important. Citizens want to discuss right away to see what's right. These people can't keep this on their toes. Discussion quickly turned to the specific scenario."
	"Het was erg ingewikkeld. "	"It was very complicated. "
	"De jeugd bereiken is omslachtig en lastig. "	"Reaching out to youth is cumbersome and difficult. "
	"Geen representatieve groep, beperkte geïnteresseerde groep. "	"No representative group, limited interested group. "

2) Needs in decision-making from participants' opinions

The citizen thinks that it is important to listen to the public as it is also their money that is being spend. The process managers think that it is essential for the municipality to know what people think. This can only be done if the citizens are involved in the process of analysing the data, so citizens themselves should present the public opinions to the municipality to make this a reliable process. This is related to support that another process manager mentioned, citizens feel supported by this participation. The municipality needs the arguments of the citizens to prevent i.e. the lack of feeling supported, but this also counts for the other important values. If the government is supported, this method can give a mandate to operate. Another important aspect are the concerns that people mention for the policy options, knowing the concerns and acting upon them can increase support for the municipality.

	Dutch	English
What is the value of having this information?	"Ik hoop dat ze het gaan gebruiken en de inhoudelijke basis vormt. Als ze dit meenemen hebben ze de argumentering al liggen. Daarmee heb je zicht op draagvlak. "	"I hope they will use it to form the substantive basis. If they take this with them, they'll have the arguments. This gives you a perspective on support. "
	"Vaak is het niet het terrein van de beleidsmaker, wel het geld van burgers om de dijkjes te herstellen. Als gemeente heb je hier degelijk een positie en wil je weten wat mensen willen. "	"Often it is not the domain of the policymaker, but the money of citizens to repair the dikes. As a municipality you have a solid position here and you want to know what people want. "
	"Opvallende nieuwe inzichten, misschien zijn dingen niet zo bedacht. Aannames testen is interessant. Kloppen de ideeën wel die belangrijk lijken voor de bevolking. "	"Striking new insights, maybe things have not been conceived that way. Testing assumptions is interesting. Are the ideas that seem important to the population correct? "
How to present the outcomes of PVE?	"Het burgerpanel maakt keuze wat aan te dragen. Anders zou de gemeente een ranking willen hebben van de cijfers. Het is duidelijk dat interpretatie belangrijk is, een burgerforum ertussen geeft een ander resultaat."	"The citizen panel makes a choice what to contribute. Otherwise, the municipality would like to have a ranking of the numbers. It is clear that interpretation is important, a citizen forum in between gives a different result."
	"Ambtenaren lezen rapport, mensen identificeren waardes als SWF waardes. Hoe kan beleid waardes bevestigen. "	"Officials read report, people identify values as SWF values. How can policy confirm values. "

3) Experiences with human values in the decision-making

Citizens have never experienced the use of value-based decision-making. The municipality did have two failed examples. In the first example values have been used for savings, though, only 70/170 surveys were useful, and this was not representative enough. For their second value-based project, they have rented a bus for a month to gather opinions about environment management. This was a costly process; they have visited weekly markets in many cities and only gathered 150 opinions. Both projects have never turned into action for decision-making.

The other participants mentioned one other project in Friesland, but they mainly mentioned values to be used in technological designs and not in decision making. Three participants have mentioned their expectation that value consultation only works in an early stage of a project, before policy designs have been made.

Participants did all agree that focusing on values improved the quality of this process. Most important results are that values are essentially about what people find important, not about specific ranking of choices. Also, people cannot judge each other on their values, this creates a higher sense of togetherness. If arguments are

being used, participants often try to discuss and negotiate based on them, which is not relevant when creating the vision of a policy.

	Dutch	English
What do you think about the use of values?	"Bij waardes gaat het over wat mensen echt vinden, dichter bij de mening. "	What do you think about the use of values? "Values are about what people really think, closer to the opinion. "
	"Door het kijken naar waardes geef je inwoners een stem in een moeilijk dossier. Het geeft vooral aan waar de beperkingen liggen, wij dienen iets te besluiten waar mensen betrokken worden, mensen het mee eens zijn en waar mensen in kunnen participeren."	"By looking at values you give residents a voice in a difficult file. It mainly indicates where the limitations lie, we have to decide something in which people are involved, people agree and in which people can participate. "
	"Het is lastig een scheiding tussen waardes en argumenten te maken. Het is lastig te onderscheiden, het loopt vaak door elkaar. Dit zie je terug in de principes. Een argument, is dat een reden waarom je iets wilt, of een voor/tegenargument voor waarde A. Argumenten kun je over discussiëren en onderhandelen, waardes kan je overleggen, maar niet oordelen. Zelfstandigheid is niet/wel belangrijker dan een andere waarde."	"It is difficult to distinguish between values and arguments. It is difficult to distinguish, it often mixes up. This is reflected in the principles. An argument is that a reason why you want something, or a pro / counter argument for value A. Arguments can be discussed and negotiated, values can be discussed, but not judged. Independence is not / is more important than another value. "
	"We zitten op strategieniveau, daarom wil je op niet praktisch oplossingen focussen of op belangen. Maar voor strategiebepaling is dit essentieel. "	"We are at the strategy level, which is why you want to focus on impractical solutions or on interests. But this is essential for strategy determination. "
Previous policy making project that have used values.	"4 jaar geleden bezuiniging op de begroting samen met inwoners gedaan. Hier zijn bezuinigingsvoorstellen uitgekomen, er waren 170 voorstellen, de eerste 70 meningen konden we niks mee, uiterst teleurstellend voor deelnemers… Toen naar waarden gekeken om mee te nemen voor bezuinigen."	"We cut the budget together with residents 4 years ago. Savings proposals were made here, there were 170 proposals, the first 70 opinions were useless, extremely disappointing for participants… Then looked at values to consider for cutbacks. "
	"Ook met omgevingsproces, er was een bus gehuurd met omgevingsvisie, 150 mensen hebben deelgenomen. Op markt Bolsward en Sneek, waar mensen zich niet konden voorbereiden. Er zijn wel waarden opgehaald. Mensen konden zich niet voorbereiden op de vragen, hier was het lastig om waarde aan waardes te hechten."	"Also with the environmental process, a bus was rented with an environmental vision, 150 people participated. At the Bolsward and Sneek markets, where people could not prepare. Values have been retrieved. People could not prepare for the questions, here it was difficult to value values. "

4) **Interpretation of results in Súdwest Fryslân**
All participants were able to interpret the value overview, they mainly mentioned the extremes that could be seen in the graph. That is already interesting as the municipality perceives the high percentage of leadership to

be a mandate for them to operate. The process managers did indicate that there should not be a mandate based on these results, instead it might validate other analyses that have been made, the mandate should result from the process, not from the data. The municipality also concluded from this graph that their citizens have a lot of knowledge and expertise. When realizing that all values have their arguments, they were wondering how they could involve the public even more to make use of their knowledge. Decision makers also used this graph to validate if they could identify their organization with the values. If values are underrepresented, this can be a motivation to work on these values in decision making and communication. All interviewees did agree that the graph without explanation is not valuable. They demand the data underlying the graphs and the arguments that are corresponding to the values. Also, an important limitation is the lack of sentiment, it is not sure if people are positive of negative about the values.

The word clouds that have been shown to most of the interviewees were hard to interpret. Contrastingly, by just giving them some time, they were really good in determining the arguments that could be found by making sentences from the words. One person mentioned the use for exploring the data, though most interviewees did not see any value just based on this information. For them it needs another representation and it is advised to use sentences instead of words, as the interpretation of single words can be hard. If this methodology can help to quantify arguments, policymakers are positive, though they mention to skip the step of using this visualization in the communication.

	Dutch	English
How do you interpret the value overview of Figure 28?	"Het is duidelijk dat de gemeente de leiding moet nemen."	"It is clear that the municipality must take the lead."
	"Veel mensen zeggen leiding, gemeente moet leiding nemen. "	"A lot of people say leadership, the municipality must take the lead. "
	"Gemeente neemt de leiding en biedt ondersteuning en kennis. De kennis hoeven we niet zelf te leveren, maar via een adviesbureau kan ook. Ook is samenwerking belangrijk, mensen moeten ondersteunt worden in kennis en geholpen worden in samenwerking. "	"Municipality takes the lead and offers support and knowledge. We do not have to provide the knowledge ourselves, but it is also possible through a consultancy. Collaboration is also important, people must be supported in knowledge and helped in collaboration. "
	"Er is een rol voor de overheid, mensen zien een rol. Ze zitten niet te wachten om het met elkaar op te lossen. "	"There is a role for the government, people see a role. They are not waiting to solve it together. "
	"We gaan geen besluitvorming doen zonder hier rekening mee te houden. Samenwerking met eigen gemeenschap, maar ook met overheid is belangrijk. "	"We are not going to make decisions without taking this into account. Collaboration with your own community, but also with the government is important. "
	"Duidelijk, procentuele aanduiding is goed, kleuren zouden wat beter kunnen. Kleuren geven interpretaties aan waardes. Rood zou slecht kunnen zijn en groen zou goed kunnen zijn. "	"Clear, percentage indication is good, colors could be a bit better. Colors give interpretations to values. Red could be bad and green could be good. "

5) **From values to design principles**
   The value hierachy of Figure 29 gained most attention to the participants as it could be a bridge to solve the problem of intangibility of values and the previous visualizations. Before making a value hierarchy, it is important to think about the goal and how it will be used. It should not become the truth in itself and readers must realize that the value hierarchy is never completed.

Arguments or guidelines in favour of this method were:
- This overview can help policymakers to think of their own design principles and add them to the overview. Discussing the norms and design principles will boost creativity and easily familiarizes the policymaker with the arguments of citizens.
- This would save the municipality a lot of research. They often don't know the risks in advance, this would really help to make policies and to use it as an instrument in discussions.
- This method can be used in the beginning of a participation process to familiarize the citizens with the data.
- The method shows a good overview of the topics that play a role. Use sentences and make sure the data behind the motivation is transparent.
- Counsellors would like to test assumptions, that could be done in this way.

Disadvantages / prerequisites:
- The guiding principles might become a truth by themselves and should not be used as a checklist.
- It might be misused if some policymakers use it as an instrument in favour of their own policy. If other policymakers don't have the time to dive in these results, it will be harder to come up with good arguments.
- If it is being used as a checklist, it might detect problems that would have never been a problem if the information was not available.
- It might lead to a lot of discussion if this is being used in a citizen panel, it could give direction to a discussion that is not relevant.
- There is an overlap in the values. This should not be a problem, but this should be realized. It is interesting to find these overlaps to know which values can be improved by a specific design principle.

	Dutch	English
Advantages of the use of value hierarchies	"Ik zie hier het belang van in. Vooropgesteld, het voorkomt veel uitzoekwerk."	"I see the importance of this. First of all, it prevents a lot of research. "
	"Ik vind dit een overzichtelijk geheel, waardes, waar normen onder liggen en handvatten of ontwerpprincipes om mee aan de slag te gaan."	"I think this is a well-organized entirety, values, under which standards lie and handles of design principles to get started with."
	"Het is vrij helder, mooi hoe die teruggebracht is. Het werkt vrij logisch. Zelf zou ik het ook vaak op deze manier uitwerken."	"It is quite clear, nice how it has been brought back. It works quite logically. I would often work it out in this way myself. "
	"Ambtenaren willen aannames testen, dat zou op deze manier kunnen. Er zit overlap in onderwerpen, maar dat is niet erg. "	"Civil servants want to test assumptions; it could be done this way. There is overlap in subjects, but that's okay. "
Disadvantages of the use of value hierarchies	"Je kunt niet zien of het allesomvattend is. Uitvoering en beheer mist erin."	"You cannot tell if it is all-encompassing. Execution and management are missing. "
	"De informatie kan nooit bindend zijn. "	"The information can never be binding. "
	"Je bent nooit volledig in een overzicht. Het biedt handelingsperspectief, projecten kunnen erop inspelen."	"You are never complete in an overview. It offers action perspective; projects can respond to it. "
How can value extraction i.e. value overviews or value hierarchies contribute in decision-making?	"Erover door gaan denken. Ondervraging aangaan en daarna naar aanleiding van het luisteren meer aandragen."	"Keep thinking about it. Entering into interrogation and then, as a result of listening, suggest more."
	"Uitdiepen waar behoefte aan is, zaken aanraken wat kan prikkelen en aandragen waar behoefte aan is."	"Explore what is needed, touch things that can stimulate and provide what is needed."
	"Dit is een goede manier. Centrale waarde, er liggen bepaalde normen onder en handelingen die men zelf in kan vullen. "	"This is a good method. Central value, there are certain standards behind and actions that people can fill in their selves."
	"Met deze overzichten kun je concreet aan de slag gaan met projecten, waar hebben we het over, wat gaan we doen. Het moet niet in een laatje belandden. Projecten uitrollen die hierover gaan."	"With these overviews you can get started with projects, what are we talking about, what are we going to do. It should not be forgotten. Roll out projects that deal with this. "
	"Je wilt informatie hebben en handelen. Je wilt aan bepaalde waardes invulling geven, handreikingen voor bepaalde waardes."	"You want to have information and act. You want to give substance to certain values, guidelines for certain values. "
	"Je kunt hiermee van waarden naar handelingsperspectief. Ambtenaren kunnen het overzicht aanvullen. "	"You can move from values to an action perspective. Civil servants can supplement the overview. "

6) Advice

Some general advice and feedback have been given on the process, most of the suggestions have been done related to one of the topics above. The other thought and recommendations are:

- Create different PVEs for the amount of time people have.
- Use sentiment in the analyses, you can create the same value overview, but divide the bar in positive and negative.
- Use other colour schemes in bar plots, a rainbow could influence the emotion that is being perceived by different values.
- Try to combine a value hierarchy with quantification to show the relevance of the different arguments. Keep also the unique arguments that are not often mentioned but that can be valuable.
- It might be interesting to use personas and present information via the principle of storytelling. This will increase participation, especially for the youth and lower educated citizens.
- The participation process involved citizens in the design of the PVE and in the interpretation and analyses of the PVE, this highly increases the trust in this method.
- It is scientifically interesting to look at conflicting values. How do your personal values relate to the values of the 'average' citizen?

