Utrecht University

THE ENDLESS JOURNEY TO INFOTOPIA

A Critical Inquiry of Dutch Open Government Data Policies, Using Textual Content Analysis and Metaphor Analysis

Author: Siem van Boxtel (3798275)
First Reader: Imar de Vries
Second Reader: Michiel de Lange
Master: New Media & Digital Culture

UTRECHT UNIVERSITY

The Endless Journey to Infotopia

A critical inquiry of Dutch Open Government Data policies, using textual content analysis and metaphor analysis

First Version

13 August 2015

Abstract: This thesis aims to shed light on the discourse surrounding open data in relation to Dutch governmental organizations in order to provide a better understanding of how Dutch governments currently conceive of- and frame open data. Two research methods are used: Textual content analysis and metaphor analysis. The claims and presuppositions about open data within five Dutch government policy documents have been identified using textual content analysis and have been juxtaposed to findings of academic studies. Through these comparisons it becomes clear that government's conception of open data is rather utopianesque building myths in order to convince others (public) of open data's potential. Data are always constructed yet this is sometimes neglected within government policy. Through metaphor analysis it becomes clear how this construction of open data is concealed and it becomes visible why certain ideas are more prevalent and seem more natural than others. Finally this thesis elucidates to what extent government's notions about open data, instead of realistic, can better be understood as idealizations of information which is called 'ideal information'. Concluding that openness within open data itself is a metaphor. Although myths might be necessary and may be indispensable when drafting policy, this thesis argues to be careful of creating an infomaginary in which all current dissatisfactions of social reality and desires for a better society are projected on information sources such as open data.

Author: Siem van Boxtel
Student nr.: 3798275
Master thesis: New Media and Digital Culture
Faculty: Humanities
Tutor: Imar de Vries
Second reader: Michiel de Lange

Preface

Before you are going to read this master thesis you should understand that its title forms a productive metaphor. What you are about to read is a story and not just an objective master thesis. A story is characterized for its subjective character.

Every letter, word and sentence that I put on paper is constructed, thought through, maybe re-written and finally those sentences make up my thesis. These words and sentences determine what kind of information I will present within the pages that are waiting to be read by you.

The moment you will read my words, the words change. Not literally on paper of course but in your mind. You will have to interpret 'my' information as it first starts out as data from your perspective. The moment you are able to read the data it becomes information and once internalized it becomes your knowledge. Data becomes information, information becomes knowledge. And there is a fair chance that your interpretation differs from others who read the same words and sentences, meaning that information and knowledge differs.

In our current society we tend to forget the constructed nature of data and information. It is not that every bit of information is a lie or false neither should we accept it as one truth. The story you are about to read is about people and governments who nevertheless believe in information as if it is unconstructed and objective, and thereby ideal information. Read what might happen when we skate over reality too fast…

Table of contents

Chapter 1: Introduction

In 2009 Barack Obama, president of the United States of America, signed the memorandum on transparency and open government. Short after Obama signed the memorandum, data.gov was created: an open data portal hosting hundreds of federal datasets for the public to access and reuse. Obama's signing and the creation of data.gov worked as a catalyst for an open government data movement (Davies 2012). Not long after the creation of data.gov, the United Kingdom followed and launched data.gov.uk in 2010. Continuing in this line the Dutch European Commissioner Neelie Kroes believes Europe should not lag behind. From the start of 2010, Kroes actively advocates an open data policy as part of Europe's digital agenda (Van Barneveld 2013). The Dutch government decides too to not stay behind and one year later, 2011, data.overheid.nl is launched. Numerous other open data portals have been created and are being created, but why? What arguments are used in favour of open data and what is the reasoning process behind these arguments?

When I started writing this thesis, the concept of open data was still in an early stage of development. To give you an idea of the subject matter I will use a definition of open data from the Open Data Handbook (ODH) by Dietrich et al. (2014). They define open data as "data that can be freely used and redistributed by anyone – subject only, at most, to the requirement to attribute and sharealike" (Dietrich et al. 2014). Many definitions confine open data primarily to be (public) data collected by means of public resources (i.e. tax). Thereby they automatically link open data to open government data (cf. Janssen et al. 2012; Schouten 2012; Hansen et al. 2013).[1] The ODH exemplifies how open data can be seen as a broader concept and potentially includes e.g. open science data as well or commercial data. Within this thesis open data is used in reference to open government data and Dutch open government data to be specific.

The idea that (public) data can be freely used and redistributed inspires many to be open data advocates. Amongst them is Kroes who argues that open data are an untapped goldmine of socio-economic value (European Commission 2011). And the authors of the Dutch Digital Agenda who concur with Kroes and regard open data to be a feedstock for innovative services (Digitale Agenda.nl 2011). The authors of Visie Open Overheid (VOO) take on a different perspective and describe how open data enables the sunlight to shine through growing a more transparent and thus a more democratic society (VOO 2013). While Kroes and the Digital Agenda mainly relate open data to economic and innovative benefits,

[1] When you consult the Dutch Wikipedia webpage related to open data, you are confronted with a governmental related definition.

the authors of VOO refer to its democratic potential. All sources agree (European Commission 2011; Digitale Agenda.nl 2011; VOO 2013) that open data is potentially of great value. Dietrich et al. claim opening up public data leads to a greater availability of information which potentially holds great beneficial value, enabling new services that "improve the lives of citizens and make government and society work better" (Dietrich et al. 2014). In other words Dietrich et al. argue the greater the availability and accessibility of public information, read open data, the greater the merits will be. Open data or open 'information' ultimately leads to a better society. Does it sound utopian? This thesis reveals why.

Dutch discourse on open government data

Currently the open data trend reached the consciousness of local Dutch governments and semi-governmental organisations. On a local, national and international level governments cooperate with organisations such as The Green Land and Open State foundation (OSF) and develop open data policies. Many open data projects and portals pop up. Recent projects within the Netherlands include; amsterdamopendata.nl, utrechtopendata.org, data.eindhoven.nl, rotterdamopendata.nl, hackdeoverheid, openspending, Waterschappen en Open Databeleid and Noord-Holland slimmer.

Parallel with governments developing open data policies and the rising number of open data portals there is an increasing number of research reports focusing on open data (Halonen 2012, 13). This is no surprise as open data provides for many research opportunities covering and connecting a multitude of different aspects, e.g. socio-economic or democratic impacts of open data on society, aspects related to law and legislation and implications of OGD concerning privacy and licenses, ICT related aspects that concern data portals, networks, open standards and dataset formats.

With a growing body of academic, political and popular literature concerning open data, the arguments covering the merits of open data are getting more established, accepted and widespread (Kitchin 2013). Benefits include contentions that open data, opening up data, leads to increased accountability and transparency; increases the efficiency of agencies; promotes public participation and social innovation, enhances public governance and fosters economic innovation and wealth creation. Empirical evidence to underpin the benefits however often seems to be lacking and potential problems or negative consequences often are neglected and less well examined (Ibid.). So if arguments are getting more established while the empirical evidence is lacking, to what extent are these arguments realistic? If the latter

part of the question appears to be true, would not this make governmental policies simplistic and even unrealistic? This begs the question if government policies are not overly utopian.

1.1 Purpose

Research within the field of new media studies has shown that *discourses* surrounding new media and communication technologies always come with utopian expectations fuelled by myths that surround a technology or medium (De Vries 2012). Discourses involve the "connection with culturally ingrained and institutionally powerful ways of looking at, experiencing and understanding particular areas of social life" (Deacon et al. 1999, 147). Because of this connection discourses deeply permeate what is allowed as legitimate knowledge in domains of social life and at the same time exclude other possibilities and other perspectives on those domains (Ibid.). Assuming that Dutch governmental institutions are powerful authorities within society, their way of looking at and understanding open data translated into policies plays an important role in constructing the overall Dutch open data discourse. Governmental policies co-determine society's conception of open data. Hence this thesis is particularly concerned with which knowledge about open data is allowed and counts as legitimate (utopian?), as well as which knowledge or perspectives are excluded and deemed irrelevant through government policy (realistic?).

In *Tantalisingly Close* Imar de Vries, assistant professor New Media & Digital Culture at Utrecht University, delineates how discourses of mobile wireless technologies are surrounded by utopian expectations. He explains that the interesting part of utopian expectations does not rely in whether such utterances or predictions eventually tell the truth or might come true, but the interesting part resides in language; in analysing the images and metaphors that are used to convey ideas of what such technologies are for. De Vries postulates: "However unprecedented the potential of mobile communication devices actually is to accomplish social, political, and economic equality, the gush of revolutionary and utopian claims that accompanies their incessant rise is not unique" (De Vries 2012, 15). Utopian narratives and expectations are visible throughout the evolution of media i.e. the telegraph, radio, telephone, television and finally mobile wireless communication technologies.

Every new medium comes with utopian expectations steered by recurring *myths* which give voice to persistent unrealizable beliefs and desires of reaching. This recurring discursive pattern is what De Vries defines as *ideal communication* (ibid. 17). In line with John Durham Peters (1999) and Katz & Aakhus (2002), De Vries defines ideal communication "as perpetual contact, the fulfilment of "sharing one's mind with another"; in

other words, as ubiquitous and pure communication without misunderstanding" (De Vries 2005, 1). Peters (1999) described it as communication like angels. Similar utopian expectations now seem to be present within surrounding discourses of open data.

The subject within this thesis concerns discourses surrounding open data and focuses specifically on Dutch governmental policies. From this perspective it might seem problematic to align with De Vries' work concerning ideal communication in relation to open data. Ideal communication has been conceptualised in relation to mobile wireless communication technologies such as the smartphone, tablet computer or more recent Google Glass. Ideal communication appears through the advent of new media or media technologies that were once new. How does this match the concept of open data?

It is difficult to determine exactly what media are and if open data fit the description as do the previous mentioned media technologies. Yet this thesis is not concerned with whether or not open data can be defined as new media but concerns the discourse surrounding open data, the arguments and conceptions underlying government policy about open data. In this light, could it well be possible that interesting parallels exist with what De Vries dubbed ideal communication? Might we need a variant of this concept which we could call ideal information? Since this new phenomenon, open data, is largely about information I propose a new variant in this thesis that sufficiently grasps what is underlying the discourse about open data and possibly other forms of new and old media. In order to investigate this matter, the following research question has been developed:

To what extent does ideal communication fit the discourse about open government data and to what extent does ideal information as a new concept help to better understand the open data discourse? The purpose of answering this question is not just to find out if government policies present a realistic view about open data and, if not, we will risk creating an *imaginaire* as Patrice Flichy (2007) would say. In the next paragraph I will go into more detail and delineate Flichy's conception of the imaginaire. Yet the overarching argument seeks to provide more insight into the pattern with which language is used within government open data policy and to characterise how meaning concerning open data is given direction.

To provide an answer on the above matters, I decided to focus on five Dutch government policy documents. These documents are part of the open data discourse and thereby give meaning to our conceptions and actions related to open data. Two methods account for the investigation of these documents: textual content analysis and metaphor analysis. Those two methods will be explained in detail in section 1.2.1 and 1.2.2. First however the theories that inspired me writing this thesis and their relevance will be explained considering the purpose and relevance of this thesis.

1.2 Relevance

Utopian narratives presenting new media to be the next self-evident solution in the quest for reaching ever-improved communication proliferate through everyday life. They can be seen in commercial ads, press releases, websites and so on. De Vries (2012) uses the commercial advertisements of NTT DoCoMo, a Japanese communications giant, as an example of utopian narratives. In these advertisements the company presents a seemingly perfect world made possible by the technologies and services the company created or will create in future. NTT DoCoMo's advertisements are just one of many. Therefore it is not interesting in whether these advertisements ultimately tell the truth. Instead we should look at the patterns shaped by the language used, the images and metaphors, to convey ideas of what technologies and new media are intended to do. Utopias presented by NTT DoCoMo and others do not merely paint a pretty picture but readily tap into our deepest desires for completeness which we ordinary mortals hope can be fulfilled by improving communication.

According to De Vries (2012) such ideas of progress and utopia are rooted in the human urge to hope and search for purpose and meaning. This urge expresses itself in *necessary fictions*, which are "regulative narratives that foster a sense of improvement by positing two situations that differentiate in time and then presenting the latter of those two situations to be in some way better than the former" (De Vries 2012, 27). It is a two-pole system which is being sustained by powerful myths. On the one end we live in a world in which we are confronted with mysterious and age-old questions about the purpose of our existence and on the other hand there is the place where those questions are answered, made redundant, or even deemed irrelevant. While this latter place, that is the utopian sublime state, cannot ever be fully reached;

> …the crux of these myths is that at least there appears to be a road that can be travelled. No matter how unreal or irrational it may seem to actually arrive at the other end of the two-pole system, necessary fictions tell us it can be done. (De Vries 2012, 164)

If we ever would like to reach the communications sublime we should follow our guide NTT DoCoMo. Walking along the road oriented to a future sublime means that our idea of ideal communication is getting closer, future equals progress. Myths are important considering their part in upholding this sublime regulative necessary fictions guide us to. They constitute human hope and search for purpose and meaning of life (De Vries 2012, 16). De Vries refers to Vincent Mosco (2004) when he addresses the myth:

> Useful as it is to recognize the lie in the myth, it is important to state at the outset that myths mean more than falsehoods or cons; indeed, they matter greatly. Myths are stories that animate individuals and societies by providing paths to transcendence that lift people out of the banality of everyday life. They offer an entrance to another reality, a reality once characterized by the promise of the sublime. (Mosco 2004, 3)

A myth allows us to experience something of 'another world' which in case of utopia is aimed at perfection. We are allowed to experience something of this perfection while perfection actually remains out of range and untouchable.

In *The Internet Imaginaire* (2007) Patrice Flichy clarifies the concept of the myth with reference to Roland Barthes' semiology; "the myth is a metalanguage that takes an existing sign as signifier and gives it another meaning as signified" (Flichy 2007, 7). Myths are based on real events and are distinct from ideology as they do not hide the real, rather myths deform the real (Ibid. 8). Flichy addresses Howard Rheingold's *The Virtual Community* to exemplify what he means with a myth. According to Flichy this book can be regarded as one of the founding myths of the Internet. While Rheingold's experiences at the electronic community, The Well, were actually little more than a social experiment, the author took the sociotechnical frame of the electronic community and invisible college and placed them in the different sphere of ordinary sociability. Short after, The Well became the reference model for the internet and became the prime example of a virtual community for those who wanted to launch out in the Internet business (Ibid. 90-1). This shows how a particular social context that made the experimentation possible is forgotten and subsequently how a "local technique" is presented as the "basic technique of a new social functioning" (Ibid. 11). The shift from a local technical experiment becoming a basic technique for a new social functioning is performed by the myth.

Mosco's and Flichy's conception of the myth are complementary to each other. Mosco shows that myths should not be regarded simply as lies or fiction, instead they offer an entrance to another reality characterised by the promise of the sublime. Flichy emphasizes the link between myths and (existing) reality. It is here, reality, where both Flichy's and Mosco's definitions of the myth coalesce. Myths are based on real experiences and do not necessarily hide the real. Powerful myths deform reality, offer entrance to another reality, a better reality, and as such they uphold the belief that is created through necessary fictions that we can reach utopia.