	Dutch	English
Advice or tips for future PVE's	" Jongeren via sportorganisaties betrekken. (kaatsvereniging, voetval) "	"Involve young people through sports organizations. (handball club, football) "
	"Misschien kan je met personalia werken. "	"Maybe you can work with personal details. "
	"Het gemiddelde opleidingsniveau in Súdwest Fryslân is MBO, op een totaal van 100 punten uitkomen kan een uitdaging zijn. "	"The average level of education in Súdwest Fryslân is MBO, reaching a total of 100 points can be a challenge. "
	"Sentiment, positief en negatief uit tekst halen zou een goede toevoeging zijn. "	"Getting sentiment, positive and negative out of text would be a good addition. "
	"Als bewoners het eindresultaat niet herkennen ontbreekt het draagvlak. Zorg met alle methodes dat dit duidelijk is. "	"If residents do not recognize the end result, there is no support. Make sure this is clear with all methods. "
	"Iedereen kiest de argumenten uit zijn stelling. Als hij de analyse doet, het burgerforum doet het advies, zij zouden de data moeten kunnen zien. Burgers willen best struinen. Toegang is fijn, maar het proces moet goed georganiseerd zijn. "	"Everyone chooses the arguments from their position. If he does the analysis, the citizen forum gives the advice, they should be able to see the data. Citizens are willing to stroll. Access is nice, but the process must be well organized. "
	"Hoe hangen jouw ideeën samen met dat van anderen. Als je als doel hebt om zo goed mogelijk het beleid te maken, dan wil je weten wat je zwaktepunten zijn. Als je kijkt naar waardes van mensen kan je leren wat de samenleving belangrijk vind. Waar kijken we naar in onze communicatiestrategie, en klopt dit. "	"How do your ideas relate to those of others. If your goal is to make the best possible policy, then you want to know what your weaknesses are. If you look at people's values, you can learn what society considers important. What do we look for in our communication strategy, and is this correct. "
	"Wat als je minpunten toe zou kunnen kennen in de PWE?"	"What if you could assign negative points in the PWE?"

Conclusion of interviews

There is a demand in getting insight in the qualitative data, it seems that these qualitative results have been interpreted to be the most valuable within the process of PVE in Súdwest Fryslân. The interviewees see a lot of potential in the use of PVE for evaluation and participation. The main advantage is that there are a lot more people involved to whom the municipality can listen. These participants do have a lot of knowledge that can and should be used in decision making according to the decision-makers. The participants propose prerequisites and design principles that can be connected to values. Using values for decision making has never been used this explicitly in Súdwest Fryslân, it gave people the possibility to discuss about something that you cannot judge about. It did deliver a valuable contribution as the values resulted in practical guiding principles. The expectation is that values mainly work in an early stage of the project, at that moment there is not yet a discussion about a windmill in your back garden but is more about the vision of the policy. There is demand for using automated data processing to retrieve values, as long as the process is transparent. Stakeholders should be able to get insight in the motivations of people. One promising method is by creating value hierarchies, this can either be done for all values, or this can be done upon request. These value hierarchies are perceived useful if they are not presented as being complete and the truth, the hierarchy should not become a checklist on its own. They could be a contribution when discussing how to involve certain values into the policy. Thinking and adding design principles to the hierarchy will increase creativity according to the interviews. The main concerns of the stakeholders are related to representativity of the participants and regarding the reliability of the analysis due to the lack of sentiment. However, the lack of sentiment did not highly affect the process in Súdwest Fryslân as interviewees agree that PVE and its results should be a transparent to be used for interpretation of stakeholders. To conclude, the interviews did validate the usefulness of (automatically) extracting values from the qualitative data of participants.

9. Conclusion

Due to the demand for public participation, Participatory Value Evaluation (PVE) has been proposed to fulfil the needs of involving many participants in a scalable way in the process of participation and evaluation. PVE it is unique in the way that it allows people to be a virtual decision maker in the specific context. PVE is not a decision-making tool in itself but allowing participants to make their quantitative decisions and asking them to motivate their answers qualitatively results in a lot of data. In the case studies, the qualitative data has mentioned to be the most valuable, but the previously used method cannot process al the qualitative data. This research has the goal to investigate the possibilities of applying automated data processing for the qualitative data of PVE. To be able to transfer this knowledge to other PVE's, the main interest has been on to categorize this data based on human values, but also to validate language processing methods that are not applicable for values. The following knowledge gaps have been addressed:
- There was no solution to include all qualitative data in the analysis of PVE in a structured way for large scale public participation.
- This is the first PVE analysis that structures human values extracted out of all qualitative data.
- The practical usage of automatically generated data has been validated via interviews.

In order to address these knowledge gaps, the PVE method has been applied to two case studies in which value-based decision-making is of interest. There is not one definition of the minimal data that is required for applying techniques. This depends on the complexity of the data and the required accuracy. A guideline of at least 1.000 participants has been used. The first case study is the participation process of the municipality of Súdwest Fryslân with 1356 participants regarding their policy on who should take the lead in the Regional Energy Strategy. The second case study is the evaluation process conducted for The Dutch Ministry of Health and Environment. >30.000 participants in this PVE decided on the policy how to reduce the Covid-19 measures. By using the two case studies the main research question could be explored.

How to support public participation by using automated qualitative data processing in Participatory Value Evaluation to extract values that can be used in decision-making?

First the background in public participation and the need of qualitative information from the participants has been discovered (SQ1, SQ2). Then the methods for automated data processing have been discovered and these methods have been explored within the case studies (SQ3, SQ4). Lastly, a validation study based on interviews has been performed to reflect with stakeholders on the results from the qualitative data and to discover which opportunities they see to apply this in their policy making processes (SQ5).

To be able to conduct public participation at a large scale, PVE can be suitable as an online evaluation tool that has no limits in the number of participants. The results of PVE can help policymakers in legitimacy, support and quality of decision-making. The qualitative output of this method results in context specific arguments, prerequisites and design principle. Especially in an early stage of policymaking, decision-makers want to prevent discussion and negotiation about specific policy choices. For this reason, is has been validated that focusing on values bridges this gap as people cannot judge each other based on their values. It is however essential to have insight to the arguments, prerequisites and design principles that are related to these values. First, this is needed to familiarize and discuss about the values. Secondly, this is needed for decision-makers to translate the values into actions for their policy. They want to make use of the expertise of the public to incorporate the values within policies.

Currently, the manual approach of qualitative data processing within PVE is not structured in a way that can be quantified. For this reason, a semi-automated process has been performed in which all language processing of the PVE in Súdwest Fryslân has been done manually. By doing this the potential results of automated data processing could be simulated to reflect on the advantages and disadvantages of results. Being able to process the quantitative results of qualitative data makes it possible to create visualizations. It can be concluded that manual language processing involves a lot of uncertainties due to biases. This does not have to be a problem in conceptualisation of policies. Still, caution of the human uncertainties is needed if this data will be used as an input to train automated methods. Also, the sentiment is essential to interpret the data correctly and use them within Natural Language Processing (NLP) techniques. In this research basic NLP techniques have been explored that have led to the

following results: First, it is possible to create insight in the data by making a value taxonomy guided by NLP, this save 75% of this time as the data can be analysed more efficiently. Secondly, within five minutes the positive and negative topics within policy options can be defined to familiarize with the motivations of participants. Lastly, NLP techniques will be able to perform topic modelling based on values; the current techniques for detecting values are limited so this requires training of the algorithms in Dutch on which research is still being done at the department of Intellectual Intelligence at the Technical University of Delft. It will be essential to train the algorithm based on manual annotated data that includes the keywords related to the annotated values. To increase the accuracy of NLP, it is always better to have more participants. An explorative semantic analysis to find positive and negative relations resulted in nine times more relations above the threshold in Covid-Exit due to the fact that this research has 21 times the number of participants.

Interviews have been conducted to let stakeholders reflect on the results from the qualitative data and to discover which opportunities they detect to apply this in their policy making processes. All interviewees agree on the added value of using values in the case study. These values should be chosen based on the context of the policy problem. According to the interviews, generic values in this context would not have been useful to be used as guidelines in policymaking, the values must be recognizable within the context. Value overviews can give a mandate to the stakeholder if they are i.e. associated with leadership. These interpretations of values can only be made in a transparent process in which the arguments, prerequisites and design principles corresponding to the values are available. Preferably, this should be done by phrasing the exact quotes of participants to increase legitimacy. One promising tool to give insight in the values and their arguments is creating a value hierarchy. Stakeholders in the case of Súdwest Fryslân perceive this as a valuable way to guide discussions, increase creativity and to reflect their policy based on design principles mentioned by citizens, the experts of their society.

To conclude on how to support public participation by using automated qualitative data processing in Participatory Value Evaluation to extract values that can be used in decision-making? The main requirement for public participation is the process, PVE is a tool that makes it easier to involve many people. The quality of participation can be increased by incorporating 'the public' in the creation and interpretation of PVE. Automated qualitative data processing offers a lot of opportunities to quantify qualitative data and to get more insights in the values, arguments, sentiments and prerequisites. Natural Language Processing will have the highest potential, but all stakeholders should be aware of the uncertainties and limitations, a transparent method is advices. The expertise of citizens can be used by policymakers to familiarize with 'the public opinion', to boost creativity of their policies and to create conditions for the policies that will be designed. Values can be used in early stage policymaking to involve the values and corresponding arguments in further policymaking.

10. Discussion and recommendation

Strengthening of PVE

First, a short summary will be given of possibilities for using automated data analysis within PVE. Next, the limitations will be given as well as the points of attention for the analyst.

As can be seen in chapter 6, automated data analysis can help the analyst in the interpretation of high amounts of qualitative data. First, it can help to guide through the motivations in such a way that the analysts read through a diverse set of motivation. This decreases the occurrence of reading to duplicate arguments and helps in familiarizing with the data. The data can be loaded into a programming script to use it instantly. If the policy analyst has a short amount of time to process the data, this method will be highly valuable. Secondly, automated data processing can be used to find the sentiment of topics. The most positive and negative words can be found in a policy option, this can be done instantly and it advice if time is limited. Next, different methods can be used for topic modelling. The models differ in how they have been setup. The result for all of them will be a set of motivations that correspond to a specific label. In both case studies, the labels have been values. A policy can be made based on the values; these values can become a prerequisite for the policymaker. This topic modelling can be done unsupervised, this takes less time, but it is less accurate as the analyst has not been able to define keywords corresponding to the values. The supervised method of topic modelling takes a lot more time, it is estimated to take about 20 hours to annotate a part of the data, to define keywords, and to validate the outcomes. However, as there is ambiguity in values, it is advices to involve different analyst. The time for each analyst is expected to be the same. This does not include the time for reporting. It is advised to conduct topic modelling if policymakers are interested in the argumentation behind the different values and if they would like them to be quantified. This can be done based on value sensitive design, this process helps policymakers to get familiarized with the data. Also, it will propose design principles by participants in such a way that the policymakers can add their own design principles. In this way the policy is being co-created by citizens.