Finally myths might transform a utopia into what Flichy defines a *mask ideology*, when projects turn into a mask ideology aspects of reality are concealed in order to promote the new technology (Flichy 2007, 11). When such promotion is done successfully and people

comprehend a new technology to be offering complete solutions for society's problems, we may speak of what new media scholars define to be a *technological imaginary*. This term claims its origins from Jacques Lacan's psychoanalytical theories, referring to the *imaginaire* (Lister et al. 2009, 66). In *New Media: A Critical Introduction* Lister et al. (2009) describes the imaginaire within Lacan's psychoanalytical theories as referring to images, representations, ideas and intuitions that fulfil our desire as being incomplete and fragmented humans. Lacan's notion of the imaginaire is related primarily to the individual.

Flichy conceptualized the term to denote a 'collective imagination' shaped by ideas; utopias, myths and ideologies such as the immediacy discourse concerning the socio-technological construction of a new medium (cf. Fichy 1999, 34). The technological imaginary draws attention towards the way with which dissatisfactions with social reality and desires for a better society are projected onto technologies (Lister et al. 2009, 67). Just like Flichy I am concerned with the role of language in shaping a technological imaginary, in this case concerning open data. This research offers a critical understanding to how government policy is potentially shaping society's conception of open data, which language is being used more prominent and how to characterize this particular language pattern. Therefore this research is relevant for society. Scientifically this research is relevant for it extends the existing work of mainly two new media scholars, De Vries and Flichy. By adding ideal information as a concept and through critically investigating open data policies regarding the possibility of shaping an imaginaire this research broadens existing research in the field of new media studies. Until now most research papers stem from scholars with an economic or ICT background, considering open data is conceived to be such a recent new media phenomenon this research from a new media studies perspective is a necessary addition concerning the field of new media studies and inspires more new media scholars to inquire open data.

1.3 Method

To answer the research question different analytical approaches to the study of written texts as a form of discourse analysis have been adopted. Many forms of discourse analysis exist cf. Foucault (1980), Laclau and Mouffe (1985), Norman Fairclough (1995) and Jorgensen and Phillips (2003). Many different research directions acquainted with discourse analysis exist, all having their different conceptions and approaches emphasizing different aspects. As I mentioned within my theoretical approach I decided to align myself to De Vries and Flichy and inquire the role of language in shaping an imaginaire characterised through ideal information. In order to research the open data discourse I rely on two research methods that

relate to discourse analytical methodology: textual content analysis and metaphor analysis. Both methods will be further outlined in the following paragraphs, after which I will finalize this introduction through a brief explanation of this thesis' structure.

1.3.1 Textual content analysis

This thesis aims to shed light on the discourse surrounding open data in relation to Dutch governmental organisations (from this point on: government).[2] By investigating policy documents on open data the objective is to understand the 'common' presuppositions regarding open data within these open data policies, which thereafter will be inquired regarding what arguments are underpinned by empirical evidence. To find out which presuppositions with regards to open data are most common I rely on textual content analysis. This method will be used because this is "a research technique for the objective, systematic and quantitative description of the manifest content of communication" (Berelson 1952, 147 in Deacon et al. 1999, 115). Although we can always pose the question to which extent a method is objective, content analysis is well suited to "quantify salient and manifest features" of texts by statistics that "are used to make broader inferences about the processes and politics of representation" (Deacon et al. 1999, 116).

Coding frame

Content analysis is a method that aims to produce a "big picture" which "measures" the degree to which the total body of texts are "slanted" towards a particular perspective (Deacon et al. 1999, 114). In order to systematise data and categories a *coding frame* should be developed. The coding frame can be understood as a methodological tool to systematically count and categorise manifest features within the texts that have been identified in accordance with the determined research objectives (Ibid. 120-1). The coding frame enables the quantification of identified content to be structured according to (pre)-determined categories in order to reveal patterns of which topics are dominating the open data policies.

Utilising the coding frame, presuppositions (benefits, drawbacks, requirements and barriers) about open data as manifest in government policy documents will be elucidated as presuppositions will be distributed according to categories. The coding frame helps shaping the 'big picture' and elucidates which presuppositions (categories) are most common.

After this coding work, the arguments and claims of empirical data that underpin presuppositions will be further inquired. Claims will be juxtaposed and compared to academic studies and other empirical sources. This move not just opens up the possibility to

[2] Within this thesis the words 'governmental organisations' refer to *Dutch* government/ governmental organisations. Any exception will be explicitly mentioned to prevent semantic confusion.

assess the legitimacy of these presuppositions but also functions as a first step in identifying a possible 'pattern' of underlying ideas of idealised open data. Ultimately the method is quantitative but goes beyond counting coded content and involves context in order to unveil a pattern previously concealed. Decisions related to the identification and quantification of salient features of text (presuppositions) and the formation of categories are integral to the analytical process of content analysis. Hence the construction of the coding frame and the development of categories will be explicated in chapter two.

Selection of policy documents

The corpus of research material consists of five (textual) policy documents that stem from governmental institutions.[3] These documents are government reports that present concrete plans of activity and/ or a vision with regards to the construction and sustainment of an open data policy. At the website of the national government, a search for national policy documents on "open data" resulted in 41 hits. Yet, only one document actually related to open data and functioned as a policy document, the Digitale Agenda.nl published by the Ministry of Economic Affairs (2011). Therefore I searched for more documents to include into the sample to get a more representative sample in terms of different government levels and departments. I broadened my search criteria and managed to find two more documents on the website of the national government itself, Actieplan Open Overheid (AOO) and Visie Open Overheid (VOO). Not yet satisfied I decided to use Google's search engine and typed in the keywords "open data beleid" and "open data beleidsnota". This yielded enormous search results varying from 10.400 to 229.000 links. Most links referred to open data portals or open data related websites, open data research reports and popular as well as academic articles related mainly to open data. Nevertheless several policy documents also showed up, and I decided to include two more documents into the sample.[4]

Finally, of five policy documents, two documents originate from provincial and semi-governmental institutions; i.e. the province of Noord-Holland, the IPO and the

[3] The documents included in the sample are: 'Noord-Holland geeft open data de ruimte' (waag society 2012); Digitale Agenda.nl ICT voor innovatie en economische groei' (Ministerie van Economische Zaken, Landbouw en Innovatie 2011); 'Actieplan Open overheid' (Ministerie van Binnenlandse Zaken en Koninkrijksrelaties 2013); 'Visie Open Overheid' (Ministerie van Binnenlandse Zaken en Koninkrijksrelaties 2013); & 'Rapport Waterschappen en Open Databeleid - een afwegingskader' (Informatiehuis Water 2013).

[4] The first document that I included appeared on page one, a report published by Informatiehuis Water (IHW) Rapport Waterschappen en Open Databeleid (2013). IHW represents a collaboration of the state, the Interprovincial Agreement (IPO) and the 'water authorities' (Waterschappen). Secondly I included a policy document, published by Waag Society that popped up on search page two, Noord-Holland Geeft Open Data de Ruimte (2012). This document concerns the open data policy of the province of Noord-Holland and completes the corpus.

Waterschappen. The other three documents are published by two ministries that have most influence on the development of open data policy within national government according to Forum Standaardisatie (the Ministries of Economic Affairs and the Interior).[5] As the five policy documents represent different levels of government authority, different departments and different agenda's this small sample is assumed to provide a broad enough ranger of governance perspectives.

1.3.2 Metaphor analysis

In chapter three I will analyse the concept of open data by means of analysing its accompanying metaphors. Open data policy documents, press releases, popular and academic articles about open data, all seemed to be full of metaphor. By means of metaphor analysis I will aim to provide an explorative understanding of how the concept of open data is framed.

Two different kinds of metaphor analysis are mobilised to analyse the metaphors that appear in policy documents; first the theory which is known as the *Conceptual Theory of Metaphor* (CTM) developed by Lakoff and Johnson (1980). This theory is widely considered as the dominant paradigm in metaphor analysis. In this research the CTM is extended through mobilising the notion of discourse metaphor (cf. Van den Boomen 2005, 2014; Nerlich & Jaspal 2012). Both theories will be briefly explicated below.

Conceptual Theory of Metaphor

Often metaphors are regarded as just a figure of speech, rather innocent rhetorical ornaments; a matter of language instead of action and thoughts. In *Metaphors we live by* written by George Lakoff and Mark Johnson, this conception of the metaphor is highly contested. According to them our human conceptual system is metaphorically structured and they therefore defined metaphors as pervasive and ubiquitous within our everyday life in thought and action (Lakoff & Johnson 1980, 3-6).

To illustrate the pervasiveness of metaphors in everyday life Lakoff and Johnson refer to many examples of mundane speech; for instance in modern western cultures we often make use of the metaphor ARGUMENT IS WAR. In utterances like, "she *attacked* every *weak point* in his argument" and "he changed *strategy* to *win* the argument" this metaphor is prevalent. And as we speak of our income that *rose* or when the Dutch government decides to restrict pubs selling beer to people who are *under* eighteen years old, we use metaphors that adhere to the conceptual metaphor MORE IS UP and LESS IS DOWN (Ibid. 4).

[5] See the following webpage Forum Standaardisatie, http://www.forumstandaardisatie.nl/themas/open-data/.

Lakoff and Johnson (1980) explain that in a metaphor two domains are connected; a target domain X constituted by the immediate subject matter (ARGUMENT and MORE) connects with the source domain Y, in which important metaphorical reasoning takes place and that provides the source concepts used in that reasoning (WAR and UP) (Lakoff & Johnson 1980, 265). A metaphor allows us to comprehend one or more aspects of a concept, the target, in terms of another, the source; ARGUMENT in terms of WAR and MORE in terms of UP, or OPEN DATA in terms of an UNTAPPED GOLDMINE. In other words, metaphors provide means to express abstract concepts into more concrete familiar forms (Van den Boomen 2014, 44).

CTM defines metaphors as cross-domain mappings, as sets of conceptual correspondences across conceptual domains "creating a bridge between a source domain from which a concept is drawn and a target domain to which it is mapped" (Puschmann & Burgess 2014, 1696). Not all elements from the source and the target are used in the metaphor, a selection suffices, and non-used aspects are downplayed (Van den Boomen 2014, 44). For instance, when the concept of open data is being described with the notion of an untapped goldmine certain aspects of an untapped goldmine will be ascribed to the concept of open data. Gold is a precious metal, relate it to luxurious jewellery or perhaps to winning a gold medal at major sports event such as the Olympic Games. Gold has connotations with luxury, success and precious value. Discovering an *untapped* gold*mine* therefore would mean that you discovered a great source of precious value that through fabrication can be used for various applications and that, maybe more important, it will deliver great wealth and success. The metaphor thus connects the target domain OPEN DATA with a valuable (re)source, not yet discovered or at least untapped. In doing this, *they hide or downplay other aspects* of the concept. The goldmine metaphor for instance hides the plurality of data. Where gold is more or less a singular entity in that it is one and the same metal, data(sets) on the contrary are by virtue plural and might differ significantly from one another with respect to their content, formats and subsequently their onto- epistemological being. Far from merely being an innocent rhetoric ornament, the goldmine metaphor literally assigns meaning to the concept of open data and acts as

> [a] guide for future action. Such actions will, of course, fit the metaphor. This will, in turn, reinforce the power of the metaphor to make experience coherent. In this sense metaphors can be self-fulfilling prophecies. (Lakoff & Johnson 1980, 156)

Metaphors frame the way we think, talk and act about open data. They emphasize particular aspects while downplaying others and by doing this the metaphors used in governmental

open data policies inform how others, politicians, journalists and the wider public, understand the concept of open data and act upon it; with a possibility that many elements have been downplayed while other elements prevail. For instance, elements that frame open data to be the next logical step to take in order to create a more successful society.

Discourse Metaphors

Lakoff and Johnson's theory of metaphor is based on a notion of conceptual transference. According to their conceptual or cognitive theory of metaphor, conceptual metaphors share similarities across languages by virtue of being governed underlying cognitive principles that relate embodied human experience and abstractions by means of analogy (Puschmann & Burgess 2014, 1696). In 'Metaphors of Big Data' Cornelius Puschmann and Jean Burgess observe that Lakoff and Johnson discuss metaphors first in relation to the mind. Therefore CTM leaves a gap that concerns the question of agency in how metaphors are strategically chosen (Ibid.). This has also been noticed by Marianne van den Boomen. In 'Networking by Metaphors' she concurs with Iina Hellsten (2002) that CTM is too confined to the individual mind, too quasi-universal and too static. Van den Boomen argues that conceptual metaphors never operate in a social or cultural vacuum;

> No matter how determined they are by human sensory-motor experiencing or embodied in material-semiotic machinery, metaphors – as rhetoric devices, as ordinary expressions, as unconscious concepts, and as material switches – operate in discursive formations regulating networks of power, truth and knowledge (Foucault 1980). Metaphors are embedded in discursive formations, and thereby constitutive of certain views of the world, of society, of how things work. *(Van den Boomen 2006, 12)*

CTM falls short in this regard since it "does not provide a vocabulary which can give an account of cultural differences in concept interpretation and mapping, neither of social transformations in metaphorical concepts over time" (Hellsten 2002 in Van den Boomen 2006, 12). Instead Van den Boomen proposes the notion of discourse metaphors which has been addressed by Jörg Zinken, Iina Hellsten and Brigitte Nerlich (2008).

Zinken, Hellsten and Nerlich (2008) define discourse metaphors as relatively stable metaphorical mappings that are considered to be key framing devices within a particular discourse over a certain period of time. This notion acknowledges that metaphors are never innocent and neutral ways of expression, but neither do they support dominant discourses unambiguously. According to Van den Boomen,

> Discourse metaphors frame and organize shared narratives (be it in the form of public opinion, political agendas, research programs, or world views), but most of all they organize and install standards, rules, norms, and procedures – in short, material discursive formations of power, truth, and knowledge. *(Van den Boomen 2014, 78)*

CTM considers metaphors as part of our conceptual human system, framing cognitive assemblages of thought, articulation and action. The notion of discourse metaphors on the contrary acknowledge that metaphors might be taken up, negotiated, rejected or reformulated; they might be elaborated, criticised and redesigned in order to achieve particular goals, by framing social assemblages of thought, articulation and action (Van den Boomen 2006, 12).

By moving beyond Lakoff and Johnson's conceptual metaphors, through the extension of discourse metaphors, I will aim to provide an explorative mapping that concern the discursive politics of 'metaphors that open data' as manifest in Dutch governmental discourse.[6] In order to find out to what extent discourse metaphors are at stake, the corpus material will be extended with various press releases, articles and texts published by key advocates of open government data within the Netherlands of which Neelie Kroes is one. I will address several discourse metaphors that appear to be most salient within open data policies of Dutch governmental organisations: Open data as a natural resource, a feedstock and raw material, and finally I will discuss the conduit and container metaphor in relation to open data.

1.4 Structure

In chapter two I will first construct the coding frame necessary for the textual content analysis related to five policy documents of national and local Dutch governmental institutions. The coding frame tabulates benefits and barriers as mentioned in the corpus. Secondly this chapter compares the claims within the policy documents with academic studies. As will become clear, government policies lack empirical proof and their value claims, that contain an urge of progress, are not enough aware of potential contradictory implications concerning open data. They are underpinned by what I call ideal information instead.

Chapter three examines the metaphors that surround open data. I agree with Brigitte Nerlich and Rusi Jaspal (2012) as they argue that metaphors are the mind's eyes and society's tools. It will be argued that the metaphorical framing, in which an urge of progress and

[6] This shortcoming becomes salient when Lakoff and Johnson discuss Michael Reddy's description of the conduit metaphor, a metaphor that according to Reddy plays in important part in the way our language about language is structured. Whereas Reddy proposed an alternative metaphor, the toolmakers paradigm, Lakoff and Johnson only address the former. This has been noticed by Marianne van den Boomen: "By ignoring the alternative proposal they also ignore discursive competition between metaphorical frames, i.e. the politics of metaphor" (Van den Boomen 2014, 94).

wealth is projected on open data, potentially results in what earlier accounts of new media studies have identified as technological imaginaries (Patrice Flichy 2007).

The urge of progress that characterizes the discursive framing, as revealed through textual content analysis and metaphor analysis, will be further elucidated, discussed and criticised in chapter four. This concluding chapter is aimed to discuss and extend the theoretical implications concerning the discourse of open government data in relation to an *ideal* of *information* pointing towards INFOTOPIA.

Chapter 2: Unravelling Open Data Policies

The first section of this chapter explains in great detail the implementation of textual content analysis within my research. This includes the construction of a coding frame necessary to range and sort different indicators according to various categories and varying levels. Section 2.2 and 2.3 present the results of the textual content analysis and give an overview of benefits and barriers regarding open data that have been traced throughout the five policy documents. In section 2.4 I will discuss the problematic implications of the presented benefits and drawbacks of current open data policies. Finally in section 2.5, by turning back to a definition of data, information and knowledge, I will argue that current policies are a result of what I call a pattern of ideal information.

2.1 Constructing the coding frame

Textual content analysis relates to theme analysis, which means that the coder has to recognise certain themes or ideas in a text which from then on have to be allocated to predetermined categories (Beardsworth 1980 in Deacon et al. 1999, 119). Every document has been read and analysed according to predetermined research objectives. These objectives have been translated into relevant questions in order to identify salient pieces of content. In deciding what to 'count', I decided to ask the following questions from a government perspective: *Why should we start with an open data policy? According to what benefits? And, how are these benefits made possible?* The categorisation, by means of the aforementioned questions, resulted in a 'three level structure' surrounding benefits and requirements. I will use an example from the coding frame (see appendix) to clarify this structure, based on the following quote:

> Bovendien laat de overheid aan burgers zien welke gegevens worden gebruikt in welke processen en geeft hen de gelegenheid deze gegevens te corrigeren wanneer nodig. (Actieplan Open Overheid 2013, 15) [7]

This quote concerns a document by the ministry of the Interior who argues open data enables people to review and adjust data and therefore benefits a "doe-democratie" which translates into DIY Democracy or participatory democracy (Ibid. 14). This policy document ascribes a certain democratic/political value to open data. The political/ democratic value is categorised as an *overarching value* (first level, overarching) that includes various other benefits which

[7] A literal translation of the quote: "Moreover, the government allows citizens to see what data are used in relation to which processes and gives them the opportunity to correct the data if necessary".

are second level.[8] In this case the quote explicitly refers to the possibility for citizens to review and correct government data because of open data, which has been identified as leading to an increase of democratic accountability, a *benefit* (second level, beneficial). Finally, democratic accountability is obtained through more transparency and equality as the government exposes their data and enables citizens to act on the same data government. Transparency and equality therefore function as *requirements* (third level, conditional) to assure democratic accountability. Summarized, the presupposition that the concept of open data has a (1) *democratic and political value* is substantiated through the claim that open data will lead to an increase of (2) *democratic accountability* which is achieved through (3) *transparency and equality*.

The observant reader might have noticed that I didn't mention barriers and drawbacks yet as part of constructing the coding frame. As a researcher you want to avoid walking around with blinders on. Sometimes you are so focused on discovering a particular kind of pattern that you tend to have only eyes for one part of the subject and forget the other. A similar fallacy momentarily took hold of me causing me to forget to take into consideration the drawbacks and only look and count benefits. 'Benefits, Adoption Barriers and Myths of Open Data and Open Government' by Marijn Janssen, Yannis Charalabidis and Anneke Zuiderwijk (2012) reminded me that barriers and drawbacks too are part of government policies.[9] Ignoring these would create a distorted picture of government open data policy in which open data is plain beneficial and has no (potential) limitations.

Re-constructing the coding frame

The categories identified by Janssen et al. (2012) resemble many of the categories identified by me. In case of difference this often seemed to be a matter of sheer nomenclature. Therefore I decided to adopt most of their nomenclature for benefits as well as barriers and drawbacks in case I thought it would be of relevance. For example "accountability" has been adjusted to "democratic accountability" emphasizing government's aim for a more democratic society through open data. A category I didn't include at first: "Development of new products and services" was added into the coding frame alongside innovation for cases

[8] The other overarching categories that have been recognized are; positivistic notion, economic value, governmental value, societal value, technical value and environmental value. They are delineated as the analysis progresses throughout this chapter.

[9] There are many parallels between their research and this study as Janssen et al. studied benefits and adoption barriers of OGD and related these to myths after juxtaposing them to empirical data. They analysed adoption barriers and benefits of open data through an organized group session and interviews with 14 key persons represented from different organisations including Ministry of Economic affairs, Hack de Overheid and Ministry of Interior.

that specifically refer to the development of applications and services. Innovation as a category is broader and could also refer to new innovative insight.

A major difference between Janssen et al. and my categorisation is that, contrary to the three-level-structure I constructed, Janssen et al. (2012) developed a coding-frame consisting of two levels; (1) overarching categories and (2) benefits or barriers/drawbacks which link to these overarching categories. During the coding process I experienced that most, if not all, statements identified on a conditional level could as well be identified on a beneficial level. For instance, transparency could be identified as a requirement enhancing democratic accountability. However, it could also be understood to a be beneficial in itself since it is often presented as a direct solution to reach a 'more' democratic and 'equal' society (Fung 2013, 185). This overlap hasn't been exceptional and therefore I decided to efface the hard line separating benefits from requirements and partly reconstruct the coding frame. All indicators previously separated according to both levels and structured underneath the overarching categories are now accumulated together.[10]

To prevent more of the above coding difficulties, I decided to use a two level structure regarding *barriers and drawbacks* thereby resembling the structure of Janssen et al. (2012). Most of the nomenclature developed by Janssen et al. was adopted and adjusted or complemented in case necessary. Whereas Janssen et al. (2012) identified barriers concerning "information quality" such as "lack of information" I also developed a category "Data Quality" with a barrier that concerns the "lack of data" since data and information are two different concepts. Notwithstanding the fact that in popular as well as scientific texts and debates these terms often conflate and are used indiscriminately. Without proper definitions both terms are often equated to knowledge. In section 2.5 I exemplify how data, information and knowledge often are conflated in government policy and extrapolate on how I argue this results in what I have identified to be a pattern ideal information. Section 2.2 presents an overview of the benefits of open data that have been traced in the five policy documents by means of the coding frame.

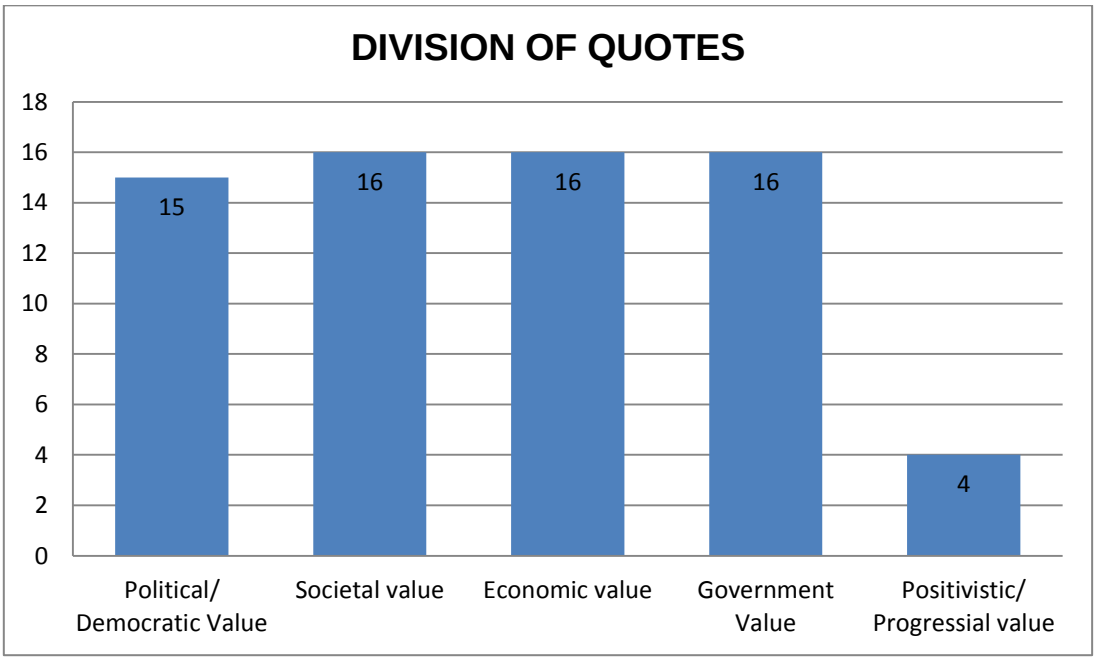
2.2 Open data benefits

If we should believe the government policy documents, open data generate value, even more than when selling these data sets. The opening up of non-personal (public) data in machine-readable formats under such a licence that allows both commercial and non-commercial use, re-use and distribution by anyone, a move towards an open data policy, yields numerous

[10] The coding-frame document is included in the appendix. A separate excel document can be downloaded at this website: siemvanboxtel.wordpress.com

benefits. The benefits assumed by governments were clustered into five overarching categories (1) Political and Democratic value, (2) Societal value, (3) Economic value, (4) Governmental value and finally a Positivistic cluster (5). This fifth category includes utterances of the government referring to certain 'positive effects' or 'progressive value' of open data without further clarification of what exactly these positive effects are or in what respect competitiveness is improved by open data. In total 67 bits of content have been identified and categorised according to the overarching values (figure 1).

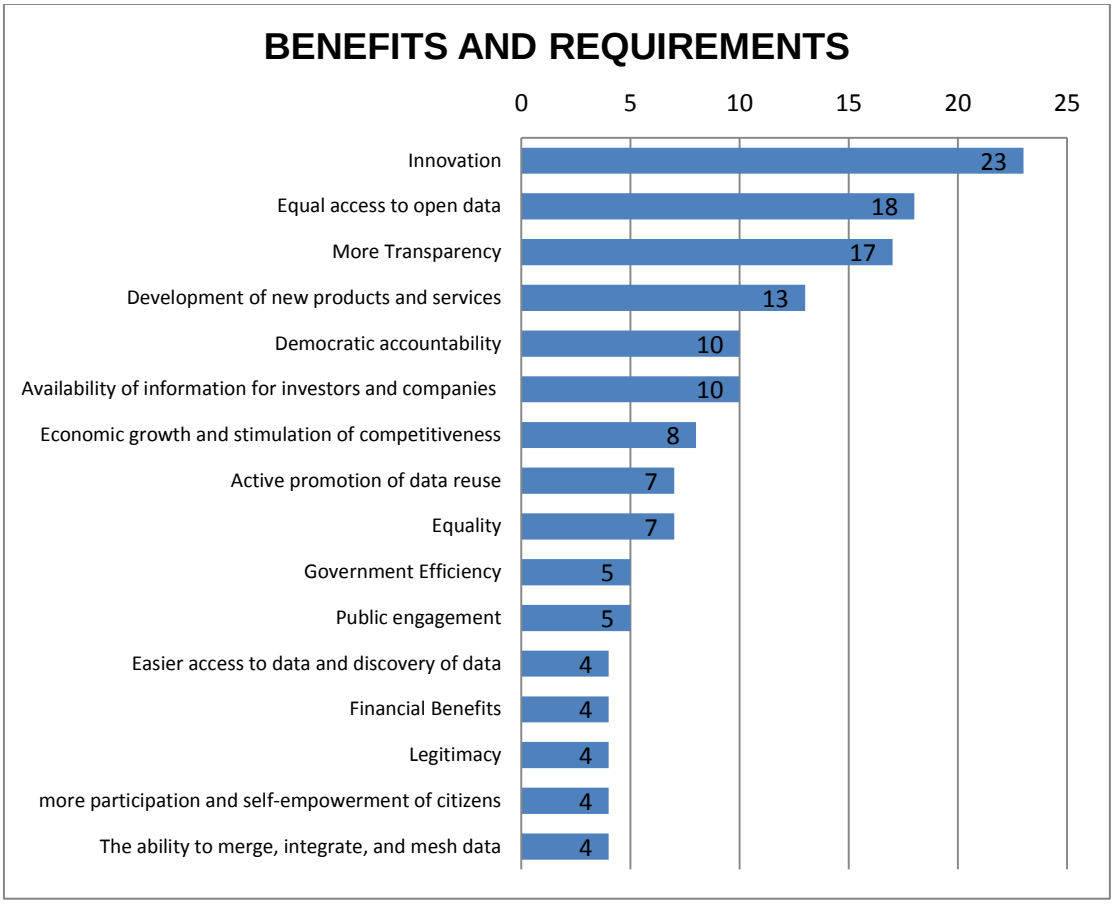
In the same order, 15 quotes are classified underneath the political and democratic cluster, 16 quotes in the societal cluster. Another 16 in the economic cluster and again 16 bit of content refer to the governmental value of open data. Only 4 quotes relate to the positivistic cluster.



Figuur 1: 67 quotes ranged according to five overarching values

Before I effaced the categorical separation between benefits and requirements I identified 25 different benefits that occur 83 times identified underneath the 67 quotes. Accidentally the number of requirements is exactly the same as the number of benefits; I identified 25 different requirements occurring 104 times. Altogether there are 42 different benefits/requirements. Eight indicators directly overlap. In total the 42 benefits[11] have been quantified 187 times (figure 2), referring to 67 quotes structured according to five overarching values.

[11] With benefits I mean both benefits (2nd level) and requirements (3rd level).



Figuur 2: A selection of the most salient benefits and requirements (16 out of 42).

2.2.1 Political and democratic value

Every policy document except for the 'Digitale Agenda'[12] mentions open data in direct relation with its political and democratic value. According to those documents open data enhance, ensure and sustain (modern) democracy. Fifteen quotes have been identified expressing political and democratic value in relation to open data. Eleven different benefits have been identified in relation to these fifteen quotes and are present 45 times. Of all eleven benefits four stand out and are quantified 34 times. This means that 34% account for 76% of the total 45 quantified benefits. On a political level the concept of open data is beneficial mainly because it enhances transparency (12), democratic accountability (9), equality (7) and equal access to open data (6).