The reason for this research was to find out how to overcome the limitations of manual qualitative data analysis. Scalability is one of the limitations for PVE to be used in bigger contexts as well as in grand challenges. One of the valuable results of qualitative data in previous PVEs has been the citation of unique, inspiring and interesting quotes by citizens. Natural Language Processing techniques will make it possible to guide the analysts through the motivations to minimize the analysis of duplicate arguments. This process does not guarantee that the best, most unique, or most representative quotes will be found. It will increase diversity of the quotes based on the text they contain. Secondly, stakeholders would like the data, arguments and annotations to be transparent. In this thesis, the insight in the data has been selected manually to be able to explain the selection of a certain value or argument within a motivation. If stakeholders gain insight in the data that has been labelled automatically, it might be hard to explain why a decision for a specific label has been made. This has not been in the scope of this research and caution is needed when applying the technology in decision-making. If the insight in this automated data will be accurate enough to explain its relationships, it will be a great contribution to the process of PVE. First, stakeholders can familiarize with the data. Secondly, they can use the important values in discussions and policy design to increase creativity. Lastly, the design principles and arguments will help to validate and reflect their policies. This only works if the stakeholder accepts the uncertainties, and they will need to trust the experts, analysts and methods used. For this reason, interviewees mention that the design of the participation process is more important than the PVE tool. If stakeholders do not trust the process, due to for example the bias of experts, then the results will not be accepted as well.

Contributing in solving grand challenges

Grand challenges are complex, they confront organizations with radical uncertainty and grand challenges are evaluative, implicating multiple criteria of worth and revealing new concerns when solving them (Ferraro, Etzion, & Gehman, 2015). Both case studies, for the energy transition in Súdwest Fryslân as well as how to reduce the measures taken against the Covid-19 virus do fulfil these three requirements of grand challenges. Many stakeholders are involved in both issues and there is not one solution for these complex issues. I think it is essential to involve the public in the decisions that have to be made. These participants might be the experts of the society and their contribution can lead to guiding principles for policymakers. Also, participation will grow support, increase legitimacy and can increase the quality of decision-making. Grand challenges are not only practical challenges, but they are emotional challenges as well in which a focus base on values could be valuable. This research does not include the

relationship between values and grand challenges, but they can also not be solved based on normative discussions and negotiations.

Model uncertainties

All data analysis face uncertainties, they should not be prevented, but awareness is essential to deal with uncertainties. If manually annotated data is being used to train an automated method, then these uncertainties should also be taken into account. The main human uncertainty is cognitive bias. As can be seen in chapter 6, the annotators highly differ in their annotations. A Kappa Cohen score of 0.42 has been reached. This might raise the question is this is reliable. According to theory this Kappa score is substantial, this will be due to the emotional character of values. Perception of values within textual data differs, they do not have one definition, even if this have been defined before annotating. It is important that annotators agree upon their different perception, the perception can differ due to cognitive bias (Jansen et al., 2019). This may be considered an important limitation of the current study. Also automated data processing faces uncertainties. In this exploratory research the methodological uncertainties have not been researched extensively. When testing a word embeddings model, the accuracy has been between 30-50%, which is not significant for policymakers. This accuracy will increase when Interactive Intelligent experts will create this model based on semi supervised topic modelling. Also, the model to determine relations between words and human values will have to increase in quality, this can be done by training the model based on PVEs. Another uncertainty is that English models are currently more accurate. For this reason, motivations are being translated at the cost of their semantics. Selection bias, label bias, model overamplification and semantic bias should be taken into account when working on the technologies (Shah et al., 2020).

The scale of the PVE also highly affects the accuracy of automated data processing. One interesting result is that the analysis for Covid-Exit results on average in 184 positive relations based on the threshold of 0.3, for SWF there are 20 positive words. This threshold is related to the certainty of the relation. This is due to the amount of data that the PVE of Covid-Exit contains compares to the PVE of Súdwest Fryslân.

Another uncertainty is related to the sentiment of the topics or labels. The sentiment has not been including in the topic modelling based on values in this research. Without sentiment, the analyst or policymaker cannot interpret the quantitative data correctly. An example is the case study of Súdwest Fryslân in which the municipality might use the mentioning of leadership as a mandate to take the lead in this process. This sentiment has been validated manually, though not quantitively. In an early process of decision-making, this lack of sentiment might not be essential as often familiarization with the opinions is the goal. Still, the municipality does not know the percentage of people that mention leadership in a positive way. This limitation has to be solved when using qualitative data as a mandate to make the policies and is recommended to take into account within Natural Language Processing.

Scientific recommendations for future research

Value-based deliberation

Based on these values, the method of Value Sensitive Design (Van der Poel, 2016) has be used. An example is given in Figure 30 in which the value 'collaboration' has been translated to the given norms of participants. The design principles are examples of principles that the municipality can decide upon. This approach by Van der Poel will more specifically present a way of translating values of a PVE into actions or policy-design requirements. However, this method has been created for technological designs. At this moment, it has not been validated in literature how this can contribute in the process of designing policies. For this reason, the theory has been proposed in the interviews to reflect on the usability within this context. Based on the interviews there is a demand for creating an overview of the arguments within values and an example of Value Sensitive Design has been verified to be useful. However, this research did not verify if the design of the value hierarchy based on Van der Poel can also be applied in the political context. Interviewees already mentioned that the hierarchy is never complete, future research can focus om combining the output of automated data processing with the theory of Van der Poel to construct a practical guide on how to incorporate values in decision-making.
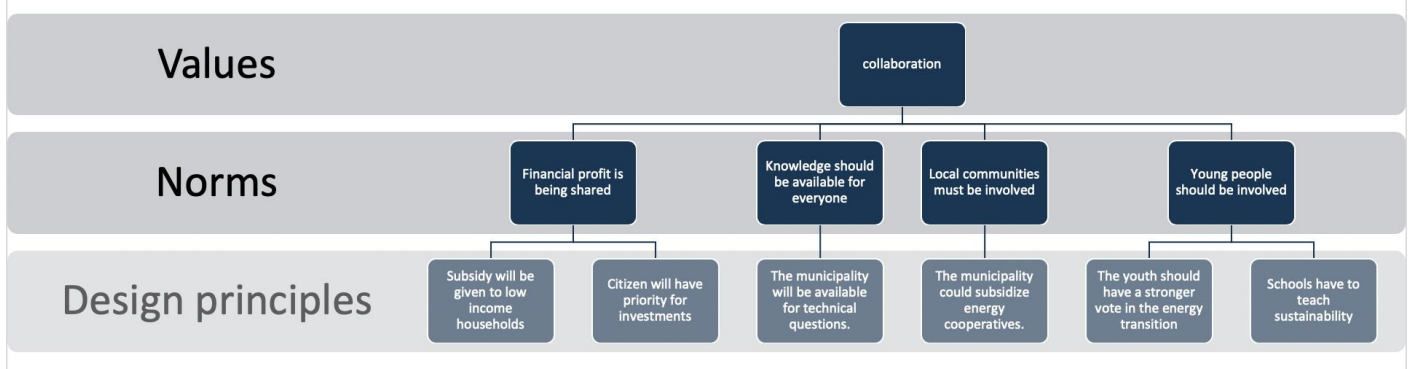
Figure 30 – An example of Value Sensitive Design applied for the value 'collaboration'

Using value-based deliberation in agent-based modelling

Within the master program of Engineering and Policy Analysis, the technique of agent-based modelling has been practiced. Agents can be trained based on values. The interesting combination with PVE could be made as PVE could provide a framework of values and corresponding arguments to train these agents for value-based deliberation. Within the course Data Analytics, such kind of agents have been trained by human input of argumentation. The field of Natural Language Processing offers huge potential in combination with PVE to provide these arguments. Literature shows some examples in which agent based modelling based on values has been use to simulate inhabitants' behaviour in energy management (Atkinson et al., 2011; Kashif, Hoa, Le, & Dugdale, 2013). First, static deliberation could be tested, in the next stage, dynamic deliberation among actors could be researched.

Creating a universal approach to determine value taxonomies.

The value taxonomy in this research has been created iteratively as described in chapter 6. This is a methodology that has not been used before for creating a value taxonomy. It seemed the context specific values are most relevant for decision-making. This is however at the cost of being able to compare a generic set of values in different context. A combination should be made in which some values are more generic than others, so that they can be used across contexts. I assume a kind of dynamic value taxonomy can be created, maybe based on Graham et al. (1991) that can be used to select values for a specific context. Within Natural Language Processing word embeddings are being trained across topics and contexts. For this reason, more generic use of values will increase the possibilities of transfer learning and increases the accuracy of Natural Language Processing.

Practical recommendations for future PVE

There are a few practical recommendations to be shared when performing PVEs in which human values will be extracted. There are some adaptions that could be made in the design of PVE.

- Most important and essential recommendation is the sentiment analysis based on the values that will be selected. At this moment, detecting values gives insight in the values that need extra attention. However, it does not yet give insight in the sentiment. People mention a specific values as they think this is an underappreciated value or they can mention a value as it corresponds to appreciation of that value for a specific policy option.
- It is recommended to let participants select options or to let them make other trade-offs instead of dividing 100 points. Two interviewees mentioned that dividing 100 points can be hard for (lower educated) people to participate. The second reason is that people do not have to make trade-offs. Some participants want to say something about all options, so they give some policies just a few points to be able to give qualitative reasoning. This reasoning is often negative and influences the analysis of qualitative data. It is advised to only gather the qualitative data from people that are in favour of the option as automated data processing will increase in quality when knowing the overall sentiment of the motivation. The concerns about the policy-option can be mentioned later in the PVE if people really want to give their opinion. Another method that would even be better for algorithmic purposes, but that will be time intensive is to ask people why they choose an option, but also why they did not choose an option. Knowing why an option is negative can be used to improve the word embeddings in Natural Language Processing.
- It would be interesting to develop the PVE in such a way that participants can be confronted with contrasting opinions. So, if they i.e. select a few values from a list that they think are important, the algorithm can

propose motivations of contrasting values. Doing this will broaden the view of participants and it might increase support as they are confronted by other opinions that they might not be confronted with in their daily life. This could increase the acceptance of an outcome as people can familiarize with the outcome based on their own opinion, as well as the opinions of others that they have read in the PVE.

- It is advised to annotate based on a smaller set of values. Showing a bar plot containing 18 different items did scare some of the policymakers/participants as they find it hard to detect the main message that can be practically used. Also, if future work will try to automatically detect values and maybe confront participants with contrasting opinions, the accuracy of NLP will increase when the number of labels decrease.

- It is advised to involve participants creation of the PVE, as well as in the interpretation of the data and results. PVE has been a tool in the process of Súdwest Fryslân, but the validity of this public participation process is due to the involvement of citizens from the start until the delivery of the final report. This advice mainly holds for PVE's that have a high desire for public participation.

- It would be interesting to offer participants the option to take a PVE of different durations. This could add more information by the people that do have time and interest and want to spend an hour for sharing expertise. On the other hand, this could also involve citizens that might not be interested enough to spend 20 minutes on the survey, but who just want to join for a five minutes survey. The process of data analytics can be adjusted based on this process. It will also validate the assumption of some of the interviewees that only highly interested people join this extensive PVE.

- For Natural Language Processing, it would be ideal to ask for arguments on separate lines, in that way every argument can be processed to be a separate topic. In that way, every argument can be labelled based on its type, norm and values. If the Value Sensitive Design (Van der Poel, 2016) has the potential that many interviewees see, then a more detailed automated data process has to be designed to detect the different layers of information within motivations.

- For more practical recommendations, the interviews can be requested to gain insight in the advice given by interviewees.

Bibliography

Arnstein, S. R. (1969). A Ladder Of Citizen Participation. Journal of the American Institute of Planners. *Journal Of The American Institute Of Planners*, *35*(November 2012), 37–41. https://doi.org/10.1080/01944366908977225

Artstein, R., & Poesio, M. (2008). Inter-Coder Agreement for Computational Linguistics. *Computational Linguistics*, *34*(4). Retrieved from www.essex.ac.uk/Research/nle/arrau/

Atkinson, K. M., Bench-Capon, T. J. M., Cartwright, D., & Wyner, A. Z. (2011). Semantic models for policy deliberation. *Proceedings of the International Conference on Artificial Intelligence and Law*, 81–90. https://doi.org/10.1145/2018358.2018369

Bakarov, A. (2018). A Survey of Word Embeddings Evaluation Methods. Retrieved from http://arxiv.org/abs/1801.09536

Basu, S., Bilenko, M., & Mooney, R. J. (2004). A probabilistic framework for semi-supervised clustering. *KDD-2004 - Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 59–68. https://doi.org/10.1145/1014052.1014062

Bishop, P., & Davis, G. (2002). Mapping public participation in policy choices. *Australian Journal of Public Administration*, *61*(1), 14–29. https://doi.org/10.1111/1467-8500.00255

Boyd, R. L., Wilson, S. R., Pennebaker, J. W., Kosinski, M., Stillwell, D. J., & Mihalcea, R. (2015). Values in words: Using language to evaluate and understand personal values. *Proceedings of the 9th International Conference on Web and Social Media, ICWSM 2015*, 31–40.