According to 'Rapport Waterschappen en Open Databeleid' (hereafter RWOD) equal access to (open) data on an egalitarian basis is important since it enhances the transparency of governmental policy and its processes. Democratic decision making processes are supported

[12] The digital agenda is published by the ministry of Economic Affairs and focused on possibilities to ensure innovation and economic growth through ICT of which open data is part of.

because the same information is accessible to both governmental organisations and the public (RWOD 2013, 6).[13] 'Visie Open Overheid' (hereafter VOO) argues that transparency is a conditio sine qua non in order to ensure and sustain trust in government from the public;

> Transparantie en openheid van de overheid is een voorwaarde voor burgers om de overheid en haar activiteiten in de publieke sector te kunnen bewaken en controleren (VOO 2013, 9).

Transparency is a core value of a modern democracy that enables citizens to monitor and control government decision making processes and activities (Ibid. 9). So while transparency and equality are considered inherently beneficial they also enhance democratic accountability. Open data not just open up government policy but empowers those who were previously excluded or significantly less empowered, concerning political decision making. It enables citizens to take part in politics and policy making on an "egalitarian information position" (AOO 2013, 13). A participatory democracy is enhanced in which the government apparatus is organized less hierarchical and more "horizontal" which fits today's "network society" that according to 'Boom en het Rizoom' (2012) is presumed to be organized horizontally (in VOO 2013, 9). To exemplify the participatory democratic power of open data, 'Noord-Holland Geeft Open Data de Ruimte' (hereafter NHGODR) refers to the Danish Folketsting.dk and the Dutch variants argumenten.nl and politiekinzicht.nl. These are all websites created through using open data sets. The websites allow citizens to see which decisions are taken, based upon which data (sets) and information and which members of parliament are involved and make it easier for citizens to get informed and involved in politics whereas previously they were excluded and could not access this sort of information (NHGODR 2012, 8). Such examples imply a certain participatory democratic value of open data, but actually the examples refer to processed data. Nevertheless, processed or raw, the reasoning is that these innovative democracy enhancing applications would not have been created and benefits would not have occurred if data have not been opened up.

2.2.2 Societal value

Where transparency, equality and democratic accountability are important benefits of open data, creating a more participatory democracy it becomes clear that the argument concerning innovation is at least equally important in promoting open data. Open data stimulates innovation which goes beyond merely the political domain and appears most notably

[13] Quote translated from Dutch to English: "Een laatste belangrijke baat is dat het openstellen van data de transparantie van het beleid vergroot, immers zowel het publiek als de waterbeheerder kan over dezelfde informatie beschikken. Op deze manier worden democratische besluitvormingsprocessen met betrekking tot waterbeheer ondersteund" (Rapport Waterschappen en Open Databeleid 2013, 6).

contributing to the societal and economic value of open data. In the coding frame twelve different benefits are ranked among societal value of open data, quantified 43 times. The most salient benefits include; innovation (11), equal access to open data (10), the development of new products and services (5) and public engagement (4). 33% of the identified benefits account for 70% of all benefits accumulated regarding open data's societal value.

Government's policy documents argue that opening up data stimulates innovation and yield new products and services that benefit society; "Bovendien opent de vrije beschikbaarheid van geodata de mogelijkheid te komen tot andere toepassingen" (Digitale Agenda.nl 2011, 18). Earlier studies concerning open data and government policy concluded innovation to be one of the key themes that informs policy making (cf. Davies 2010, 5; Janssen et al. 2012, 260). Therefore it is not much of a surprise that innovation is one of the most salient benefits with regard to open data. Again documents often refer to best practices which serve to exemplify the societal value that comes about through the development of innovative applications, services or even insights. For example 'NHGODR' refers to BuitenBeter;

> Kapotte lantaarnpalen, vieze bushokjes of losliggende stoeptegels: applicaties als BuitenBeter en Verbeter de Buurt bieden de mogelijkheid melding van het probleem te maken. Dergelijke applicaties zouden echter niet mogelijk zijn zonder Open Data. *(NHGODR 2012, 10)*

BuitenBeter.nl facilitates the public to cooperate with governments or other organisations in addressing and solving practical societal issues. Those applications are considered to exemplify the means by which 'ordinary' citizens are getting more empowered with respect to their influence and participation in society.

2.2.3 Economic value

Open data's economic value manifests in every policy document. Sixteen different benefits have been identified and quantified 50 times. Four benefits (25%) stand out and account for almost 70% of all quantified indicators; innovation (12), economic growth and stimulation of competitiveness (8), the development of new products and services (7) and the availability of information for investors and companies (7). The development of innovative applications, services and insights sparks economic growth as creative entrepreneurs are granted access to open data;

> Binnen de overheid zijn veel data beschikbaar die relevant kunnen zijn voor ondernemers om te worden hergebruikt. Verschillende studies

> laten zien dat de economische waarde van de innovatie en bedrijvigheid die ontstaat op basis van open data groot is. Met open data als grondstof kunnen nieuwe toepassingen en diensten ontwikkeld en vermarkt worden. *(Digitale Agenda.nl 2011, 14)*

Whereas equal access to open data previously was considered vital in order for open data to be of political and societal value concerning the wider public, the focus regarding open data's economic value is clearly on the availability of information for investors and companies, creative entrepreneurs that are able to innovate through merging, integrating and meshing data with other data. Entrepreneurs should be able to grow and open data is presented as the "soil" entrepreneurs can use to achieve this growth and stimulate competitiveness (Digitale Agenda.nl 2011, 14; AOO 2013, 22).

Government policies claim that open data are of great economic value which is estimated to be millions of euros each year. 'NHGODR' refers to a macro-economic study conducted on behalf of the European Commission that estimates the economic value of open data to be 140 billion euro each year for the EU. At the same time, though, governments acknowledge that the value of open data remains unclear. Open data have no real value in itself. Data only become valuable when used. Making it hard to predict potential applications, possible "killer" applications are even more difficult to predict (Janssen et al. 2012, 260).

In section 2.4 I will discuss this issue in relation to other questionable and/or problematic implications within these government policy documents. First we will continue further delineating the last two categories, as well as the drawbacks and barriers in the next section 2.3.

2.2.4 Governmental value

Governments consider open data to be of economic, societal and political/democratic value, but open data is also considered valuable for government itself; government institutional and operational value. The governmental value of open data most notably is discussed in two policy documents: 'RWOD' and 'NHGODR'.[14]

One of the important benefits of open data is that it might lead to an improvement of government efficiency. Firstly, the opening up of government data according to *open standards* makes it easier for governments to interoperate internally between multiple departments of government as well as external of government. According to the 'Open Data

[14] An explanation could be that compared to the other policy documents which concern national politics, the documents published by the Water Authorities and the province Noord-Holland still have to convince civil officials. Whereas 'Visie Open Overheid' and 'Actieplan Open Overheid' more or less regard open data inherently of value with regards to government. The 'Digital Agenda' furthermore isn't aimed on governmental value, but merely on societal and economic value of ICT.

Handbook' (ODH) open standards entails the publication of data free of license restrictions in formats that are machine readable; and ensure data are accessible and reusable for multiple purposes (ODH 2014). Secondly, knowing that data are public, there is an outside scrutiny factor which offers government a chance to be more critical at their own methods and spending (VOO 2013, 11). An enhancement of government efficiency could lead to substantial financial benefits, i.e. savings, as well as organizational benefits:

> Open data biedt de kans om een kleinere overheid te realiseren waarbij de markt zelf diensten kan optuigen die goed aansluiten op de informatievragen vanuit de maatschappij zonder dat de overheid hierin hoeft te sturen of financieel hoeft bij te dragen. *(RWOD 2013, 21)*

More participation and self-empowerment of the public creates a situation in which 'the market itself' is increasingly able to establish services that connect to information requests from society without government interference. Hence RWOD argues that open data contribute to a smaller but more efficient and effective government apparatus.

At the same time the legitimacy of government is argued to be enhanced. Previously we saw that transparency is argued to be an essential characteristic to ensure trust in government. Yet openness is not the only by means legitimacy is increased. Open data will lead to an increase of data quality and sustainability of these data. Since data will be used more often, feedback arises through both internal and external quality checks. Future decisions based on the available data will therefore be better substantiated as data is of higher quality. Furthermore, if people are more actively involved in the decision making process, the resulting policies are more widely supported while the government apparatus became smaller since government efficiency is increased. And finally, more transparency and democratic accountability benefits government's legitimacy and thus open data will contribute to the public's satisfaction in government.

2.2.5 Positivism and progress

Lastly I identified four utterances within government policy that on the one hand are aimed on the positive effects yielded through the opening up of data and their reuse (RWOD 2013, 13). The statements imply that a move towards open data is equivalent to making progress. The message is that if we decide not to proceed further with exploiting open data we will soon lag behind. Similarly, connecting with the Open Government Partnership (OGP) in relation to open data is explained to be a sustainment of competitiveness (VOO 2013, 7). The rapid growth of the OGP according to VOO demonstrates that governments all around the world realize they should connect to a movement which progresses towards more

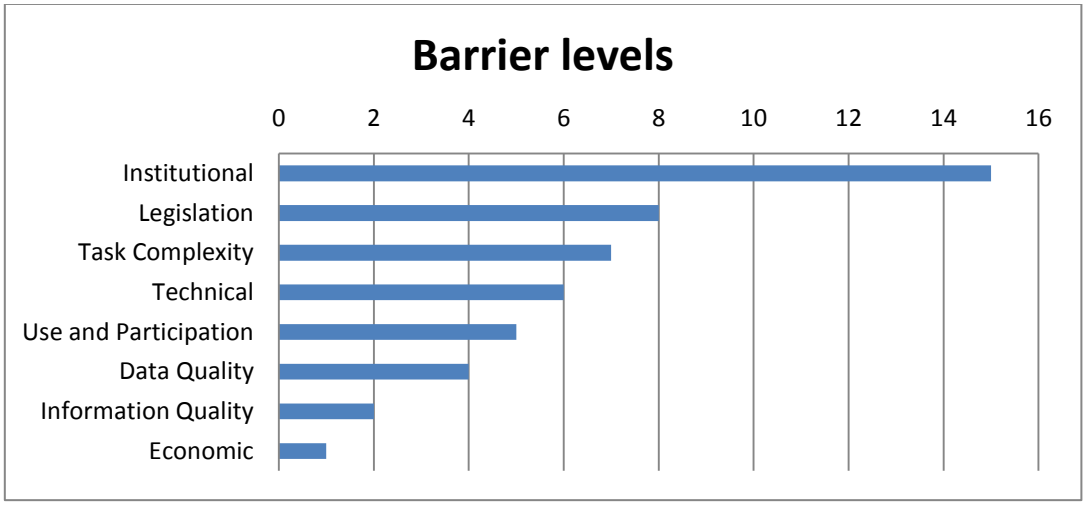
transparency in society. Open data happen to be one of the many means in order to realize more transparency.

In summary, the opening up of government data yields numerous benefits; transparency, accountability, equality, efficiency, legitimacy and innovation. The exact value of open data however remains vague; still this does not withhold governments of adopting open data policies. After all, the ways and motivations for using open data are broad; societal, political, economic, government institutional and operational, and even plain positivistic; for if we do not follow and jump into the open data train now we will lag behind.

2.3 Barriers to overcome

Analysing the policy documents I noticed that the barriers I identified resembled many of the barriers identified by Janssen et al. (2012). Most of the nomenclature and categorical structure therefore stems from their research, complemented with two categories and a few extra barriers. Janssen et al. found that "barriers are related to either data providers (resulting in not wishing to publicize data) or data users (resulting in an inability to use the data in an easy manner)" (Janssen et al. 2012, 261). Barriers have been discerned according to eight different categories. The barriers on an *Institutional*, *Economic* and *Legislative level* concern barriers from a data provider's perspective. Barriers on the levels of *Task complexity* and *Use and participation* concern the user's perspective. And barriers that concern *Information quality, Data quality* and *the Technical level* are relevant for both user and provider.

The coding frame includes 48 quotes that refer to 30 different barriers, which have been quantified 68 times. The three most important categories are: institutional barriers, legislative barriers and barriers concerning task complexity. Barriers on the institutional level appear most frequently and widespread. 15 quotes referring to institutional barriers appear in four policy documents. The barriers on the legislative level include 8 quotes that appear in 3 documents. On the level of task complexity the barriers are identified 7 times and appear in 4 documents. Figure 3 shows a graph of which barriers appear to be present most frequently.



Figuur 3: Quotes ranged according to different barrier levels

38 quotes of the total 48 quotes that have been identified belong to two documents NHGODR (25) and RWOD (13). This is not much of a surprise since the other three documents include open data to be part of a bigger subject; open government (VOO & AOO) and as means for innovation and economic growth (Digitale Aganda.nl). NHGODR and RWOD specifically aim at open data. They try to persuade internal government officials of open data's usefulness.

It is at the level of the institutional barriers that government policy emphasizes a change of culture in order to accomplish a successful open data policy. According to the policy documents government institutions should be more open to external initiatives and entrepreneurship instead of maintaining the current risk-averse culture (VOO 2013, 9; NHGODR 2012, 19). Such change of culture can only occur if there is sufficient internal support. Employees have to acknowledge the potential value of open data and not be precarious about it. A uniform policy for publicizing and maintaining data has to be created to prevent the publication of data that shouldn't be published. This could entail a loss of government's public image and the creation of public distrust or dissatisfaction regarding government. On the legislative level this could result into privacy violation and possible litigations (RWOD 2013, 19-20).

A change of culture and structures however entail costs which form barriers as well, especially when there is not enough internal support. Costs have to be well contemplated;

> Het realiseren van openheid zal echter ook investeringen en kosten met zich meebrengen die in tijden van bezuinigingen niet vanzelfsprekend zijn. Daarom is het belangrijk om stapsgewijs toe te werken naar meer openheid. *(VOO 2013, 11)*

‘Digitale Agengda.nl’ furthermore asserts that currently many of the published government data aren’t open. The data are difficult to locate and cannot be read by machines (Digitale Agenda.nl 2011, 14). Barriers on the technical level are related to barriers on the level of task complexity and use and participation as well. When you publicize incomplete and insufficiently described datasets, the tasks to retrieve and reuse open data could become more complex. Data remain ‘closed’ and obstruct the potential innovative benefits in case of otherwise opened data to be realized.