Braman, E. C., & Jones, R. L. (2002). How to manage and display qualitative data using 'Framework' and MicrosoftÒ Excel. *Journalism Quarterly*, *12*(4), 610–612. https://doi.org/10.1177/107769905303000408

Charmaz, K. (2003). *Constructing Grounded Theory*. https://doi.org/10.16309/j.cnki.issn.1007-1776.2003.03.004

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, *20*(1), 37–46. https://doi.org/10.1177/001316446002000104

De Argumentenfabriek. (2012). Handreiking Waardenkaart Ruimtelijke Ordening.

de Haas, M., Faber, R., & Hamersma, M. (2020). How COVID-19 and the Dutch 'intelligent lockdown' change activities, work and travel behaviour: Evidence from longitudinal data in the Netherlands. *Transportation Research Interdisciplinary Perspectives*, *6*, 100150. https://doi.org/10.1016/j.trip.2020.100150

de Vries, W., van Cranenburgh, A., Bisazza, A., Caselli, T., van Noord, G., & Nissim, M. (2019). BERTje: A Dutch BERT Model. Retrieved from http://arxiv.org/abs/1912.09582

Dekker, T., Koster, P., & Mouter, N. (2019). The Economics of Participatory Value Evaluation. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3323645

Dietz, T., & Stern, P. C. (2008). *Public participation in environmental assessment and decision making. Public Participation in Environmental Assessment and Decision Making*. https://doi.org/10.17226/12434

Dutch Ministry of Economic Affairs and Climate. (2019). *National Climate Agreement-The Netherlands*. https://doi.org/10.1016/J.ENG.2016.04.009

Ferraro, F., Etzion, D., & Gehman, J. (2015). Tackling Grand Challenges Pragmatically: Robust Action Revisited. *Organization Studies*, *36*(3), 363–390. https://doi.org/10.1177/0170840614563742

Fritzsche, D. J., & Oz, E. (2007). Personal values' influence on the ethical dimension of decision making. *Journal of Business Ethics*, *75*(4), 335–343. https://doi.org/10.1007/s10551-006-9256-5

Gallagher, R. J., Reing, K., Kale, D., & Steeg, G. Ver. (2017). Transactions of the Association for Computational Linguistics. *Transactions of the Association for Computational Linguistics*, *5*(0), 529–542. Retrieved from https://www.transacl.org/ojs/index.php/tacl/article/view/1244/275

Gemeente Súdwest-Fryslân. (2018). Agenda Duurzame Ontwikkeling SWF. Retrieved from https://pdc.sudwestfryslan.nl/htdocs/wp-content/uploads/2019/02/Agenda-Duurzame-Ontwikkeling-2018-2022.pdf

Gerton, T., & Mitchell, J. P. (2019). Grand challenges in public administration: Implications for public service education, training, and research. *Journal of Public Affairs Education*, *25*(4), 435–440. https://doi.org/10.1080/15236803.2019.1689780

Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., & Ditto, P. H. (2013). *Moral Foundations Theory: The Pragmatic Validity of Moral Pluralism. Advances in Experimental Social Psychology* (1st ed., Vol. 47). Copyright © 2013, Elsevier Inc. All rights reserved. https://doi.org/10.1016/B978-0-12-407236-7.00002-4

Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. (1991). Mapping the Moral Domain. *The Journal of Nervous and Mental Disease*, *179*(7), 444. https://doi.org/10.1097/00005053-199107000-00016

Haidt, J. (2013). *The Righteous Mind: Why Good People are Divided by Politics and Religion*. Vintage Books. Retrieved from https://books.google.nl/books?id=U21BxGfm3RUC

Hanel, P. H. P., Wolfradt, U., Lins de Holanda Coelho, G., Wolf, L. J., Vilar, R., Monteiro, R. P., … Maio, G. R. (2018). The Perception of Family, City, and Country Values Is Often Biased. *Journal of Cross-Cultural Psychology*, *49*(5), 831–850. https://doi.org/10.1177/0022022118767574

Heath, H., & Cowley, S. (2004). Developing a grounded theory approach: A comparison of Glaser and Strauss. *International Journal of Nursing Studies*, *41*(2), 141–150. https://doi.org/10.1016/S0020-7489(03)00113-5

Hendriksma, M. (2020). Mienskip neemt regie windmolens over - Binnenlands Bestuur. *Binnenlands Bestuur*, (15). Retrieved from https://www.binnenlandsbestuur.nl/ruimte-en-milieu/nieuws/mienskip-voert-regie-over-windmolens.14008714.lynkx

Hills, M. D. (2002). Kluckhohn and Strodtbeck's Values Orientation Theory. *Online Readings in Psychology and Culture*, *4*(4), 1–14. https://doi.org/10.9707/2307-0919.1040

Hofstede, G. (2011). Dimensionalizing Cultures: The Hofstede Model in Context. *Online Readings in Psychology and Culture*, *2*(1), 1–26. https://doi.org/10.9707/2307-0919.1014

Hoogendoorn, J. (2017). *Onderstroom*.

Hoving, W. (2020, January 30). Alle bezwaren tegen windpark Nij Hiddum-Houw van tafel. *Friesch Dagblad*. Retrieved from https://frieschdagblad.nl/2020/1/30/alle-bezwaren-tegen-windpark-nij-hiddum-houw-van-tafel?harvest_referrer=https:%2F%2Fwww.google.com%2F

Irvin, R. A., & Stansbury, J. (2004). Citizen Participation in Decision Making: Is It Worth the effort? *Public Administration Review*, *64*(1), 55–65. https://doi.org/10.1111/j.1540-6210.2004.00346.x

Itten, A. (2017). Context and Content toward Consensus in Public Mediation. *Negotiation Journal*, *33*(3), 185–211. https://doi.org/10.1111/nejo.12182

Jansen, T., Claassen, L., van Kamp, I., & Timmermans, D. R. M. (2019). Understanding of the concept of 'uncertain risk'. A qualitative study among different societal groups. *Journal of Risk Research*, *22*(5), 658–672. https://doi.org/10.1080/13669877.2018.1503614

Jansons, A. (2020, June 16). Opinie: Energiebelangen provincie botsen met belangen dorpsbewoners. Retrieved from https://frieschdagblad.nl/2020/6/16/opinie-energiebelangen-provincie-botsen-met-belangen-dorpsbewoners

Joseph, K., Friedland, L., Hobbs, W., Lazer, D., & Tsur, O. (2017). Constance: Modeling annotation contexts to improve stance classification. *EMNLP 2017 - Conference on Empirical Methods in Natural Language Processing, Proceedings*, 1115–1124. https://doi.org/10.18653/v1/d17-1116

Kashif, A., Hoa, X., Le, B., & Dugdale, J. (2013). Agent Based Framework To Simulate Inhabitants ' Behaviour in Domestic Settings for Energy Management 31503 Agent Based Framework To Simulate Inhabitants ' Behaviour in Domestic Settings for Energy, (March), 190–199. https://doi.org/10.5220/0003150301900199

Kenter, J. O., Jobstvogt, N., Watson, V., Irvine, K. N., Christie, M., & Bryce, R. (2016). The impact of information, value-deliberation and group-based decision-making on values for ecosystem services: Integrating deliberative monetary valuation and storytelling. *Ecosystem Services*. https://doi.org/10.1016/j.ecoser.2016.06.006

Khatua, A., Khatua, A., & Cambria, E. (2019). A tale of two epidemics: Contextual Word2Vec for classifying twitter streams during outbreaks. *Information Processing and Management*, *56*(1), 247–257. https://doi.org/10.1016/j.ipm.2018.10.010

Klijn, E. H., & Koppenjan, J. (2000). *Interactive decision making and representative democracy: institutional collisions and solutions*. (J. T. O van Heffen, W J M Kickert, Ed.), *Governance in Modern Society*. Kluwer Academic Publishers,. https://doi.org/10.1007/978-94-015-9486-8

Kuylen, K. (2007). Levenswaarden en proactieve persoonlijkheid Universiteit Gent Faculteit Psychologie en Pedagogische Wetenschappen. Retrieved from https://lib.ugent.be/fulltxt/RUG01/001/289/889/RUG01-001289889_2010_0001_AC.pdf

Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, *33*(1), 159. https://doi.org/10.2307/2529310

Lieder, F., Griffiths, T. L., Quentin, Q. J., & Goodman, N. D. (2018). The anchoring bias reflects rational use of

cognitive resources. *Psychonomic Bulletin and Review*, *25*(1), 322–349. https://doi.org/10.3758/s13423-017-1286-8

Liu, H., Huang, Y., Wang, Z., Liu, K., Hu, X., & Wang, W. (2019). Personality or value: A comparative study of psychographic segmentation based on an online review enhanced recommender system. *Applied Sciences (Switzerland)*, *9*(10). https://doi.org/10.3390/app9101992

Maio, G. R., Hahn, U., Frost, J. M., Kuppens, T., Rehman, N., & Kamble, S. (2014). Social values as arguments: Similar is convincing. *Frontiers in Psychology*, *5*(AUG). https://doi.org/10.3389/fpsyg.2014.00829

Marissing, E. van. (2005, May). De aloude inspraakavond voldoet niet meer. *Aedes Magazine*, 64–67.

Meinshausen, N., & Bühlmann, P. (2010). Stability selection. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, *72*(4), 417–473. https://doi.org/10.1111/j.1467-9868.2010.00740.x

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *1st International Conference on Learning Representations, ICLR 2013 - Workshop Track Proceedings*, 1–12.

Mostert, E. (2003). The challenge of public participation. *Water Policy*, *5*(2), 179–197. https://doi.org/10.2166/wp.2003.0011

Mouter, N., Annema, J. A., & van Wee, B. (2015). Managing the insolvable limitations of cost-benefit analysis: results of an interview based study. *Transportation*, *42*(2), 277–302. https://doi.org/10.1007/s11116-014-9539-0

Mouter, N., De Geest, A., & Doorn, N. (2018). A values-based approach to energy controversies Value-sensitive design applied to the Groningen gas controversy in the Netherlands A values-based approach to energy controversies: Value-sensitive design applied to the Groningen gas controversy in the Neth. *Energy Policy*, *122*, 639–648. https://doi.org/10.1016/j.enpol.2018.08.020

Mouter, N., Koster, P., & Dekker, T. (2020). Participatory Value Evaluation versus Cost-Benefit Analysis : comparing recommendations in the context of urban mobility investments.

Mouter, N., Koster, P., Dekker, T., & Borst, P. (2018). Een Participatieve Waarde Evaluatie voor de Lange termijn ambitie rivieren, 1–72. Retrieved from http://ienw.participatie-begroting.nl/.

Mouter, N., Koster, P. R., & Dekker, T. (2019). *An Introduction to Participatory Value Evaluation* (TI Discussion Paper Series). *SSRN Electronic Journal*. Amsterdam. https://doi.org/10.2139/ssrn.3358814

Mouter, N., Spruit, S., Itten, A., Hernandez, J. I., Volberda, L., & Jenninga, S. (2020). *Als eenheid uit de intelligente lock-down Resultaten van een raadpleging onder intelligente lock-down*. Delft.

Mouter, N., Spruit, S., Itten, A., Shortall, R., Hernandez, J. I., Collewet, M., … Borst, P. (2020). *Bewoners kiezen aardgasvrije wijken*. Delft.

Nyborg, K. (2014). Project evaluation with democratic decision-making: What does cost-benefit analysis really measure? *Ecological Economics*, *106*, 124–131. https://doi.org/10.1016/j.ecolecon.2014.07.009

Ose, S. O. (2016). Using Excel and Word to Structure Qualitative Data. *Journal of Applied Social Science*, *10*(2), 147–162. https://doi.org/10.1177/1936724416664948

Pigmans, K., Aldewereld, H., Dignum, V., & Doorn, N. (2019). The Role of Value Deliberation to Improve Stakeholder Participation in Issues of Water Governance. *Water Resources Management*, *33*(12), 4067–4085. https://doi.org/10.1007/s11269-019-02316-6

Raaij, W. F. Van, & Verhallen, T. M. M. (1994). Domain-specific Market Segmentation, *28*(10), 49–66.

Ramteke, J., Godhia, D., Shah, S., & Shaikh, A. (2016). Election Result Prediction Using Twitter sentiment Analysis. *International Conference on Inventive Computation Technologies (ICICT)*, 1–5. https://doi.org/10.1109/INVENTIVE.2016.7823280

Richards, L. (2002). Qualitative computing—a methods revolution? *International Journal of Social Research Methodology*, *5*(3), 263–276. https://doi.org/10.1080/13645570210146302

Rietveld, R. (2020). *De Energietransitie: We doen het samen.*

Ruder, S., Peters, M. E., Swayamdipta, S., & Wolf, T. (2019). Transfer Learning in Natural Language Processing, 15–18. https://doi.org/10.18653/v1/n19-5004

Ruiz-Villaverde, A., & García-Rubio, M. A. (2017). Public Participation in European Water Management: from Theory to Practice. *Water Resources Management*, *31*(8), 2479–2495. https://doi.org/10.1007/s11269-016-1355-1

Sabini, J., & Silver, M. (1983). Dispositional vs. Situational Interpretations of Milgram's Obedience Experiments: "The Fundamental Attributional Error." *Journal for the Theory of Social Behaviour*, *13*(2), 147–154. https://doi.org/10.1111/j.1468-5914.1983.tb00468.x

Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. *Advances in Experimental Social Psychology*, *25*(C), 1–65. https://doi.org/10.1016/S0065-2601(08)60281-6

Schwartz, S. H. (2012). An Overview of the Schwartz Theory of Basic Values. *Online Readings in Psychology and Culture*, *2*(1).

Shah, D. S., Schwartz, H. A., & Hovy, D. (2020). Predictive Biases in Natural Language Processing Models: A Conceptual Framework and Overview, (January), 5248–5264. https://doi.org/10.18653/v1/2020.acl-main.468

Shuyang, Z., Heittola, T., & Virtanen, T. (2018). An active learning method using clustering and committee-based sample selection for sound event classification. *16th International Workshop on Acoustic Signal Enhancement, IWAENC 2018 - Proceedings*, 116–120. https://doi.org/10.1109/IWAENC.2018.8521336

Sofia, D., Gioiella, F., Lotrecchiano, N., & Giuliano, A. (2020). Cost-benefit analysis to support decarbonization scenario for 2030: A case study in Italy. *Energy Policy*, *137*(March 2019), 111137. https://doi.org/10.1016/j.enpol.2019.111137

Stålhammar, S., & Thorén, H. (2019). Three perspectives on relational values of nature. *Sustainability Science*, *14*(5), 1201–1212. https://doi.org/10.1007/s11625-019-00718-4

Steeg, G. Ver. (2017). *Unsupervised Learning via Total Correlation Explanation*. Retrieved from http://github.com/gregversteeg.

Steel Eye. (2020). Retrieved August 19, 2020, from https://www.steel-eye.com/news/steeleye-to-monitor-social-media-for-trade-reconstruction

Stoecker, R. (1991). Evaluating and rethinking the case study. *The Sociological Review*, *39*(1), 88–112. https://doi.org/10.1111/j.1467-954X.1991.tb02970.x

Straver, F. (2018, April 12). Ambtenaren op cursus: hoe verkoop je de burger een energietransitie? *Trouw*.

Thomopoulos, N., Grant-Muller, S., & Tight, M. R. (2009). Incorporating equity considerations in transport infrastructure evaluation: Current practice and a proposed methodology. *Evaluation and Program Planning*, *32*(4), 351–359. https://doi.org/10.1016/j.evalprogplan.2009.06.013

United Nations. (2019). World Population Prospects - Population Division - United Nations. Retrieved December 6, 2019, from https://population.un.org/wpp/Graphs/Probabilistic/POP/TOT/900

Van Asselt, M. B. A., & Rijkens-Klomp, N. (2002). A look in the mirror: Reflection on participation in Integrated Assessment from a methodological perspective. *Global Environmental Change*, *12*(3), 167–184. https://doi.org/10.1016/S0959-3780(02)00012-2

van Oudheusden, M. (2014). Where are the politics in responsible innovation? European governance, technology assessments, and beyond. *Journal of Responsible Innovation*, *1*(1), 67–86. https://doi.org/10.1080/23299460.2014.882097

van Wee, B. (2012). How suitable is CBA for the ex-ante evaluation of transport projects and policies? A discussion from the perspective of ethics. *Transport Policy*, *19*(1), 1–7. https://doi.org/10.1016/j.tranpol.2011.07.001

Ver Steeg, G., & Galstyan, A. (2014). Discovering structure in high-dimensional data through correlation explanation. *Advances in Neural Information Processing Systems*, *1*(January), 577–585.

Ver Steeg, G., & Galstyan, A. (2015). Maximally informative hierarchical representations of high-dimensional data. *Journal of Machine Learning Research*, *38*, 1004–1012.

Verbeek, H., Gerritsen, D. L., Backhaus, R., de Boer, B. S., Koopmans, R. T. C. M., & Hamers, J. P. H. (2020). Allowing Visitors Back in the Nursing Home During the COVID-19 Crisis: A Dutch National Study Into First Experiences and Impact on Well-Being. *Journal of the American Medical Directors Association*, *21*(7), 900–904. https://doi.org/10.1016/j.jamda.2020.06.020

Verschuren, P., & Doorewaard, H. (2010). *Designing a research project* (2nd editio). The Hague: Eleven International Publishing. https://doi.org/10.1075/btl.99.05rus

Wickham, M., & Woods, M. (2005). *Reflecting on the Strategic Use of CAQDAS to Manage and Report on the Qualitative Research Process. The Qualitative Report* (Vol. 10). Retrieved from http://www.nova.edu/ssss/QR/QR10-4/wickham.pdf

Wiedemann, G. (2016). Opening up to Big Data : Computer-Assisted Analysis of Textual Data in Social Sciences Author (s): Gregor Wiedemann Stable URL : http://www.jstor.org/stable/24142701 Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Us, *38*(4), 332–357.

Wilson, S. R. (2019). *Natural Language Processing for Personal Values and Human Activities. A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy (Computer Science and Engineering) in the University of Michigan.* University of Michigan.

World Health Organization. (2020). Mental Health and Psychosocial Considerations During COVID-19 Outbreak. *World Health Organization*, (January), 1–6.

Appendices

Appendix A. Policy options PVE Súdwest Fryslân

	1. The municipality takes the lead and unburdens you.	2. Inhabitants do it themselves.	3. The market determines what is coming.	4. Large-scale energy generation will occur in a small number of places (as opposed to being distributed).	5. Betting on storage (Súdwest Fryslân becomes the battery of the Netherlands).	6. Become an energy supplier in the Netherlands (Súdwest Fryslân helps the rest of the country in generating more sustainable energy).
Target	Relieving and helping residents	Generate energy independently	Let companies work effectively on sustainability	Limit nuisance to a few number of places	Renewable energy for later use	Help the rest of the Netherlands
Direction	Municipality of Súdwest-Fryslân	Residents, energy cooperatives	The market, energy companies	Energy companies, Dutch government	Municipality of Súdwest-Fryslân	Municipality of Súdwest-Fryslân
Cost	Public investments (governments)	Households bear costs	Private investments (business)	Private investments (business)	Private investments (business)	Public investments (governments)
Benefits	Inhabitants, municipality of Súdwest-Fryslân	Residents, energy cooperatives	External developers	Energy companies, local residents	Municipality of Súdwest-Fryslân	The region
Effects	Consequences up to the home	Large number of places with consequences	Unknown number of places are impacted	Some places with consequences	Large number of places with consequences	Number of places with consequences

Súdwest-Fryslân

Welkom

Welkom op de online raadpleging van de gemeente Súdwest-Fryslân. Wij nodigen je uit om je mening te delen over de energietransitie in onze gemeente. Deze raadpleging werkt het beste op een desktop of een laptop computer. Kleine schermen met lage resolutie kunnen een probleem geven.

VOLGENDE

Wat vind jij belangrijk?

Nederland wil in 2030 ongeveer de helft van de energie duurzaam opwekken, in 2050 is dit zelfs 100% duurzame energie. Dit staat in het klimaatakkoord. Súdwest-Fryslân moet ook een bijdrage leveren aan duurzame energie. De precieze hoeveelheid is nog niet vastgesteld, maar elke gemeente van Nederland wil proberen om de helft van zijn eigen energiegebruik zelf duurzaam op te wekken in 2030.

Súdwest-Fryslân wil hier een plan voor bedenken en wil weten wat voor haar inwoners belangrijk is. Daarom nodigen wij je uit voor deze raadpleging. Na deze raadpleging stellen we een burgerforum samen die een advies maakt over de uitkomsten van de raadpleging. De gemeenteraad van Súdwest-Fryslân zal in september het advies van het burgerforum behandelen.

Gemeente **Súdwest**-Fryslân

2. Inwoners doen het zelf

→ BEKIJK DETAILS

In wijken en dorpen werken inwoners van Súdwest-Fryslân samen om duurzame energie op te wekken. Ze organiseren zich bijvoorbeeld in energiecoöperaties om samen zonnepanelen te kopen. Of buren investeren samen in een dorpsmolen, een buurtbatterij of gedeelde warmtepomp. Ze helpen elkaar om energie te besparen. De winst komt aan de gemeenschap ten goede.

- Doel: **Zelfstandig energie opwekken**
- Regie: **Inwoners, energiecoöperaties**
- Kosten: **Huishoudens dragen kosten**
- Baten: **Inwoners, energiecoöperaties**
- Gevolgen: **Groot aantal plaatsen met gevolgen**

Súdwest-Fryslân
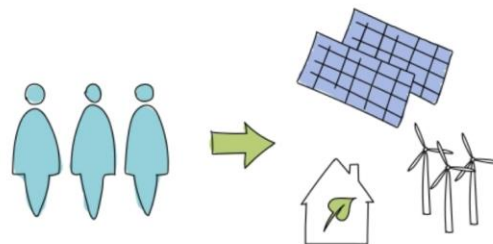
⊘ HELP   **→ VERGELIJKEN**   **→ BEVESTIG KEUZE**

Toegekend: 100

Tegoed: 0

Naam	Vergelijken	Selectie	
1. De gemeente neemt de leiding en ontzorgt	⚪	- 20 +	INFO
2. Inwoners doen het zelf	🟢	- 20 +	INFO
3. De markt bepaalt wat er komt	🟢	- 20 +	INFO
4. Op een klein aantal plekken grootschalige energieopwekking	⚪	- 40 +	INFO
5. Inzetten op opslag	⚪	- 0 +	INFO
6. Energieleverancier van Nederland worden	⚪	- 0 +	INFO

20 — 1. De gemeente neemt de leiding en ontzorgt

20 — 2. Inwoners doen het zelf

20 — 3. De markt bepaalt wat er komt

40 — 4. Op een klein aantal plekken grootschalige energieopwekking

Motivatie

Je hebt zojuist de volgende punten toegekend aan de mogelijkheden. Kun je aangeven waarom je de punten op deze manier hebt verdeeld?

Motiveer je selectie:

1. De gemeente neemt de leiding en ontzorgt   `INFO`

Motivatie

2. Inwoners doen het zelf   `INFO`

Motivatie

3. De markt bepaalt wat er komt   `INFO`

Motivatie

Appendix C. Policy options PVE relaxation of corona measures in the Netherlands

	1. Nursing and care homes allow visitors.	2. Employees in contact professions (barbers, beauticians, etc.) go back to work.	3. Young people may come together in groups.	4. Social contact within families is allowed again.	5. Hotels, restaurants, cafes and entertainment industry re-open.	6.Restrictions are lifted in Friesland, Groningen and Drenthe	7. Businesses open again, except for hotels, restaurants, cafes and contact professions	8. All restrictions are lifted for people who are immune.
Extra health pressure	10	15	8	10	25	30	15	20
Increase in deaths 70+	2.000	200	200	2.000	600	1.000	400	1.000
Increase in deaths under 70	100	500	100	750	500	500	750	500
Increase in the number of people with permanent physical health problems	500	3.000	500	10.000	5.000	5.000	2.000	3.000
Decrease in number of people with permanent mental health problems	60.000	7.500	7.500	30.000	60.000	30.000	2.000	5.000
Decrease in the number of households with long-term income loss	50	50.000	50	50	50.000	20.000	20.000	10.000

Raadpleging versoepeling coronamaatregelen

Welkom

Welkom bij deze online raadpleging over het versoepelen van coronamaatregelen in de periode 20 mei 2020 tot en met 20 juli 2020.

VOLGENDE

Na de uitbraak van het nieuwe coronavirus COVID-19 in Nederland heeft de overheid verschillende maatregelen genomen om de verspreiding van het virus te controleren, om risicogroepen zoals ouderen en mensen met een zwakke gezondheid te beschermen en om te voorkomen dat verschillende onderdelen van het zorgsysteem overbelast raken.

Nu de maatregelen effect lijken te hebben, is het mogelijk om een aantal van deze maatregelen te versoepelen. **Wilt u dat er tussen 20 mei en 20 juli maatregelen worden versoepeld? En zo ja, welke maatregelen moeten er als eerst worden versoepeld?** De overheid wil graag dat een grote groep van 5.000 Nederlanders een advies geeft over de wenselijkheid van deze 'versoepelopties' en u bent voor dit onderzoek geselecteerd.