The barriers related to *task complexity, use and participation, information quality, data quality* and *technology* show up mostly in NHGODR and RWOD. Both policy documents concern not only visionary plans but refer to concrete steps of how to proceed with open data. In ‘NHGODR’ a list of imperatives act as a guidance on “How to get started with Open Data” and overcome barriers (NHGODR 2012, 17). This list of imperatives implicitly includes barriers on the level of task complexity, information and data quality, use and participation and technology. To briefly summarize the list; before you publish government data you have to make sure each dataset includes a clear description about its original purpose and entails sufficient meta data that clarify; the meaning of columns and rows and the frequency with which the dataset is updated. Data should be published according to open standards such as XML and CSV instead of PDF and XLS without further licence restrictions. Erasing these barriers, makes it easier for potential providers to publish ‘open’ data and for users to actually find, reuse and enrich data. The NHGODR policy document asserts that task complexity and user participation can be enhanced through providing user incentives and establish a support helpdesk in case of questions from users as well as provider (NHGODR 2012, 17-19). The RWOD policy document includes similar barriers, also framed as imperatives on how to overcome or prevent the manifestation of barriers:

> Data die de waterbeheerders kunnen gaan vrijgeven als Open Data is verzameld in het kader van een overheidstaak en dus binnen een context en met een betekenis. Derhalve moet de publicatie van data altijd vergezeld gaan van een beschrijving van deze context en de betekenis. Het meegeven van meta-data zorgt er voor dat de data ook vindbaar is. *(RWOD 2013, 21)*

On the legislative level, the policy documents mainly emphasize the possibility of dispute and litigations due to privacy violation or unrestricted use and/or publication of data. Due to these legislative risks some officials (potential providers) get cold feet to publish government data. Both ‘NHGODR’ and ‘RWOD’ however ascertain that, if enough precautionary measures are

taken risks are limited regarding government liability (NHGODR 2012, 18; RWOD 2013, 21). According to the report 'Aansprakelijkheid en Open Data' by Marc de Vries the risks for government itself however are confined (De Vries 2012 in NHGODR 2012; RWOD 2013). Risks are only analysed directly from a government perspective and not placed in a wider perspective regarding society for instance. The same is true for the publication of data which might lead to unfair competition. Although unfair competition presumably implies a certain economic risk, RWOD merely emphasizes the risks concerning government liability which 'fortunately' are considered small (RWOD 2013, 21).

Compared to benefits there are significantly less barriers present in government policy; there are 187 occurrences of 42 different benefits, and 68 occurrences of 30 different barriers. An analysis of both coding frames (benefits and barriers) reveals that government policy is especially aimed on delineating the benefits and requirements by means the value of open data can be maximally obtained. This is quite remarkable since Janssen et al. (2012) identified 57 different barriers and only 32 benefits. To summarize, governments do acknowledge the existence of potential barriers regarding open data but they are all framed as *barriers that* can and *should be overcome*. Possible legislative risks 'resulting from an open data policy' eventually turned out to be small risks for government itself. Institutional risks such ass loss of public image or public distrust in government can be prevented through drafting up a solid policy that, indeed, entails financial as well as non-financial investments but eventually open data will generate greater benefits (income?). Apart from a few risks, that can be mitigated, there are no drawbacks identified that directly result from open data. The emphasis of government policy thus is on open data's benefits and most barriers mainly act as requirements in order to ensure potential benefits to happen and open data to become truly valuable.

2.4 Implications of open data policies

In the previous two paragraphs benefits and barriers as considered by government policies have been thoroughly delineated using the coding frame. Juxtaposing the benefits of open data to barriers led to the conclusion that benefits by far outweigh the barriers. Government policies speak of the many benefits of open data for society, politics, economy and government itself. Open data is presented as some kind of new technological development or feature of existing technologies which we did not yet exploited and will benefit us in several ways if we do. In this section we will confront these (ideal) assumptions about open data with academic studies and empirical data.

We will discuss questions such as: Which sources (research reports, organisations and advocates) have been consulted and to what extent do these 'count' as evidence? To what extent do anecdotes prove anything and are there anecdotes and examples related to open data that lead to contradictory situations regarding the society we desire it to be? With each step I aim to uncover a bit more of the underlying pattern concerning the concept of open data, and its related *ideal* assumptions, that forms around information and points to a situation where *information* is *pure*, *accessible* and *available for all* (section 2.5).

2.4.1 Where is the evidence?

Government policies acknowledge the value of open data is not exactly clear (NHGODR 2012, 7; RWOD 2013, 22; VOO 2013, 11).[15] Open data is still quite a recent phenomenon, how are we able to know its exact value? An often heard saying regarding new technological developments is; *we certainly did not all foresee the immense growth and usefulness of the internet at the end of the 20th century*. NHGODR asserts; "Het ligt in de aard van innovatie dat ontwikkelingen vaak uit onverwachte hoek komen" (NHGODR 2012, 7). Innovation frequently comes from unexpected sources; it's in the nature of innovation (ibid.). Some policy documents argue this uncertainty might be a potential barrier as government officials could be reluctant to see the potential value in opening up data sets (insufficient internal support) especially when investments have to be made (financial costs). Nonetheless the documents give little room for doubt that open data will eventually generate benefits and is valuable in numerous ways i.e. economic, political, societal, and government institutional. The question, as derived from government policy, does not concern whether or open data generates value, but merely concerns the quantity of this value? The generation of value is taken for granted.

There are so many factors which could differ regarding an open data policy, regarding data itself, suppliers, potential users and the situations in which they all reside. Nevertheless, numerous arguments, examples and benefits delineated in section 2.2 demonstrate how value is assumed to be acquired through open data. It is remarkable how most arguments underpinning the value of open data completely lack empirical evidence. Of all five policy documents only three explicitly refer to external studies or research reports that should account for the empirical evidence underpinning their value claims. VOO and AOO have not mentioned any academic study or research report in combination to underpin the value claims they make. This is not to say their arguments are completely false. It means that

[15] Frequently they refer to the particular economic and innovative value of open data; however this entails other categories as well.

their arguments are not directly underpinned by empirical data. Similarly a reference to external studies does not mean a value claim is sufficiently underpinned empirically.

For instance the NHGODR document included a reference to a report conducted on behalf of the European Commission. This report concerns a macro-economic study that estimates the value of open data to be at least140 billion each year for the EU countries. Huijboom & Vd Broek (2011) noticed that government policies often use such macro-economic studies to justify their open data strategies. It is not my intention to doubt the conducted macro-studies; however I do contend the evidence provided by these macro-economic studies is not sufficient to legitimise an open data policy concerning the economic value of open data;

> [t]he precise economic impact of open data for their country, and specific sectors or organisations, remains largely unclear. Impact studies at both the meso and micro levels are lacking and, since the macro studies use different indicators to estimate the economic impact, the calculations differ substantially. (Uhlir 2009 in Huijboom & Vd Broek 2011, 10)

Janssen et al. agree with Huijboom & Vd Broek (2011) and argue more research into the benefits, barriers and value of open data is necessary (Janssen et al. 2012, 266).

Digitale Agenda.nl (2011) also argues that open data is potentially of great economic value, it sparks innovation and economic activity. This policy document was published the same year Huijboom & Vd Broek presented their conclusions and recommendations. Surprisingly the document even refers to their report when claiming open data is potentially of great economic value. How are we supposed to take their value claims seriously? If Huijboom & Vd Broek argue that "evidence of economic, social and democratic impacts of open data policy is still immature or lacking" and their research team didn't find sound evidence of the impact of open data policy in any of the countries investigated (Huijboom & Vd Broek 2011, 9-10). Both policy documents, Digitale Agenda.nl and NHGODR, neglect the critique stated above and draw conclusions on insufficient empirical evidence. They seem to merely follow a trend set out previously by other organisations from other countries depicting an image of open data causing enormous amounts of economic prosperity and financial benefits as claimed by reliable authoritative sources.

NHGODR and RWOD also name another study in their bibliographies, 'De Waarde van Open Data' (2012) by Kronenburg et al. on behalf of Zenc. In this report seven different open data strategies of Dutch public organisations have been analysed; each case included different datasets incorporating different strategies (Kronenburg et al. 2012, 9). The

complexity of different strategies that concerned each dataset made it impossible for Kronenburg et al. (2012) to discuss open data strategies of organisations in general. It demonstrates the complexity of open data (policies). I admit the report describes several different financial and non-financial benefits depending on each case and dataset. Yet the report also strongly points to potential 'market-distorting' effects caused by potential (re)use resulting in the development of open data applications causing unfair competition and various different interpretations of open data. Kronenburg et al. assert that in practice it is difficult to ascribe costs and benefits to open data strategies (Ibid. 12,31). Contrary to the impression given by government policy, beneficial effects of open data are far from self-evident.

It is too extreme to state that government policy argues benefits will arrive automatically simply through opening up data. Several barriers are mentioned, and the policy documents acknowledge the exact value of opening up data is unclear but government's appear to believe that if data are opened up according to open standards, raw, as complete as possible and supplied with sufficient meta-data and information about the original purpose, if there is sufficient internal support, in short; if we abide by the rules and draft a sound policy, *if we eliminate the ifs*, barriers on the supply side and on the users' side will be eliminated. Such an elimination of barriers paves the way for benefits to happen.

2.4.2 A double edged sword

So far we have discussed the lack of empirical evidence mainly of open data's potential economic value considered by government policy of which innovation is one of the most important drivers. The claim of open data's economic value within government policy is legitimized through referring to macro-economic studies that claim to estimate the economic value generated through open data to be many billion(s) of euros each year. Indeed, these studies might not be suffice to actually prove the economic value of open data, however the fact is they do. Government policy does create an image of wealth creation through open data based on external scientific authorities that, presumably by vast majority, count as evidence.
Claims about societal, political/ democratic and institutional beneficial effects of open data are much more difficult to justify in this way. In these domains "even less evidence is available" (Huijboom & Vd Broek 2012, 10). Such value claims therefore largely depend on examples of open data applications and anecdotes instead of empirical studies.

VOO refers to open data applications in the US and UK that increase transparency of government spending subsequently enabling citizens to monitor government spending and

check if spending is conform government policy (VOO 2013, 11).[16] NHGODR mentions the yet out dated application *www.schoolvinder.nl* comparable to *www.10000scholen.nl*. These websites provide information about schools and allow citizens to compare different schools using various information criteria derived from the data, for instance the drop-out rate, average grades and test scores compared to the nationwide average, the percentage of students that live in a poor neighbourhood and so on. Government policy explains these apps to be plain beneficial, demonstrating that innovation not just leads to economic value, but as well is able to enhance democracy and be of societal value. The examples demonstrate that citizens are empowered because it is easier for them to be informed and participate in politics, and to allow better opportunities for instance for deciding a future school. Previously this information was closed and citizens were excluded from this particular information, obstructing the creation of such applications and obstructing the benefits from happening.

The examples above can also be viewed from a different perspective. In 'Open Government and Open Society' Archon Fung and David Weil warn for efforts that mostly focus on creating an "open government" (Fung & Weil 2010, 106). Especially those who tend to solely focus on government accountability, such as applications exposing government spending. For instance the web application wheredoesmymoneygo.org visualizes government spending in the UK, but it is difficult to get a grasp on the context and background that accompany the spending decisions. Fung & Weil argue such efforts rather create an effective "gotcha machine" instead of creating an effective government and more democratic society (Fung & Weil 2010, 108).

Another critical account 'Open data: Empowering the empowered or effective data use for everyone?' by Michael Gurstein describes how the opening up of digitized data led to situations where the already empowered got even more empowered and the marginalized become further marginalized (Gurstein 2011). In this regard, 10000scholen.nl might as well promote a situation where schools whose population of students are already declining will even further decline caused by the choices parents make based on the information criteria that were shown to them. To what extent does an average exam grade lower than nationwide tell anything about the teachers' qualities working at these schools? To what extent low exam grades tell anything about the quality of care the children will be given? While such open data projects generate particular information if citizens interact with them, they do not tell the whole story and their information is definitely not 'objective' if that would be possible at all.

[16] Cf. Wheredoesmymoneygo.org

Gurstein furthermore describes two cases of open data projects in Bangalore (India) and Nova Scotia (Canada) where newly digitized and openly accessible land records created the opportunity for the empowered to challenge land titles, exploit gaps in titles and take advantage of mistakes in documentation (Gurstein 2011). Whereas Fung & Weil question the supposedly beneficial effects of transparency and democratic accountability, Gurstein makes us aware of possible contradictory and repressive effects of open data projects.

Many other scholars (cf. Bates 2012; Johnson 2013; Kitchin 2013; Shah 2013), are also critical about the supposed beneficial effects and value claims surrounding open data. Their critique varies in numerous ways; addressing the open government data movement for promoting a politics of the benign (Fung & Weil 2010, Shah 2013), exemplifying how open data initiatives are empowering the empowered as part of a digital- or data-divide (Gurstein 2011; Johnson 2013), and Jo Bates is critical about open data initiatives facilitating the neo-liberalisation and marketization of public services (Bates 2012). Bates is concerned about the appropriation of the open data movement by political parties and business "on behalf of dominant capitalist interests under the guise of a Transparency Agenda" (Ibid.). Bates' critique mainly is focussed on the UK and it should be clear stated that the open government data movement is not driven by any one party but is diverse and consists of a range of constituencies with different agenda and aims (Kitchin 2013). Still, neo-liberalistic utterances are also present in the corpus of policy documents within this paper;

> Open data biedt de kans om een kleinere overheid te realiseren waarbij de markt zelf diensten kan optuigen die goed aansluiten op de informatievragen vanuit de maatschappij zonder dat de overheid hierin hoeft te sturen of financieel hoeft bij te dragen. (RWOD 2013, 21)

Under the guise of increased transparency and self-empowered citizens the quote above advocates for a reduction of government spending and the realization of a smaller (more efficient) government in order to enhance the role of the private sector, promoting self-regulating markets.

Contrary to what one might expect, the abovementioned critique stems from open data advocates. All agree that potential value creation by means of open data and the appropriation of open data is far from self-evident. Their critique unveils the double-edged-sword that resides as part of the open data concept and coincides with initiatives surrounding open data supported by the open government data movement which is gaining momentum.

2.5 Pattern of ideal information

Government policies persistently present the opening up of government data (e.g. traffic, weather, geographical, business, public sector budgeting datasets, and datasets about food-, safety-, and education-quality) to yield numerous benefits.[17] Despite the lack of evidence and the many contradictory implications of current open data initiatives the policy documents insist on the enormous amount of value which open data potentially *contain*. Before I continue to delineate the pattern that is underlying governmental reasoning, I will first clarify the definitions of data, information and knowledge and their interrelationships. So that we are able to answer the question what exactly are data? And when will data become information/knowledge? Anthony Liew provides useful definitions in this context;

> Data are recorded (captured and stored) symbols and signal readings. As symbols, 'Data' is the storage of intrinsic meaning, a mere representation. The main purpose of data is to record activities or situations, to attempt to capture the true picture or real event. Therefore, all data are historical, unless used for illustration purposes, such as forecasting.
>
> Information is a message that contains relevant meaning, implication, or input for decision and/or action. Information comes from both current (communication) and historical (processed data or 'reconstructed picture') sources. In essence, the purpose of information is to aid in making decisions and/or solving problems or realizing an opportunity.
>
> Knowledge is the (1) cognition or recognition (know-what), (2) capacity to act (know-how), and understanding (know-why) that resides or is contained within the mind or in the brain. The purpose of knowledge is to better our lives.[18] In the context of business, the purpose of knowledge is to create or increase value for the enterprise and all its stakeholders. In short, the ultimate purpose of knowledge is for value creation. (Liew 2007)

Activities and situations are the source of both data and information. Situations and activities generate information for it could imply relevant meaning to someone and when situations or activities are captured they become data. Once captured and stored "data can be

[17] This entails the wipe out of all barriers on the user's side and supply side of open data.