Het onderzoek wordt uitgevoerd door onderzoekers van de Technische Universiteit Delft in samenwerking met onderzoekers van andere universiteiten en onderzoekers van het RIVM. Beleidsmedewerkers van het Ministerie van Volksgezondheid, Welzijn en Sport en het Ministerie van Financiën hebben ook meegekeken. De uitkomsten worden onder meer gedeeld met het RIVM en andere wetenschappers die meedenken/adviseren over het overheidsbeleid rond corona.

Wij willen u alvast hartelijk danken voor uw medewerking aan deze raadpleging!

Instructie Onderdeel 1

**We vragen u om de overheid te adviseren over twee zaken:**

1) Moet de overheid coronamaatregelen in de periode 20 mei – 20 juli **wel** of **niet** versoepelen.

2) Indien u vindt dat de overheid coronamaatregelen wel moet versoepelen tussen 20 mei en 20 juli, voor welke versoepelopties moet er dan worden gekozen?

Belangrijk:

1) U kunt een beperkt aantal versoepelopties adviseren. De druk op het zorgsysteem mag **niet meer dan 50%** toenemen.

2) U mag de overheid adviseren om **geen** maatregelen te versoepelen tussen 20 mei en 20 juli.

Effect = verschil tussen snel versoepelen en later versoepelen

We geven u informatie over de effecten van de verschillende versoepelopties. Als we het hebben over effecten, dan hebben we het over **verschillen** tussen het versoepelen van de maatregel in de periode 20 mei – 20 juli 2020 en het versoepelen van de maatregel in de periode 20 juli 2020 en 31 december 2020.
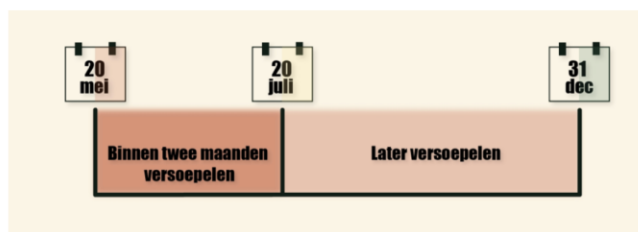
Druk op de knop onderaan om de instructievideo te bekijken. In de instructievideo leggen wij uit wat u in Onderdeel 1 moet doen. Wij vragen u om de **gehele instructievideo** te bekijken. U adviseert de overheid over een ingewikkeld en belangrijk

Raadpleging versoepeling coronamaatregelen

Rangschik op: Kies een effect:

Maximale druk op zorgsysteem: 50%

Extra druk op zorg: 34%

Extra druk	Versoepeling	Vergelijken	Selectie	
15%	Bedrijven gaan weer open (horeca en contactberoepen zoals kappers blijven nog wel gesloten)			INFO

Toename sterfgevallen 70+:		600
Toename sterfgevallen onder 70 jaar:		150
Toename lichamelijk letsel:		7.500
Afname psychisch letsel:		2.000
Afname inkomensverlies:		75.000

8%	Werknemers in contactberoepen (o.a. kapper) mogen weer werken			INFO
15%	Horeca en entertainment gaan weer open			INFO
25%	Verzorgingstehuizen staan bezoek toe			INFO
30%	In Friesland, Groningen en Drenthe worden beperkingen opgeheven			INFO
4%	Jongeren hoeven onderling geen 1,5 meter afstand te bewaren			INFO
15%	Directe familieleden hoeven geen 1,5 meter afstand te bewaren			INFO

INFO

15	Bedrijven gaan weer open (horeca en contactberoepen zoals kappers blijven nog wel gesloten)
4	Jongeren hoeven onderling geen 1,5 meter afstand te bewaren
15	Horeca en entertainment gaan weer open

91

Motivatie

U hebt zojuist de overheid geadviseerd om te kiezen voor de volgende versoepelopties. De overheid wil erg graag weten wat uw argumenten zijn om voor deze opties te kiezen. Kunt u per versoepeloptie aangeven waarom u hiervoor hebt gekozen?

Motiveer aub uw selectie:

Bedrijven gaan weer open (horeca en contactberoepen zoals kappers blijven nog wel gesloten) `INFO`

Motivatie

Horeca en entertainment gaan weer open `INFO`

Motivatie

Jongeren hoeven onderling geen 1,5 meter afstand te bewaren `INFO`

Motivatie
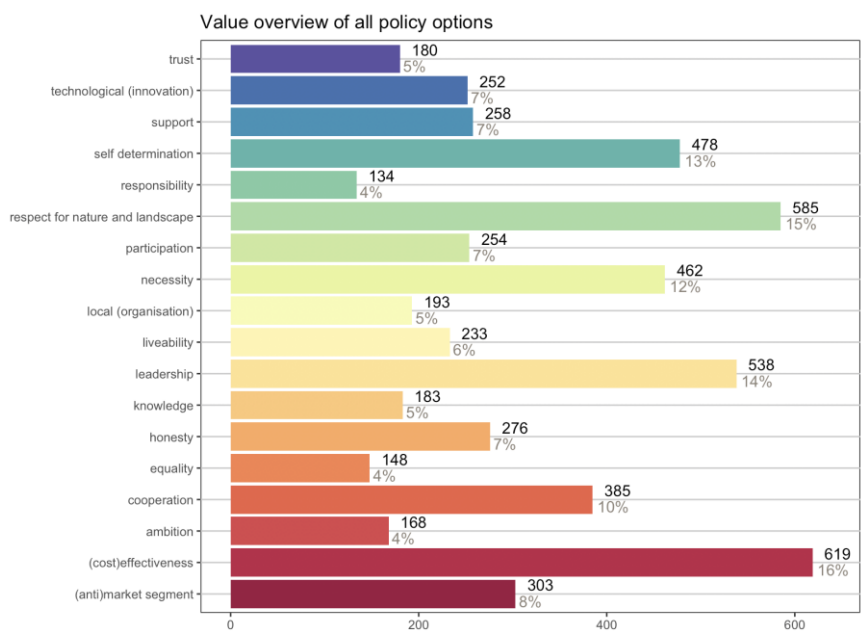
92

Appendix E.   Value overviews Súdwest Fryslân

Value overview of all policy options

Table 16 - Kappa scores for manual value annotation

Policy Option	ambition	fairness	equality	participation	knowledge	cost-effectiveness	landscape	liveability	leadership	local	(anti) market	necessity	support	cooperation	technology	responsibility	trust	self-determination
1	-0,01	0,19	0,76	0,54	0,80	0,76	0,39		0,69	0,28	0,31		0,41	0,43	0,00	0,44	0,55	0,48
2	-0,01	0,65	-0,01	0,76	0,66	0,49	0,43	-0,01	0,60	0,51	0,24		0,37	0,41	0,43	0,65	0,18	0,65
3	-0,01	0,53	-0,01	-0,01	0,72	0,68	0,66		0,76	0,50	0,37	0,23		0,60	0,44		0,19	0,32
4		-0,02	0,66	0,56		0,35	0,71	0,58		0,49		0,48		-0,02	0,32	0,66		
5	0,24			0,66		0,36	0,28	0,80		-0,01	1,00	0,70		0,60				0,33
6		0,50	-0,02	1,00		0,65	0,56	0,66		0,66	1,00	-0,01		-0,01		0,39		0,33
7	-0,01	0,21	0,53	-0,02	0,74	0,39	0,70	0,24	0,28	0,34	0,84	0,49	0,50	0,23	0,27	0,49	0,12	0,63

Policy Option	Most positive words	Coefficient	Most negative words	Coefficient
De gemeente neemt de leiding en ontzorgt	gemeente	0,558	opslag	0,396
	leiding	0,438	plekken	0,396
	regie	0,414	eigen	0,394
	overheid	0,396	markt	0,394
	centraal	0,394	ruimte	0,394
	kennis	0,39	zie	0,39
	mensen	0,39	landschap	0,372
	rol	0,386	draagvlak	0,364
	overzicht	0,382	opslaan	0,35
	centrale	0,38	nederland	0,326
	iedereen	0,372	vraag	0,322
	voorkomen	0,372	betrokkenheid	0,314
	sturing	0,368	zonnepanelen	0,312
	beleid	0,366	energie	0,302
	anders	0,364	we	0,302
	coã rdinatie	0,356		
	rdinatie	0,356		
	vertrouwen	0,352		
	burgers	0,348		
	alle	0,342		
	alle inwoners	0,332		
	houden	0,326		
	energietransitie	0,32		
	middelen	0,32		
	belangen	0,316		
De inwoners doen het zelf	inwoners	0,514	gemeente	0,396
	eigen	0,486	landschap	0,396
	draagvlak	0,422	markt	0,396
	zonnepanelen	0,42	opslag	0,396
	initiatief	0,404	plekken	0,396
	initiatieven	0,402	nodig	0,394
	betrokkenheid	0,396	we	0,392
	bewoners	0,396	lijkt	0,374
	mensen	0,39	weinig	0,368
	zie	0,386	nederland	0,366
	kleinschalig	0,37	leiding	0,356
	vaak	0,37	windmolens	0,354
	wel	0,352	overlast	0,348
	keuze	0,344	marktwerking	0,326
	inbreng	0,342	natuur	0,314
	meedenken	0,342	grootschalige	0,302
	peraties	0,342		
	dorpen	0,338		
	motivatie	0,334		

	burger	0,33		
	ideeen	0,326		
	situatie	0,31		
	gestimuleerd	0,302		
	isolatie	0,302		
	waar mogelijk	0,302		
De markt bepaalt wat er komt	markt	0,558	opslag	0,396
	marktwerking	0,4	energie	0,392
	bedrijven	0,394	inwoners	0,392
	marktpartijen	0,384	landschap	0,392
	komen	0,336	zonnepanelen	0,35
	nieuwe	0,336	bewoners	0,33
	innovatie	0,328	gemeente	0,316
	snel	0,312	beter	0,314
	prijs	0,31	eigen	0,312
	bedrijfsleven	0,306	plekken	0,312
	ondernemers	0,306	leiding	0,31
	vraag	0,304		
Op een klein aantal plekken grootschalige opwekking	landschap	0,486	gemeente	0,454
	plekken	0,486	eigen	0,396
	windmolens	0,424	inwoners	0,396
	plaatsen	0,398	markt	0,396
	overal	0,396	opslag	0,396
	natuur	0,394	belangrijk	0,392
	overlast	0,394	regie	0,384
	beter	0,386	overheid	0,362
	grootschalige	0,386	nodig	0,34
	last	0,384	opslaan	0,34
	horizonvervuiling	0,38	komen	0,328
	opwekking	0,378	burgers	0,326
	landschapsvervuiling	0,37	leiding	0,322
	grootschalig	0,366	draagvlak	0,312
	windmolenparken	0,35	duurzame	0,31
	energieopwekking	0,346	wij	0,304
	wildgroei	0,34	iedereen	0,3
	zonneparken	0,336		
	voorkom	0,334		
	omgeving	0,322		
	blijft	0,32		
	kernenergie	0,32		
	windparken	0,318		
	kleine	0,312		
	verrommeling	0,312		
	minste	0,31		
	bouwen	0,308		
	weinig mogelijk	0,308		

Inzetten op opslag				
	opslag	0,558	gemeente	0,408
	energie	0,428	inwoners	0,396
	opslaan	0,416	markt	0,396
	toekomst	0,402	mensen	0,396
	opgeslagen	0,394	landschap	0,39
	slaan	0,394	overheid	0,368
	nodig	0,378	plekken	0,368
	zon	0,366	initiatieven	0,346
	reserve	0,358	windmolens	0,338
	belangrijk	0,344		
	wind	0,344		
	pieken	0,338		
	lijkt	0,326		
	batterij	0,314		
	reserves	0,302		
Energieleverancier van Nederland worden				
	nederland	0,404	inwoners	0,396
	we	0,392	markt	0,396
	ruimte	0,362	opslag	0,396
	idee	0,354	mensen	0,348
	regio	0,35	zonnepanelen	0,342
	verdienen	0,348	bewoners	0,324
	werkgelegenheid	0,346	overheid	0,324
	banen	0,342	plekken	0,322
	friesland	0,338	regie	0,302
	nodig	0,338		
	leuk	0,33		
	energieleverancier	0,326		
	punt	0,316		
	swf	0,312		
	ambitieus	0,304		
	zie	0,3		

Appendix G. Correlation Explanation within the PVE relaxation of corona measures in the Netherlands

Table 17 - Determination of values by the use of Anchored Correlation Explanation

Policy option	Amount of arguments detected within CorEx	Amount of motivations
1 Nursing and care homes allow visitors.	6 out of 8	6200
2 Businesses open again, except for hotels, restaurants, cafes and contact professions (barbers, beauticians, etc.	6 out of 7	9240
3 Employees in contact professions (barbers, beauticians, etc.) go back to work	5 out of 9	11510
4 Young people may come together in groups.	5 out of 9	7447
5 All restrictions are lifted for people who are immune.	3 out of 6	1490
6 Restrictions are lifted in Friesland, Groningen and Drenthe (Northern regions less affected by the virus).	3 out of 7	783
7 Social contact within families is allowed again.	3 out of 8	7626
8 Hotels, restaurants, cafes and entertainment industry re-open.	3 out of 8	5955