[18] Liew's definition clearly delineates the constructed nature of data and information and their interrelationships with knowledge. Nevertheless I disagree partially when he argues the purpose of knowledge is to better our lives. This part of his definition first of all assumes that information (and data) by definition accommodate possible knowledge generation. Even if this were true this might not be the case for everyone, which brings me to the second point of critique, whom exactly might benefit from the knowledge generated? Where I understand perfectly that knowledge might lead to value creation, why would I otherwise study at the University, it should be noted that knowledge, and information and data are by definition neither neutral nor innocent as they are all constructed.

processed back into information through compilation and analysis" (Liew 2007). When information is internalized the result is knowledge of which the purpose is value creation.

The interrelationships between data and information vis-à-vis information and knowledge overlap since they occupy different space at the same time (Liew 2007). Data, information and knowledge thus are interchangeable, though not in terms of their distinct definitions, which means that "…one man's data, can be another man's knowledge, and vice versa, depending on context" (Stewart 2002, 6). In case of open government data, the publication of this data is understood by government policy as a *new* situation, moving from a formerly closed system to an open system (Janssen et al. 2012, 259). This new system of open data is conceived beneficial because it entails an open system that presumably enables the public to use and create *information* which was previously obscure and thus open data contributes to the creation of new or improved knowledge that leads to the generation of value… Yet, many examples of open data initiatives yield exactly the opposite of what government policies documents argue or desire to yield. There is no hard evidence that there is really a causal relation between an increase of innovation/ economic growth and the publication of government data. Perhaps if we decide to pontifically close data it might 'also' spark innovation.

According to Jeffrey Alan Johnson (2013) one of the main reasons that causes these problems is governments' failure to understand the constructed nature of data. Johnson understands data as the product of what he calls *datized moments* in which information about an interaction is transformed into data (Johnson 2013, 2). Open data thus only exist in relation to a broader information system that gives it meaning. In this regard, Johnson accuses the open data movement of failing to see the constructed nature of (open) data (ibid. 2). Within current five policy documents I found two direct links to open data proponents; Digitale Agenda.nl resulted directly from the European digital agenda and therefore has been influenced by European commissioner Neelie Kroes. Secondly one of the authors of NHGODR is Paul Suijkerbuijk, a known proponent of open data in Europe and the Netherlands in particular and a member of The Green Land. As many government policies are influenced by this movement they too fail to do right to the complexity of (open) data.

Johnson's critique, and that of the other scholars mentioned in the previous sections, evince that current government policies, conscious or unconsciously and some more than others, tend to skate over 'reality' too fast regarding the value of open data. Open data do not self-evidently create democratic or economic value and might even enforce further repression of minorities depending on the situation. The argument within this thesis however goes beyond their critique.

Most scholars supplement their critique with alternative solutions in order to prevent problematic implications that might result from insufficient thought-through open data initiatives and policies. Johnson (2013) for instance advocates a theory of information justice that concerns in particular an emphasis on the constructed nature and wrongly presumed neutrality of data and subsequently its information that resides in it. Gurstein (2011) attempts to efface data-divides that accompany current and potential open data initiatives, through his seven principles that ensure an 'effective use' of open data. Briefly put, they try to find ways in order to ensure open data and the information that resides in it will be truly beneficial. As much as I agree with their critique which proofs that open data are not inherently a good thing; as much I even agree with most of their solutions, I argue that we should be aware that government policies understand the concept of open data to be inherently a good thing because of its inherent potential informational value. The underlying reasoning here is that information is an inherently good thing.

This even becomes visible within the barriers presented by the policy documents. All policy documents, some more than others, present barriers which first have to be overcome. Barriers that concern the correctness and completeness of data and the aim to find ways to make sure everybody is able to access and interact with the data. Since these barriers are framed to be barriers to overcome so that everybody is enabled to be *informed.* It becomes clear that it is not just about open data, but that it is information which generates value in the form of knowledge and its outcome. Government policies thus conceive information resulting from open data distributed through our information society to be an inherently good thing. The value generated from this information is conceived self-evident. The barriers in government policies merely pose a challenge that have to be overcome in order to make sure (pure) information will be dispersed as desired.

If we take on a governmental perspective of open data: Interaction with data leads to an enhancement of current or provision of new information about (past) activities and situations that once internalized lead to more sophisticated knowledge. This knowledge is used to create value. Therefore the opening up of government data enables the public to interact with data, and acquire information that was previously closed for them or more difficult to obtain. This widespread availability and acquisition of information by the public results in a better informed and more knowledgeable public, and in more people able to reap the benefits from being more knowledgeable. Hence, information is what makes open data to generate value in various ways for society and government. It is this ideal of information which I call *ideal information* derived from De Vries' ideal communication. A belief that, firstly information resides within open data (and potentially other technologies/ media as

well) and secondly through this information value is generated in various ways, benefiting society.

Government policies never mention this ideal explicitly because this would imply utopian thinking that might impair the realising of actual open data policies which means that there will be no benefits. And secondly authors of such policies, influenced by advocates of the open data movement respectively by Neelie Kroes and Paul Suijkerbuijk, might believe in the ideal of information themselves. Government policies implicitly refer to an ideology of open data proponents which basically argues that everyone has the right to understand and know what is collected by means of public resources and counts as public data and information. Therefore government data (and what the OGD movement calls public sector information) containing valuable information has to be actively opened up, flawless and pure to do no harm to the information and to assure that everyone is enabled both access and the ability to understand, in order to preclude any form of misinformation. In this sense, *ideal information* means that when citizens, organisations, government departments, and even non-human actors (computer programs), *interact with (idealised) open data,* society becomes more democratic and spreads out more socio-economic prosperity.

Ideal information is not solely based on the assumption that there will be such a thing as pure information. Ideal information implies moreover that information provided by open data leads to e.g. a participatory democratic society. It implies that a generation and distribution of (pure) information throughout society leads to a more knowledgeable society into a society where all our current problems are solved into a society we all desire; it entails not just a move towards a system that is more open through open data, but it entails a reasoning that through open information (since data results in the generation of information) we will become more knowledgeable and we will progress to our so desired destination. This perspective that resides within the OGD movement and informs the policy documents[19] adheres to what new media scholars call a technological imaginary, which I introduced and explained in section 1.2.

In chapter four the problematic implications within the policy documents, as a result from a univocal voice regarding open data (bringing forward only best practices, neglecting the flipside of the coin through referring to macro-economic studies that cannot account sufficiently as empirical data/ evidence but nevertheless act as authority) will be discussed in relation to the creation of *myths* and potentially an *imaginaire.* In this concluding chapter I

[19] And even to some extent ideal information also underlies the academic accounts. Whilst as Johnson aims for a theory of information justice, he implies that information if formatted justly it will lead to society of justice as well.

will discuss what it means for governments' and subsequently our conception of open data if we ignore the ideal that is manifest within OGD movements and policies. But first I will exemplify how various metaphors play an important role in upholding the sublime, which is ideal information. How metaphors point to a certain direction and amplify a belief in the potential of open data or ideal information.

Chapter 3: Metaphors That Open Data

Chapter three focuses on the use of metaphors regarding the (supposed) value of open data. A metaphor might be used subconsciously or strategically by journalists, politicians, policymakers and lawmakers, but what it does is that it gives "a familiar shape to something abstract" (Puschmann & Burgess 2014, 1698). Advocates of open data use metaphors and refer to domains most people are familiar with (source domain) and relate these domains to the value of the concept of open data (target). I will commence the first section by addressing various salient conceptual metaphors that manifest within the five policy documents covered in the previous chapter.

The analysis exemplifies how particular conceptual metaphors can be identified within the policy documents framing a necessity, an urge of progress and an intrinsic goodness in relation to open data. However, there are more metaphors at stake, discourse metaphors, which frame and shape our conception of open data and its supposed (intrinsic) value in relation to information enforcing an ideal of information identified in the former chapter.

3.1 A conceptual path guiding towards open data

In *Metaphors We Live By* it becomes clear that the most fundamental values of a culture are coherent with the most fundamental concepts in a culture. To clarify this matter Lakoff and Johnson refer to some cultural values and exemplify how these are coherent to UP-DOWN spatialisation metaphors which they frequently use as an example. For instance "more is better" is coherent with MORE IS UP and GOOD IS UP (Lakoff & Johnson 1980, 22-24). If we earn more money we are able to put more into our savings account and the amount of savings will *rise* (go up) which is generally conceived to be a good thing. When you earn a promotion, not only your salary will go up but you will go up one level within the organization's hierarchical pyramid and may be considered the new and promising 'star' of the company.

Lakoff and Johnson (1980) explain that MORE IS UP always seems to have a higher priority over GOOD IS UP since it has a stronger physical basis. Still in most occasions when things are UP they are considered to be good,[20] for example "the future will be better" coheres with THE FUTURE IS UP and GOOD IS UP and "there will be more in the future" coheres with MORE IS UP and THE FUTURE IS UP. These statements point to the concept of progress and express values that are deeply ingrained in our culture (Lakoff & Johnson

[20] Inflation might also rise although this is generally conceived not to be a good thing.

1980, 22). The policy documents make use of plenty spatialisation metaphors that claim for the necessity of open data in terms of benefits and progress. The resemblance between metaphors in the policy documents with Lakoff and Johnson's metaphorical coherency is striking. The Dutch digital agenda frequently mentions the potential of open data in terms of economic growth (UP). According to this document opening up public government data enable entrepreneurs to use these data in order to develop new applications and services that might be commercialized (Digitale Agenda.nl 2011, 14).[21] Therefore, open data *thrives* innovation and leads to an *increase* of economic activity.

Other policy documents put emphasis on the political, democratic and broader societal value of open data. NHGODR (2012) recognizes the *added* value in disclosing *new* data sources. VOO argues that an open data policy leads to *greater* transparency, which will benefit a modern participatory democracy as *more* people participate in the political decision making process leading to *more* efficiency in terms of budgetary savings in the future (VOO 2013, 11; AOO 2013). Another striking example follows when RWOD (2013) mobilises Tim Berners-Lee's five-star-deployment-scheme. This scheme is widely acknowledged as a guide for open data publication standards.22 RWOD applies this scheme in relation to four different open data implementation scenarios (RWOD 2013, 20). It takes no genius to see that the use of stars in relation with open data coheres with the metaphors GOOD IS UP as stars are always up from a human oriental perspective; and with MORE IS UP since adding *more* stars is regarded as better, a future *starring* open data is better.

Given the metaphorical coherency between the spatialisation metaphors UP and our cultural values of progress deeply embedded in our culture it is no coincidence NHGODR (2012) claims the opening up of data as a first *step* in the *right direction*. In our culture we regard the future to be in front of us. Governments conceive the opening up of public government data for example according to Berners-Lee's five-star-model, to be part of a future goal and speak in terms of a journey.

A conception of OPEN DATA in terms of a JOURNEY implies a path to be taken in order to move forwards, progress and reach the ultimate destination. A path subsequently implies the possibility of other governments to be closer or further away from reaching the ultimate destination. In order to not fall behind, we should keep on moving forwards so that

[21] "Verschillende studies laten zien dat de economische waarde van de innovatie en bedrijvigheid die ontstaat op basis van open data groot is. Met open data als grondstof kunnen nieuwe toepassingen en diensten ontwikkeld en vermarkt worden" (Digitale Agenda.nl 2011, 14).

[22] In short data that is opened up can be conceived as one star data ranging till five star data. More stars entail more openness and interoperability regarding the data since this data is (1) opened up under an open license, (2) is structured data, (2) non-proprietary formats are used, (4) URI's are used and (5) data is linked to provide context (5stardata.info 2014).

we remain front runners. In this regard the Dutch governments contend open data is the right direction;

> De ontwikkelingen gaan dusdanig snel dat Nederland op relatief korte termijn tot de achterhoede kan gaan behoren wat betreft het breder benutten van overheidsdata, als ze daar niet in mee gaat. *(RWOD 2013, 4)*

One should not mistake open data with the destination or goal. Open data serve as the means to come closer in reaching the end destination(s), which is; a more open government, a more participatory society together with more economic growth and so on. As I explain in the previous chapter, the benefits are premised by ideal information. Data and information are often conflated and intermixed, as if there is no difference between the two. The policy document AOO advocates for an equal information position as they claim citizens increasingly want to participate in government policymaking (AOO 2013, 13). Such conflations shortcut open data to automatically result in patterns of, new, more, better and ever improved information which subsequently results in knowledge generation that is used for the good.

The general tendency regarding open data policies at first seems to be MORE OPEN DATA IS BETTER and THE MORE OPEN DATA ARE IS BETTER; yet the emphasis is actually on information. More open data is better because more open data means more information widely accessible. Secondly the more open these data are the more they are interoperable to both human and non-human actors and again the better it is in terms of information coming closer to reaching its sublime state. The concept of open data is acknowledged to be the next step in making progress. Metaphors tell us that we are able to cross over potential barriers. No matter if we have to use more steps when the distance is too big to be covered at once, eventually we will progress and reach the desired destination. The conceptual metaphors leave little room for doubt; a future with open data is better and leads to progress. Evgeny Morozov calls this attitude "an Enlightenment-era pro information" (Morozov 2013). The policy documents try to convince their readers of open data's potential and necessity for its informational value that will bring us on a higher platform, a goal we should be heading to and is within our grasp. Using metaphors of verticality such as growth, height, accumulation, stacking and stars combined with horizontal oriented metaphors that envision a path and goal lying in front of us. Conceptual metaphors seem excellent in delineating this urge of progress projected on the concept of open data. The message is simple; start now and benefit or else…

Yet there are more metaphors at stake, metaphors that are also conceptual and in most cases cohere with the metaphors just covered, but above all should be addressed as discourse metaphors. Discourse metaphors shape and organize narratives, stories and myths and frame open data in terms of other domains by which they codetermine of how we should conceive of open data and define its value as a concept. These metaphors are at stake not just within the policy documents we have discussed so far, but the discursive politics of metaphor are influenced from beyond as well and concurrently manifest in governmental related press releases, news items and various notable websites, such as Frankwatching.com and iBestuur.nl.

3.2 Discourse metaphors that open data

In discussing metaphors of big data Puschmann and Burgess postulate technology metaphors "frequently naturalize their target by choice of a physical source domain" (Puschmann & Burgess 2014, 1697). According to Wanick technological metaphors consequently are critiqued for their potential to hide certain aspects of the technology as a result of metaphor choice and making others appear natural (Wanick 2004, in Puschmann & Burgess 2014, 1697). Metaphors of big data often evoke threats of nature e.g. tsunamis, avalanches and torrents, in which the challenge resides to not only curb these threats of nature by human intervention but control them in such manner so to successfully turn them into a valuable resources (Ibid. 1699). Metaphors of open data appear 'less threatening', in a way that data needs to be set free 'raw' in the 'open' instead of being curbed at first like big data. Nonetheless metaphors that open data equally evoke sources of nature in order this data can be exploited, mined, grown so that the public including the government itself are able to reap the benefits. In the next section I will explain how open data 'became' a natural resource.

3.2.1 Discursive politics of metaphor, a natural resource

On December 15 in 2010 European Commissioner Neelie Kroes announced that she would like to see that more citizens and business make use of open – machine readable – data. According to Kroes open data proffer an enormous potential with regards to the economy, profoundly a "€30 billion market in Europe" (Kroes 2010). Near the end of 2011 she gave another speech at the Press Conference on Open Data Strategy in Brussels. During this speech she emphatically stresses the economic potential of open data and postulates; "Just as oil was likened to black gold, data takes on a new importance and value in the digital age" (Kroes 2011). The same day the European Commission (EC) published an official press release with the following headline in Dutch; "Digitale Agenda: overheidsgegevens worden goudmijn" (EU 2011). According to the EC:

> Europe's public administrations are sitting on a goldmine of unrealised economic potential: the large volumes of information collected by numerous public authorities and services. *(EU 2011)*

The comparisons of open data with gold and goldmine and the resemblance with oil (i.e. black gold) did not pass by unnoticed. Right after Kroes gave her speech December 2011 Freek Blankena posted a summary at binnenlandsbestuur.nl with the headline "Kroes: Open Data is een Goudmijn" (BinnenlandsBestuur 2011). On January 2012 the Dutch broadcasting foundation, NOS, refers to Kroes and calls open data to be the new gold in an article about the potential of open data (NOS 2012). Furthermore in November 2012 two articles were published at Frankwatching.com, a popular Dutch weblog regarding marketing and media trends, mobilising Kroes' gold/ goldmine metaphors, respectively; 'Geen woorden maar open data: voorbereiding op een hackathon' by Tijs van den Broek and Mark Bastiaans (2012) and 'Open data: de kansen voor overhead & bedrijfsleven' by Marc Hesp (2012). And finally March 2014 RTL Nieuws refers to Kroes and her statement "Data is the new gold" in article about open data (RTL Nieuws 2014).

These are just a few of many websites and articles referring to Kroes (2010; 2011) and the open data metaphors in terms of gold and goldmine. When I examined the five policy documents I noticed that there has not been one document which explicitly uses Kroes' gold and goldmine metaphors. It could be possible that the metaphor had not yet been activated during the publishing of the Dutch Digital Agenda in May 2011, Kroes first speaks publicly of open data in terms of gold at the end of 2011. Still, considering the other four documents, I found it quite remarkable Kroes' metaphors are not mobilised in contrast to the more popular discourse. While all documents do refer to Kroes' advocacy which confirms the important role of Kroes in activating the Dutch government towards open data the policy documents mobilise a wider source domain regarding open data, that of a feedstock/ raw material.

In 2010 Arjan El Fassed, currently the director of the Open State Foundation (OSF), published an article at Frankwatching concerning open data and development aid. At that moment El Fassed was a member of the Dutch House of Representatives on behalf of GroenLinks as spokesman for development aid, defence and agriculture. In his article El Fassed argues the Dutch government is sitting on a goldmine of data concerning development aid;

> Open data voor ontwikkelingssamenwerking zie ik als nieuwe grondstof, een die we niet op zullen maken, maar die een duurzame basis biedt voor een nieuwe toekomst. Ook de Nederlandse overheid zit op een goudmijn aan data over ontwikkelingssamenwerking. Deze

> goudmijn is nu nog gesloten, maar daar komt wat mij betreft snel verandering in. *(El Fassed 2010)*

Not only does El Fassed speak of open data to be a goldmine. He mobilises the source domain of a feedstock and claims open data to be a new feedstock or raw material that will not deplete. Instead El Fassed claims that open data provides a sustainable foundation for a new future (El Fassed 2010). I do not want to claim that Arjan El Fassed nor Neelie Kroes are the main distributors or creators of the above metaphors. Rather I contend it is more productive to understand their influence as part of a network of numerous other actors that consist of people, organisations, and governments involved in projects and events co-shaping the Dutch open government data discourse. Several important actors, people, projects and organisations within Dutch open government data discourse are; Digitale Steden Agenda, Nationaal Congress Open Data, Paul Suijkerbuik[23], Stichting Netwerk en Democratie, The Green Land, Ton Zijlstra[24], Valerie Frissen[25] and Waag Society. El Fassed and the OSF currently are involved in plenty of projects and partnerships with the organisations and people above. Their involvement in the OGD movement contributes to the shape of the discourse and also to the use of metaphors that frame the discourse.

Thus the OGD movement should not be understood in terms of one particular group of advocates, this movement consists out of a manifold of different groups and peoples; government organisations, companies, NGO's, hackers, engaged citizens who all have different motives to participate in and shape the open government data discourse (Scherpenisse et al. 2012). It is a movement characterised by a plurality of motives.

Because of all these different motives it is not strange that a more general raw material/ feedstock metaphor proliferates within the government policy documents and possibly within the wider open data discourse as well since various actors above are also directly involved in drafting the policy documents, for instance Paul Suijkerbuik, (NHGODR 2012; VOO 2013; AOO 2013) and Waag Society (NHGODR 2012). Contrary to gold a feedstock opens up a wider discourse that codetermines how to comprehend and understand open data and why it potentially is of value for society. Where open data in terms of a goldmine or gold points to a source domain that enables one to understand open data in terms that it has to be mined or when vast amounts of data can be sifted for nuggets of gold, a feedstock directs to sources which open up possibilities to intermix open data in terms that it

[23] Project Manager Open Data at the Ministry of Internal Affairs and member of The Green Land.
[24] Co-author of the Open Data Handbook, member of The Green Land and he wrote the first drafts for launching data.overheid.nl commissioned by the Ministry of Internal Affairs. He is closely related to Paul Suijkerbuik.
[25] Principal scientist at TNO and chairwoman of Stichting Netwerk en Democratie.

can be mined, grown and exploited e.g. of how we use the earth's soil to grow things and maybe less prominent in current discourse in terms of how we use the wind, and water and the sun as energy resources. Gold has more narrow connotations with monetary value and economics and is less related to value regarding political and societal domains. In paragraph 2.4 I described how there is little to no empirical evidence available yet that provides evidence regarding the claims that open data is of genuine value for society. Apart from various macro-economic studies that are dubiously mobilised in order to justify claims about open data to be of great economic value, there seems to be no empirical 'evidence' available when it comes to the other domains; the political, democratic, societal, environmental and so on. When it comes to these domains the policy documents not just rely on best practices and anecdotes that 'exemplify' e.g. the benefits of open data concerning transparency and accountability for participatory democracy, but on discourse metaphors as well. In this sense it is no surprise that open data discursively became a feedstock and raw material rather than merely a goldmine which has been envisioned by Kroes (2011). The feedstock fills the void of a lacking evidence and supports so-called success stories.

In this section I will demonstrate how metaphors cover up the lack of evidence. I will describe which metaphors are used in relation to open data within the policy documents and explain how they currently frame open data in terms that open data really becomes beneficial and actually already is benefiting society.

3.2.2 Unveiling parts that are naturally hidden

In 2011 the Dutch digital agenda was published in line with the European digital agenda, which explicitly states 'growth and innovation with open data as feedstock' (Digitale Agenda.nl 2011, 14; AOO 2013, 22). NHGODR claims that since government organisations are able to organise information exchanges differently other (public) organisations will also be able to reap the *fruits* or benefits (NHGODR 2012, 9). A feedstock connotes a natural resource enabling growth through e.g. cultivating crops which can be harvested or reaped. In this sense open data can be conceived in terms of economic *growth*, the crops that have grown from fertile ground being the products of seeds that are put into land which received enough nutrition, water and sunlight. To speak of a feedstock is to inscribe a certain kind of potency and fertility into open data. This potency is argued to help society through *harvesting* innovative services and products, but also relates to transparency and accountability in *cultivating* a more participatory democratic society. The policy document VOO for instance argues that open data will contribute to greater transparency and accountability by referring to Louis Brandeis' famous quote "Sunlight is the best disinfectant" (VOO 2013, 9). VOO

argues that citizens want to be more involved and participate in politics and therefore government needs to be more transparent. Transparency is a "conditio sine qua non" to stimulate trust in government (VOO 2013, 9). Louis Brandeis' quote is mobilised to enforce this presupposition for publicity and transparency will have a cleansing effect.

Despite the lack of evidence regarding open data's benefits related to transparency, accountability and public participation,[26] the analogy of sunlight activates a conception that data should be 'opened' up by framing open data as a natural resource itself in terms of a feedstock that benefits from another sustainable natural resource, the sun. Following this analogy the public (and non-human actors as well) act as the sunlight and should have access, shed light, to public government data and be enabled the ability to reuse the data in order to *disinfect* and *prevent* the cultivation of possible diseases and stimulate the (healthy) *growth* of a transparent and more participatory democracy. Notice how open data in terms of a feedstock and in terms of cultivation coheres with MORE IS UP and UP IS GOOD inscribing an inherent democratic value in open data.

Remarkably the policy documents including VOO (2013) never discuss the possibility of getting sunburned. Being exposed to too much sunlight and dangerous UV radiation hurts! While these issues might seem to be merely of metaphorical relevance, we should take metaphors seriously therefore we should literally question what happens when e.g. transparency hurts. For instance in terms of privacy issues as currently many data analysis techniques allow us to re-identify individuals in data which is supposed to count as public data (Kulk & Van Loenen 2012). And what happens when the soil is not as fertile you imagined at first, when external forces cause the soil to deplete or to be contaminated and crops are getting damaged? 'Open Data Open Gevolgen' by Scherpenisse et al. (2012) examines five cases of open government data in the Netherlands. They found that current open data initiatives through their contribution to transparency frequently enforce existing arrangements which in many cases did not encourage innovation but works as a threshold for innovation and growth (Scherpenisse et al. 2012, 42). Yet, the policy documents merely emphasize the positive effects regarding the cultivation of this new and promising feedstock neglecting that these metaphors leave room for some critical intervention as well.

[26] In *To Save Everything, Click Here* Evgeny Morozov exemplifies how it is far from self-evident transparency schemes will automatically lead to more trust and a better functional 'democracy'. Morozov furthermore points to a research of John Hibbing and Elizabeth Theiss-Morse (2002) that most citizens are not interested in making political decisions themselves or providing input to those who do or even knowing the intimate details of the decision-making process" (Morozov 2013). In this regard Huijboom & Van den Broek (2011) argue there is no sufficient empirical evidence proving citizens to be actually more involved in politics through open data.

Previously I explained that the feedstock metaphor enables a wider discourse compared to open data in terms of gold. Not just are we able to speak in terms of growth and cultivation but a natural resource in terms of a raw material enables open data to be conceived in terms of "rawness and granularity" (RWOD 2013, 22).[27] Data that can be mined and gathered from mines or springs (the sources) as if governments' data really are the new gold of today's information society residing in a goldmine as promised by Kroes (NHGODR 2012, 16; RWOD 2013, 19). In this case, we could still wonder, what if less gold appears to be present than first envisioned. Or ask ourselves, how we are going to tap into this goldmine. And when the mine is 'opened up' who are able to do the dirty work? Working in a mine literally is a dirty job, unhealthy and dangerous. Working with open data might at first sight not be as dangerous and dirty. At least not in the sense that people will get hurt physically, apart from potential RSIs after hours of programming code in relation to data creating new applications and insights or after examining an exhaustive list of datasets…

Tapping into an open data goldmine however has its dangers too. Besides that opening up data and working with it requires hard work just like working in a mine, it might as well lead to real people getting hurt in reality. Indeed you might think at the transparency issues I just mentioned, but you can also think of the argument mentioned earlier in paragraph 2.4 in which the already empowered are getting more empowered leaving less for minorities (Gurstein 2011). Not in the sense that less data will be available, but in the sense that minorities for instance in Bangalore were not aware and able to make use of the open data land records with a result of them being exploited and left with *less* land and goods as before.

Obviously current discourse metaphors shape and frame open data in terms of natural resources. What we did not discuss yet is that these metaphors also hide and downplay several aspects. According to Puschmann and Burgess the real problems of discourse metaphors lie in inferring particular properties of the target that the source possesses but do not fit the target (Puschmann & Burgess 2014, 1699). Conceptions of open data in terms of natural resources, a feedstock/ raw material; soil, granular rocks and gold, frame open data as if data exist without human intervention, as a given. This givenness of data according to Puschmann and Burgess "is analogized through the givenness of natural resources, which can be mined or grown and which act as a form of capital with no persistent ties to their creator" (Puschmann & Burgess 2014, 1699). What the discourse metaphors hide is that data eligible to become 'open data' have been constructed. The discourse metaphors enable to speak in terms of data rawness and granularity, thereby they hide the intermediating steps necessary

[27] The original quote in Dutch; "ruwheid van de data" (RWOD 2013, 22).

to; gather, store, open, process and transmit data e.g. choices for data formats, choices for which data to collect and choices related to data processing software. For instance when RWOD (2013, 19) asserts that before opening up government data it is vital to "construct a solid foundation layer of metadata" that explains the cause and goal related to the data gathering, the discourse metaphors obscure the choices related to the construction of the metadata itself. As if data and its supplied meta-data are some sort of value-neutral resources that naturally exist, entities of software programs that also have evolved naturally.

This does not mean that the metaphors totally obscure the work that needs to be done by human actors, referring back to processes of mining or cultivation. Indeed if we take those metaphors literally, which we should, both are hard work. Tapping into a (gold)mine is not that easy just like growing crops is not easy either. The issue is that the discursively chosen metaphors take the existence of data for granted just like the existence of mines and land, and subsequently the presence of gold, fertile soil, sunlight and water. These resources exist independent from human intervention. What is downplayed is the constructed nature of data itself and their existence in a system of information (Johnson 2013) which involves methods and policies for collecting, selecting, analysing, interpreting and presenting/ visualising and entails leaps of translation and an initial purpose rather than being naturally evolved.