One example of annotation of detected arguments within CorEx

Figure 31 Anchored Correlation Explanation of policy option 2: 'companies will open again'

Table 18 – Manually detected arguments for policy option 2: 'companies will open again'

Voor	Tegen
• Op deze manier kunnen we economische schade beperken • Hierdoor kunnen we de zorg ook in de toekomst betalen	• Deze maatregel brengt grote groepen mensen bij elkaar waardoor het risico op zieken toeneemt • Hierdoor zal een toename in vervoer zijn in Nederland

• Weer kunnen werken heeft een positief effect op het welzijn en de mentale gezondheid van mensen • Het is goed uit te voeren	• Thuiswerken is zo slecht nog niet

Voorwaarden
• Het moet veilig gebeuren • Er moet een optie blijven om thuis te werken • Scholen moeten ook volledig open zijn

Figure 32 - Anchored CorEx annotation on manual selected arguments

Appendix H.   Wordify PVE relaxation of corona measures in the Netherlands

Policy Option	Most positive words	Coefficient	Most negative words	Coefficient
1 Nursing and care homes allow visitors.	bezoek	0,74	bedrijven	0,536
	ouderen	0,706	economie	0,536
	eenzaamheid	0,634	economische	0,536
	leven	0,634	jongeren	0,536
	laatste	0,632	weer	0,536
	bewoners	0,536	elkaar	0,526
	eenzaam	0,536	horeca	0,502
	kwaliteit	0,536	minder	0,496
	leed	0,536	houden	0,49
	mensen	0,536	kapper	0,472
	onmenselijk	0,536	werken	0,472
	sterven	0,536	gaan	0,47
	verzorgingshuizen	0,536	beroepen	0,466
	verzorgingstehuizen	0,536	contactberoepen	0,464
	oudere	0,534	economisch	0,458
	bezoekers	0,53	blijven	0,456
2 Businesses open again, except for hotels, restaurants, cafes and contact professions (barbers, beauticians, etc.	bedrijven	0,74	bezoek	0,536
	economie	0,74	eenzaamheid	0,536
	economische	0,634	familie	0,536
	economie weer	0,536	familieleden	0,536
	economisch	0,536	groep	0,536
	kantoor	0,536	jongeren	0,536
	meter	0,536	kapper	0,536
	mogelijk	0,536	kleine	0,536
	schade	0,536	ouderen	0,536
	thuis	0,536	mensen	0,534
	thuiswerken	0,536	vaak	0,49
	werk	0,536	entertainment	0,484
	werken	0,536	leven	0,484
	open	0,534	beroepen	0,482
3 Employees in contact professions (barbers, beauticians, etc.) go back to work	beroepen	0,634	jongeren	0,634
	contactberoepen	0,634	afstand	0,536
	kapper	0,634	bedrijven	0,536
	kappers	0,634	bezoek	0,536
	mondkapjes	0,632	eenzaamheid	0,536
	beschermende	0,536	elkaar	0,536
	bescherming	0,536	familie	0,536
	beschermingsmaatregelen	0,536	familieleden	0,536
	beschermingsmiddelen	0,536	handhaven	0,536
	fysio	0,536	horeca	0,536
	fysiotherapeuten	0,536	kinderen	0,536
	goed	0,536	leven	0,536
	goede	0,536	meter	0,536
	juiste	0,536	ouderen	0,536
	klant	0,536	houden	0,534
	klanten	0,536	laatste	0,506

	kleine	0,536	contact	0,486
	maatregelen	0,536	immuun	0,486
4 Young people may come together in groups.	jongeren	0,74	bedrijven	0,536
	groep	0,536	bezoek	0,536
	handhaven	0,536	economie	0,536
	houden	0,536	economische	0,536
	jeugd	0,536	eenzaamheid	0,536
	minder	0,536	familie	0,536
	risico	0,536	familieleden	0,536
	scholen	0,536	horeca	0,536
	weinig	0,536	mensen	0,536
	moeilijk	0,534	weer	0,49
	ontwikkeling	0,534	werken	0,482
	kinderen	0,532	beroepen	0,472
	klein	0,532	contactberoepen	0,472
	school	0,526	economisch	0,472
	sporten	0,524	kapper	0,472
5 All restrictions are lifted for people who are immune.	immuun	0,58	jongeren	0,474
	immuniteit	0,536	bedrijven	0,432
	mensen	0,528	afstand	0,406
	besmetten	0,522	bezoek	0,404
	testen	0,512	contact	0,404
	gevaar	0,498	contactberoepen	0,404
	vormen	0,496	elkaar	0,404
	anderen	0,49	familie	0,404
	virus	0,478	familieleden	0,404
	getest	0,47	horeca	0,404
	risico	0,464	kapper	0,404
	verspreiden	0,464	kinderen	0,404
	logisch	0,46	mondkapjes	0,404
6 Restrictions are lifted in Friesland, Groningen and Drenthe (Northern regions less affected by the virus).	provincies	0,536	bedrijven	0,404
	besmettingen	0,532	bezoek	0,404
	regio	0,52	contact	0,404
	weinig	0,498	familie	0,404
	noorden	0,472	groep	0,404
	corona	0,47	jongeren	0,404
	land	0,462	mensen	0,402
	nederland	0,462	ouderen	0,402
	gevallen	0,454	eenzaamheid	0,388
	minder	0,426	belangrijk	0,382
	rest	0,422	werken	0,354
7 Social contact within families is allowed again.	familie	0,646	economie	0,634
	elkaar	0,634	bedrijven	0,536
	familieleden	0,634	beroepen	0,536
	contact	0,536	contactberoepen	0,536
	directe	0,536	economisch	0,536
	eenzaamheid	0,536	economische	0,536
	eigen	0,536	horeca	0,536

	families	0,536	jongeren	0,536
	gebeurt	0,536	kapper	0,536
	huis	0,536	leven	0,536
	kleinkinderen	0,536	mondkapjes	0,536
	maatregel	0,536	open	0,536
	praktijk	0,536	ouderen	0,536
	gezin	0,534	werk	0,536
8 Hotels, restaurants, cafes and entertainment industry re-open.	horeca	0,634	bezoek	0,536
	economische	0,536	contact	0,536
	entertainment	0,536	contactberoepen	0,536
	gaan	0,536	familie	0,536
	mensen	0,536	jongeren	0,536
	meter	0,536	kapper	0,536
	ontspanning	0,536	kappers	0,536
	restaurants	0,536	familieleden	0,534
	sector	0,536	ouderen	0,53
	weer	0,536	risico	0,514
	economie	0,534	eenzaamheid	0,51
	ondernemers	0,524	beroepen	0,49
	open	0,518	werken	0,474
	afleiding	0,506	weinig	0,472
	failliet	0,502	elkaar	0,468

Appendix I.   Interview invitation

Beste ...,

De RES voor de gemeente Súdwest Fryslân komt steeds een stap dichterbij. De gemeente is uniek geweest in de manier waarop zij burgers hebben laten participeren in dit proces. Dit door het evalueren van de beleidsopties door middel van de methode van Participatieve Waarde Evaluatie. Zelf ben ik in dit proces betrokken geweest voor mijn afstudeeronderzoek. Hierin kijk ik hoe we grote hoeveelheden kwalitatieve data semiautomatisch kunnen verwerken. In dit onderzoek gaat het om alle geschreven motivaties van de deelnemers waarom ze bepaalde beleidsopties gekozen hebben. Met enkele andere analisten hebben wij voor dit onderzoek de data ook handmatig doorgenomen. Vervolgens kijk ik naar de mogelijkheden om dit proces minder tijdsintensief te maken en meer bruikbare informatie uit de data te halen. Dit bijvoorbeeld met behulp van machine learning om de tekst te categoriseren.

Mijn vraag aan u is of u open zou staan voor een digitaal interview van ongeveer 45-60 minuten. In dit interview wil ik graag samen kijken naar een aantal resultaten en een aantal mogelijkheden van automatische dataverwerking. Het doel van de gebruikte methode in Súdwest Fryslân is uiteindelijk om ondersteunend te zijn voor beleidsmakers en hierbij analyses aan te bieden die het meest nuttig zijn. Het doel van dit interview is om te kijken welke mogelijkheden u ziet in deze nieuwe manier van data verwerken en hoe dit zou kunnen bijdragen aan een beter beleid. Hiervoor wil ik onder andere terugblikken op het proces van Súdwest Fryslân en wil ik een aantal visualisaties laten zien waar ik benieuwd ben naar uw interpretatie.

- Het onderzoek wordt uitgevoerd door een team van onafhankelijke wetenschappers van de TU Delft.
- De data wordt volledig **geanonimiseerd** opgeslagen. Het is niet mogelijk om te herleiden welke deelnemer wat heeft gezegd.
- We schrijven op basis van alle deelnemers een aanbeveling geven hoe organisaties in het vervolg kunnen leren van bijdrages van deelnemers van deze evaluatiemethode. Wellicht vindt dit wederom plaats binnen de gemeente Súdwest Fryslân, dit zou ook rond een soortgelijk onderwerp plaats kunnen vinden binnen een volgend onderzoek.

Indien u geïnteresseerd bent, zou ik graag in een moment afspreken indien dit gelegen komt.

Ik zie uit naar uw reactie, voor vragen kunt u mij bellen of mailen.

Met vriendelijke groeten,

Lionel Kaptein

Vragen of klachten?

Mail dan naar de onderzoekers: Shannon Spruit (S.L.Spruit@tudelft.nl)

Namens het TU Delft onderzoeksteam: Niek Mouter, Shannon Spruit, Pradeep Murukannaiah, Enrico Liscio, Lionel Kaptein en Luciano Siebert

Informatiebrief interviews rondom de Regionale Energiestrategie in Súdwest Fryslân – Delft University of Technology

<u>Auteur :</u> Lionel Kaptein (<u>l.n.kaptein@student.tudelft.nl</u>) gebaseerd op voorbeelden van <u>UK Data Services</u>
<u>Lastst aangepast:</u> 28 juli 2020.

U gaat deelnemen of heeft deelgenomen aan een interview of activiteit in het kader van het project van participatieve waarde evaluatie voor de gemeente Súdwest Fryslân. Via dit schrijven informeren wij u over de dataopslag en de omgang met de vergaarde onderzoeksmaterialen in dit project.

Doel van het project:
In het kader van het onderzoeksproject binnen de onderzoeksgroep "Participatieve Waarde Evaluatie (PWE)" willen wij graag interviews afnemen met een selecte groep betrokkenen. Het doel hiervan is inzicht te krijgen in de externe blik op de resultaten van de PWE aangaande de Regionale Energiestrategie van de gemeente Súdwest Fryslân. Door evaluatie van het proces willen we kijken op welke manier de resultaten van de (semiautomatische) data-analyse gepresenteerd kunnen worden op een manier dat dit de behoeftes vervult van deelnemers en de gemeente.

Dataopslag:
Interviews worden, afhankelijk van de voorkeur van de geïnterviewde, opgenomen met een voicerecorder of vastgelegd door middel van aantekeningen. Audio-opnames worden na afronding van het project gewist. Tijdens observaties zal er door de onderzoeker aantekeningen worden gemaakt van gespreksthema's. De data worden geanonimiseerd opgeslagen op een beveiligde server van de TU Delft. Wij willen de geanonimiseerde data beschikbaar laten voor toekomstig onderzoek. Wanneer u hier beperkingen aan wilt opleggen kunt u contact opnemen met projectleider. De persoonlijke gegevens die beschikbaar zullen zijn, zijn de volgende: de rol binnen het project. Alle andere persoonlijke gegevens zullen niet opgeslagen worden en niet gekoppeld worden aan dit onderzoek.

Toegankelijkheid van de data:
• Enkel de onderzoekers betrokken bij het project van de Gemeente Súdwest Fryslân en aangesloten bij de TU Delft of onderzoekers onder supervisie van de onderzoekers hebben toegang tot de data vergaard in dit project. Shannon Spruit is hierin de projectleider.
 • U heeft de mogelijkheid tot inzage in de opgeslagen transcripten van uw eigen interview en tot het beluisteren van de opnames van uw interview op aanvraag. Direct na het interview zal deze data geanonimiseerd worden en is opvragen niet meer mogelijk.
• Bij het delen van gevoelige informatie, kunt u een embargo-periode aanvragen. Wij zullen dan tijdelijk (periode in overleg) uw onderzoeksinformatie achterhouden.

Publicatie:
• De resultaten en quotes van dit onderzoek kunnen gepubliceerd worden in (internationale) wetenschappelijke publicaties. Dit zal anoniem plaatsvinden.
• Wetenschappelijke publicaties kunnen op aanvraag worden toegestuurd, neem hiervoor contact op met projectleider via <u>S.L.Spruit@tudelft.nl</u>.

Appendix K.   Questionnaire semi-structured interview

Naam:
Functie:

Persoonlijke introductie van elkaar, niet opnemen*

Privacy

Allereerst wil ik u vragen of ik dit interview mag opnemen. Dit zal worden gebruikt om een transcript uit te werken, zal de audio achteraf worden verwijderd.

Deelname aan het interview

Ik neem vrijwillig deel aan dit onderzoek en ga akkoord dat eventuele uitkomsten gebruikt kunnen worden voor een wetenschappelijke publicatie. Ik heb vooraf alsmede tijdens het interview de mogelijkheid gehad om eventuele vragen te stellen naar tevredenheid.

Ik begrijp dat ik vragen kan overslaan en kan stoppen met de studie wanneer ik wil.

Dit onderzoek bestaat enkel uit een interview van maximaal 60 minuten. Dit interview vindt digitaal plaats en vraagt geen andere verplichtingen van de deelnemer.

Ik geef toestemming dat de uitgewerkte data van het interview opgeslagen worden binnen de TU Delft en wellicht gebruikt kunnen worden voor verder onderzoek.

[optioneel]
Ik geef toestemming voor het maken van een geluidsopname voor de verdere uitwerking van het interview. Deze zal direct hierna verwijderd worden zodat enkel het uitgewerkte interview opgeslagen wordt.

Verder gebruik van de informatie

Ik geef toestemming dat de uitgewerkte data van het interview opgeslagen worden binnen de TU Delft en wellicht gebruikt kunnen worden voor verder onderzoek. Dit omvat de mogelijke publicatie van wetenschappelijke artikelen waar een anonieme quote in vermeld kan worden.

1) **SWF algemene inleiding aangepast op de rol van de geinterviewde in dit traject:** de ervaringen

Binnen het proces dat opgezet is vanuit het NPBO is een methode gebruikt voor de evaluatie en participatie, deze heet participatieve waarde evaluatie. Met deze methode is de online vragenlijst voor de inwoners opgezet.

- Terugkijkend op het proces van Súdwest Fryslân, wat ziet u als uitkomst van dit proces?

- Op welke manier heeft hierin de methode van PWE een rol gespeeld? Wat is hierin de toegevoegde waarde van PVE (en is dit de output die u had verwacht)?

- Wat zijn naar uw mening nadelen van PVE kijkend naar het proces van Súdwest Fryslân?

2) **Voorkeur/behoeften in de besluitvorming met de mening van de deelnemers**

- Welke informatie heb je nodig van de meningen van de deelnemers met betrekking tot de RES in SWF?

- Wat is de waarde van het hebben van deze informatie?

- Hoe wilt u die adviezen ontvangen om die te kunnen gebruiken bij het nemen van beslissingen?

- *Wat is een andere methode die je vroeger nam naar de mening van de inwoners en hoe werkte dat voor u? (Was dat nauwkeurig en gemakkelijk om mee te werken?)*

3) **We kozen ervoor om ons te concentreren op menselijke waarden,**
- Weet u wat ze zijn en wat vond je daarvan?

- Wat is volgens u belangrijk uit waarden die verband houden met besluitvorming?

- Heb je al eerder met waarden gewerkt in de besluitvorming? (Wat was uw ervaring hiermee?)

4) **Analysemethodes die voor SWF worden toegepast**
Ik zal een paar methoden /resultaten van geautomatiseerde gegevensverwerking in combinatie met een 'menselijke' selectie van waarden laten zien. Neem de tijd om een kijkje te nemen waarna ik een paar vragen zal stellen.

De gemeente neemt de leiding en ontzorgt
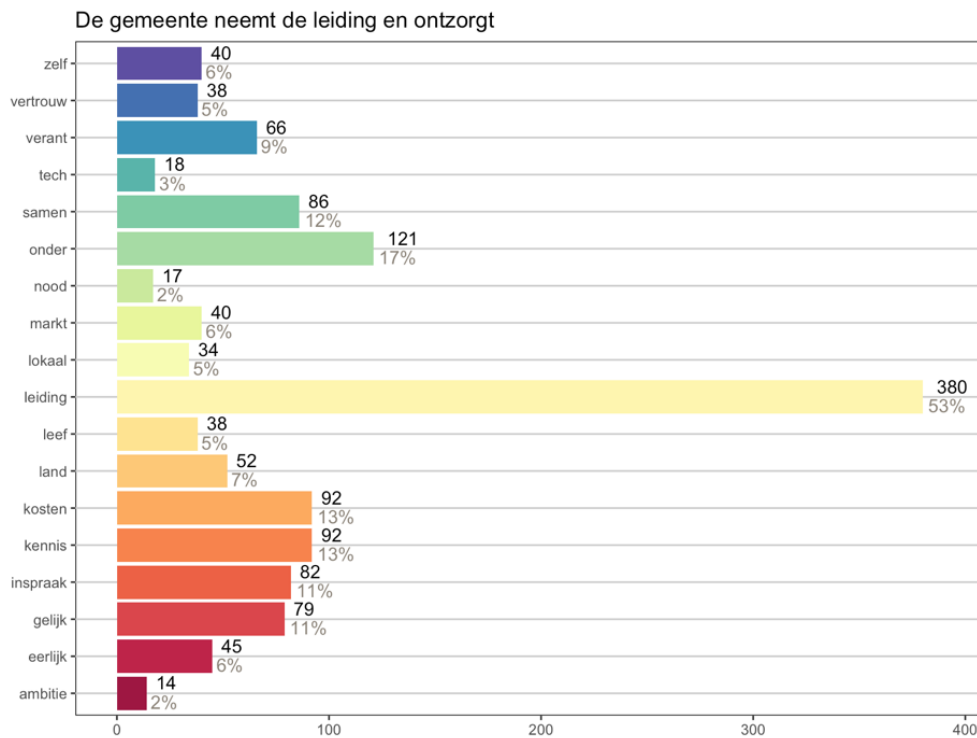
- Wat ziet u in deze grafiek? (indien lastig, probeer hardop te denken)
  - Wat betekent het?
  - u een voorbeeld geven?

- Hoe kan dit van waarde zijn om de besluitvorming voor u te ondersteunen?

- Er zijn bijna 700 motivaties achter deze grafiek, welke informatie zou je willen zien om meer informatie uit deze grafiek te krijgen?

De bovenstaande grafiek bevat waarden die specifiek relevant zijn in de context van deze PVE. Er zijn ook algemene waarden, bijvoorbeeld uit schwartz theorie die de waarden van Zelf-Richting, Stimulatie, Hedonisme, Prestatie, Macht, Veiligheid, Overeenstemming, Traditie, Welwillendheid, en Universalisme omvat.

Waardeoverzicht

context-specifiek   vs   generiek

Verantwoordelijkheid
Leiding
Kennis
Gelijkheid
Samenwerking
Ondersteuning
Vertrouwen
Kosten-effectiviteit
Ambitie
Zelfbeschikking
Landschapsbescherming
Inspraak
Eerlijkheid (van verdeling)
(Anti)markt segment
Technologie
Leefbaarheid
Noodzakelijkheid
Lokaal
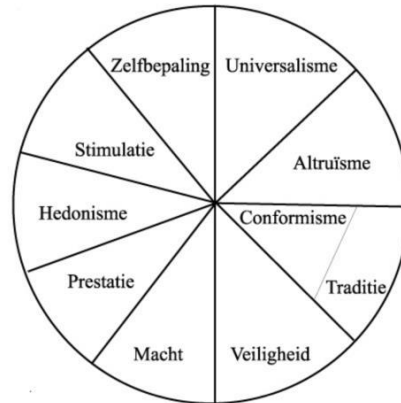
- Als we een grafiek tonen die lijkt op de bovenstaande, maar met algemene waarden in plaats van contextspecifieke waarden, hoe zou dat dan de besluitvorming beïnvloeden?


(Corex voor Beleid optie 2, doen de inwoners het zelf))

Deze cijfers zijn zogenaamde woordwolken. Een algoritme kan zoeken naar verschillende onderwerpen in alle motivaties, dat betekent een groep woorden gecorreleerd aan elkaar. Dit komt overeen met de onderwerpen van de beleidsoptie 'Inwoners doen het zelf'.

- Voor het bovenstaande onderwerp zijn deze woordwolken relevant. Wat zou je hier waardevol uit zien? Hoe interpreteer je deze informatie?

- Welke argumenten zou je uit dit woord wolk?

- 400 motivaties zitten achter deze grafiek, welke informatie zou je willen zien om meer informatie uit deze grafiek te halen?

-

5) **Van waarden tot ontwerpprincipes**

Ontwerpen op basis van waardes
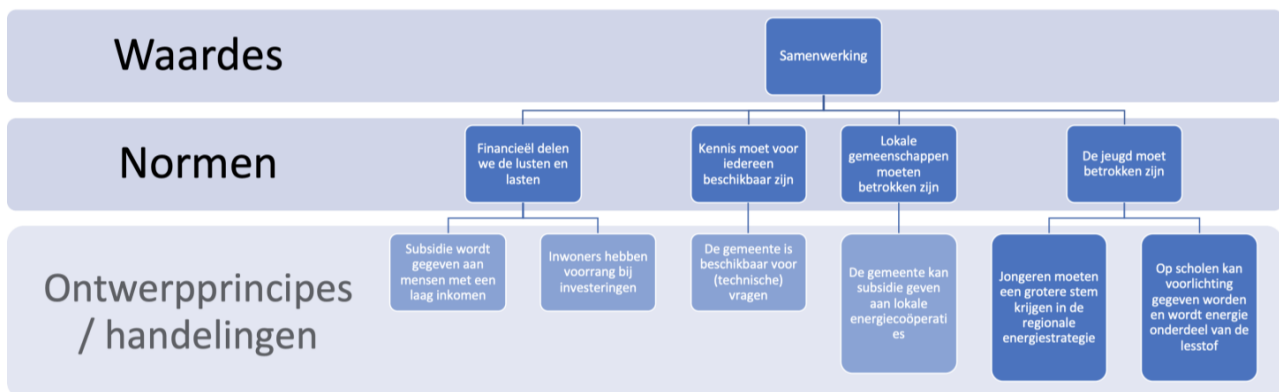
- Dit figuur is een waardehiërarchie. Wat zou je hier waardevol uit zien? Hoe interpreteer je deze informatie?

Dit is een gevolg van de PVE. Samenwerking bleek een belangrijke waarde van de deelnemers. Vaak stellen deelnemers al ontwerpprincipes voor in hun drijfveren. Door gebruik te maken van NLP kunnen de waarden worden geëxtraheerd, daarna kunnen de ontwerpbomen semi-automatisch worden gebouwd.

- Hoe kan dit overzicht dat is afgeleid van waarden de besluitvorming beïnvloeden?

6) **Nadenken over kansen** – wat mogelijk zou kunnen zijn **(vanuit het perspectief van de geïnterviewde)**
- Wat zou er mogelijk zijn door deze informatie over burgers te hebben??

- u komen met andere soorten van het resultaat / visualisaties die kunnen worden gebruikt voor de besluitvorming?

*** Vraag 7-9 zijn optioneel***

7) **Data inzichten, wat mogelijk zou kunnen zijn (optioneel, vanuit het perspectief van de onderzoekers)**

- Wilt u zelf in deze informatie duiken? Dit betekent dat u gemakkelijk waarden selecteren om de motivaties en argumentendoor telezen.

- Sommige van de bovenstaande processen zijn hybride in dat ze menselijke en machine intelligentie combineren. Bent u bereid om de rol van menselijke experts te spelen in dit proces?

8) **Publieke bewustwording (facultatief, voorstel van het idee)**
Machine learning kan het mogelijk maken om deelnemers gegevens of argumenten te verstrekken op basis van hun input. Een voorbeeld hiervan zou kunnen zijn om deelnemers argumenten te geven die tegen hun mening zijn. Daarna konden ze besluiten om een andere beleidsoptie te selecteren.

- Welke invloed heeft dit op de uitkomst van het beleid dat de gemeente voert?

9) **Specifieke vragen (optioneel)**

- Is het belangrijk dat je uitleggen wat er aan de hand is?

- Heeft u zorgen over automatische data-analyse met behulp van machine learning?

- Maakt het uit dat de data-analyse is geautomatiseerd?

10) **persoonlijke feedback (5 minuten voor het einde)**
- 'Is er iets wat je zelf wilt toevoegen'

Appendix L. Interview documentation

The documentation of the interviews is bundled and available upon request. This also includes the recording of the interviews. All are available at the local disk of the Technical University of Delft and can be requested via Dr. Niek Mouter, n.mouter@tudelft.nl