The observations by Puschmann and Burgess (2014) relating to big data apply on many levels on discourse metaphors of open data as well. An important remark in this sense is that open data as a raw material and natural resource has predictable qualities across quantities. Water (in relation to big data), gold, fertile soil and so on, whereas sheer quantity of these resources will improve the quality of predictions in a sense that more water will irrigate more crops and more gold residing in a goldmine can be used to create more jewellery or can be used for other purposes generating more income, sheer quantity does not automatically improve the quality of predictions regarding opening up of public data collected by or on behalf of governments. In contrast, the source domain of natural resources compared to open data is much more universally valuable than data when outside particular contexts, the value of data "is inscribed by analysis rather than being inherent in some sort of natural form of consumption" (Puschmann & Burgess 2014, 1699). I concur with Puschmann and Burgess who argue:

> Suggesting that the intrinsic meaning of data is, like nuggets of gold, already there, just waiting to be uncovered, means distancing the interpretation from the interpreter and her subjectivity. *(Puschmann & Burgess 2014, 1699)*

In this sense, the metaphor of goldmine, besides its connotations of a resource that exists naturally distancing its constructed nature would not be that bad if we erase the gold and leave blank what kind of material resides within the mine. This way the outcome is way more unpredictable. We still have to do the dirty work entailed with the dangerous risks in order to see if the data are potentially of use for others as well. And furthermore this metaphor acknowledges that each different raw material which is discovered demands a different treatment i.e. procedures and requirements and possibly different formats in order to open up the data which subsequently might lead to different creations and interpretations depending on the choices made. Deleting gold from the goldmine, puts emphasis on the interpretative work that needs to be done regarding identification of data, is it really gold we speak of? Or is it silver or copper? Each element requires different value assessments regarding benefits or drawbacks and subsequently entails different treatments. For example the opening up of crime statistics across different neighbourhoods might require a different treatment considering several privacy issues or the effect of this data regarding the value of properties within more criminal neighbourhoods compared to data that concern train schedules.

Nevertheless the metaphor keeps hiding essential characteristics that are part of data being digital software entities. This becomes visible when Puschmann and Burgess assert that "data can be relied on as a driver of economic growth only if its value is both predictable and stable rather than the result of ongoing interpretation and negotiation, but the value of big data seems extremely difficult to predict" (Puschmann & Burgess 2014, 1699). The opening up of data does not lessen this problematic situation of instability. As matter of fact the opening up of data rather extends and endures this process of ongoing interpretation and negotiation making the value of open data even more difficult to predict.[28] Take for instance the publication of crime statistic data across neighbourhoods and compare this to a natural resource such as silver. The mining of silver might lead to the creation of several different applications e.g. jewellery. However, whatever kind of application is created this does not change the fact that the initial resource was silver and had a particular initial value. If we melt the jewel the silver will still be silver. Now consider crime statistics.

In *To Save Everything Click Here* Evgeny Morozov (2013) describes what happened after the publication of crime statistics. Nowadays there are several maps that visualize crime

[28] I could continue on delineating various more disparities and contradictions regarding the mapping of the natural source domains in relation to their target domain. Even knowing that presumably there are plenty of other (discursive) metaphors at work. Yet this goes beyond the aims of this thesis which were foremost explorative and the disparities outlined above suffice to underpin my argument. Nevertheless I recommend further research into this area, in particular through an extension of the notion of discourse metaphors by apprehending the notion of material metaphors (cf. Van den Boomen 2014).

statistics so that the police could more effectively identify problematic areas, and at the same time it can be beneficial for people interested in making decisions about where to go and live. So far so good, if it were not for the fact that reality is a bit more complicated.

While those data visualisations might be of help for people avoiding buying properties in dodgy neighbourhoods, people already living there will have more trouble selling theirs. As a result of these 'public crime maps', people who already live in those dodgy areas might be less willing to report crimes worrying that each report of an incident will reduce the value of their properties. Morozov explains "the vary act of publishing the data will influence the quality of future data" (Morozov 2013). In other words, publishing crime statistics led to the crime statistics themselves being changed of structure and content. Whether this leads to a deviation of reality or not is not even relevant, the data are changed! Through feedback and processes of ongoing interpretation making the value of this data hard if not impossible to predict compared to raw materials such as silver.

Morozov laments the open data initiative to ignore such feedback mechanism. Some policy documents however somehow seem to include such feedback mechanisms, for instance in NHGODR;

> De datakwaliteit kan omhoog gaan doordat de organisatie gedwongen wordt om de eigen datakwaliteit te verhogen alvorens data naar buiten te brengen, of doordat er feedbackmechanismen worden ingericht. *(NHGODR 2012, 9)*

> Van belang is te beseffen dat door de data beschikbaar te maken en je een goede terugmeldfacitliteit inbouwt, de kwaliteit van de data eenvoudig kan worden vergroot. *(RWOD 2013, 5)*

Sadly NGHODR are only willing to acknowledge the dynamic ontology of open data as long as the opening up of data leads to an increase of data quality. The opposite effect is not discussed within this document.

3.2.3 Black Box

As we have seen in the previous sections discourse metaphors frame open data in terms of a natural resource, a raw material that can be used to cultivate or mine new insights, services, products which from then on can be used again to fabricate more new insights, services and products for benefit of society. Hiding processes of data being constructed and distance the interpretation from the interpreter (be it human or non-human) and its subjectivity. Frame a particular stability where there exists none, except if dynamics adhere to an increase of data-quality. In this last section I argue that discourse metaphors of open data can be identified

particularly to cohere to a *conduit discourse* which is a discourse metaphor of transmission that has originally been identified and criticized by Michael Reddy (1979) and to a discourse metaphor of storage *media as container*, and what it means to government policies' conception of open data. Both of these discourse metaphors are discussed in *Transcoding the Digital* by Marianne van den Boomen (2014).

The conduit discourse implies the idea that thoughts and feelings exist independent of any need for living human beings. A speaker puts ideas (objects) into words (containers) and sends them (along a conduit) to a hearer who takes the idea/object out of the word/container (Lakoff and Johnson 1980, 10; in Van den Boomen 2014, 92). Thoughts are pre-existing within human heads, and are able to flow either disembodied or reified in words and other representations, cf. (open) data, through ambient conduits arriving undamaged and unchanged in other human heads (Van den Boomen 2014, 92). This frame becomes visible in expressions such as 'I gave you that idea' and 'I didn't get it'. In the same line the open data discourse metaphors argue that potential informational value literally resides, like a "treasure" (NHGODR 2012, 3), within raw government data sources that only need to be 'opened up' in order to enable transmission to a wider public causing them to be more knowledgeable and society to be enriched. The conduit discourse metaphors provide little knowledge about open government data and their constructed nature. They leave open only one option when something goes wrong, that is "blame the speaker for failures" (Van den Boomen 2014, 93). In the end receiving and unwrapping of a package (data) is so passive and simple – what can go wrong? Blame the one who opened up data without using the right open standards and database formats.

The conduit discourse mixes well with the discourse metaphor of media as container (and the discourse metaphors previously mentioned). Considered that we simply have to tap into the (gold)mine or cultivate the feedstock in order to grow or mine the INTRINSIC potential of open data CONTAINING information that 'everybody' will be able to understand and become more knowledgeable, benefitting society. Apparently information can be perfected in such a way that everybody will be able to understand and access information residing in containers of (open) government data. Open data are necessary in order to create an efficient and thriving information society, as these data contain valuable information and this information should be liberated, flow freely and ubiquitously through our society in an urge of societal progress.

Current discourse metaphors frame open data as a natural resource combined with metaphors of media as container and the conduit discourse. This leads to the framing of public government data natural value-neutral resources, which inherently contain valuable

information and if opened up this information will be distributed through society benefitting current *information society*. This coheres with the conceptual metaphors that argue; the more data are opened up the more information is distributed and the better the information will be which results in more knowledge to be generated. These metaphors ignore that data are anything but intrinsically valuable like value neutral resources and that data are subject to wear and tear over time and need maintenance. The metaphors ignore that the opening up and potential reuse of data involve human and non-human interpretation and various methods for collecting and analysis depending on their purpose which subsequently influence the potential amount of information to be extracted and the extent to which the information appears to be valuable.

The policy documents present open data as a new technological development in terms of a natural resource that exists out there whether we humans decide to exploit it or not. This resource contains valuable information which therefore *should* be exploited allowing society and governments themselves to reap the benefits. In this sense the discourse surrounding open government data resembles what Van den Boomen described as the discourse of immediacy which concerns the withdrawal of mediation: "the desire for an instant result, as effortlessly as possible" (Van den Boomen 2014, 16). This discourse of immediacy is inevitably a matter of ideology and aimed at making invisible what counts, what computes, what mediates, and what materializes behind the screens (Ibid.). In a similar manner the open government data discourse is aimed at concealing the complexities that lie behind the increase of transparency, economic prosperity and a participatory democracy. Publishing government data (open data) according to the standards of Open Knowledge Foundation naturally will deliver (instant) results regarding the above matters. However what actually happens is that the metaphors used in the policy documents distract the reader from the complexity and subjectivity concerning open data. They 'black box' the relationship between the raw open data and the fulfilment of their potential regarding complex concepts such as transparency, democratic accountability and subsequently participatory democracy. The use of metaphors that stem from natural resources enforce an ideology that potential value is intrinsic in data and becomes valuable if opened up, however the metaphors hide what happens in processes before, during and after the data gathering, contextualising, manipulating, scraping, publishing, formatting, reusing and reinterpreting.

Summarizing, current metaphors close the curtains concerning processes that are related to making formerly 'closed' data 'open' and the effects afterwards of open data for instance in making society more open and transparent. To speak in terms of open thus is a metaphor itself, for 'open' downplays the many parts that construct data, which remain

closed and by means are not value-neutral or –natural, while it emphasises the aspects we 'desire' to be open, partially according to neo-liberalistic ideology as demonstrated by Jo Bates (2012) which has been referred to in section 2.4.2. These metaphors all serve the purpose of hiding the open being a metaphor for constructed data and information that are far from value-neutral or –natural in order to maintain a positivistic attitude towards open data.

Conclusion 4

In this final chapter I will elucidate how current government policies that succumb to ideas of ideal information might bring us closer to an infomaginary[29] instead of 'guiding us toward infotopia'. In order to elucidate my argument I will appeal on new media theories of De Vries (2012) and Flichy (2007) regarding ideal communication and the imaginaire which have been discussed in the introductory chapter, section 1.2.

4.1 A path to utopia?

The notion of utopia was first coined by Sir Thomas More who wrote the book Utopia in 1516. On one hand the word refers to the meaning of 'no-place' and on the other hand to 'good-place'. Both meanings merged together give a notion of an ideal place that is unreachable since it has no existence (Roest 2005, 20). Hence when something is utopian it would mean that it is actually unreachable, a desired image we could possibly come close to but remains unreachable. The concept of ideal communication puts emphasis on the utopian thinking patterns of expectations surrounding the technologies and new media rather than the specific content of these thinking patterns (Ibid.). In the introduction I referred to De Vries' example of NTT DoCoMo's advertisements which is just one example of a myth that gives voice to the necessary fictions which guide us along the road of progress. These myths tell us that sublime communication is near. NTT DoCoMo's myths logically anticipate on consumers' belief in necessary fictions. The two-pole system seduces consumers to buy the revolutionary communication technologies in order to progress. In this thesis a similar trait within and around open data policy documents has been detected.

In chapter 2.2 and 2.3 the results of textual content analysis revealed that currently government policies emphasize 'potential' benefits of open data and instead of drawbacks the policy documents merely mention barriers. 255 claims regarding open data have been categorised into the coding frame. Of these 255 occurrences 187 are benefits and only 68 occurrences refer to barriers. Inspired and influenced by open data advocates the government policy documents construct a necessary fiction with a clear starting point; the present in which huge amounts of public government data of great potential exist but where these data are 'closed', not publically available and hamper interoperability. The other end of the two-pole system exists of a sublime goal for the future; a society where at least all public government data is open and available for reuse according to the standards of the OKF.

[29] That is the imaginaire created through ideas of ideal information.

To support and sustain this necessary fiction, proponents of open data make clever use of myths that project parts of the road to the sublime that awaits us. In the introduction, section 1.2, I referred to Flichy who gave the example of The Well and demonstrated how a particular social context that made the experimentation possible is forgotten and subsequently how a "local technique" is presented as the "basic technique of a new social functioning" (Flichy 2007, 11). Flichy explains that the shift we see, from a local technical experiment to becoming a basic technique for a new social functioning, is performed by the myth. In this case thesis I unveiled similar shifts concerning open government data. In section 2.4 and 2.5 government's presuppositions, identified through content analysis, have been compared to academic studies and articles concerning open data projects. The result of this juxtaposition led to the conclusion that current policies lack empirical evidence and therefore use apparent best practices and 'success stories' to construct myths. Open data initiatives such as BuitenBeter, Verbeter de Buurt, Where does my money go, Openspending and Buienradar are mobilised to demonstrate the participatory democratic and innovative potential (power) of open government data. Just as The Well these open data initiatives become prime examples that demonstrate what open data are capable of.

Still, the most powerful open data myths probably root in academic reality. In order to legitimize open data, the policy governments rely on academic studies concerning the potential economic value of open data. Although these reports provide little to no evidence for the economic potential of open data, as has been evinced in section 2.4.1 (cf. Huijboom & Vd Broek 2011; Kronenburg et al. 2012), they still seem to perform well in constructing the myth that open data are of great economic value. A quick glance at current popular discourse through Google's search engine and various websites and news articles pop up that parrot the claims of open data proponents regarding the potential economic value and amount of economic growth it will deliver.

An example that supposed to demonstrate the economic value is mentioned in NHGODR (2012). This policy document refers to the 'fact' the Dutch economy largely depends on the services industry for 70%. To convince the reader of open data's great economic benefits, NHGODR sums up all kinds of open data examples that are connected to the services industry. The connection between the first fact or presumption and the services industry is easily made, especially if this information is combined with an actual macro-economic report that acts as an authority and gives voice to a myth that open data are of concrete great economic potential and stimulate innovation as opposed to when these data would remain 'closed'. The recent economic difficulties and the consequences that we are experiencing from this, make these myths even more powerful and appealing.

It would be faint concerning this thesis' argument to merely dismantle the myths and thereby to label the policy documents as utopian. Myths are important for policymaking as they inspire collective action (Bekkers & Homburg 2007 in Janssen et al. 2007, 264). Yet simultaneously myths mystify and blur views on reality. Janssen et al. (2012) argue that the "essence of a myth is that its existence is fictional or unproven" a legendary story without a determinable basis of fact and evidence and oppose myth to reality. Instead I argue it is more productive to understand myths to root firm in reality (cf. Flichy 2007; Mosco 2004). Myths in policy documents are not just fantastic stories that project how open data can be of value. Their roots stem from reality, and from there they deform reality, lift it up, and highlight some elements while leaving out other. This thesis clarifies in what manner language construct myths and how it gives voice to a particular direction. Furthermore this thesis explains in what manner some views seem more natural than others and metaphors play an important role.

The myths about open data thus are not entirely false nor do they tell the truth. Neither is it the case that the truth lays somewhere in the middle: myths are not neutral, they serve a purpose. In this case the myths uphold the sublime and it is according to that purpose that myths conceal, highlight and deform reality in order to allow ordinary citizens to experience part of the path and the presumed end destination that lies in front of them and make them belief there is something out there which is far better than the present. The conceptual and discourse metaphors delineated in the previous chapter serve a major role concerning the sustaining of myths and necessary fictions. In the previous chapter, section 3.1, it has been exemplified how conceptual metaphors make sure that open data connect with the human urge towards progress. Discourse metaphors furthermore 'transform' (closed) government data into a value-natural (and thus neutral resource full of potential) which need to be opened up. In doing so other discourse metaphors conceal that (open) data are constructed and construct process of 'opening up data' in terms of a conduit and a container. Conceptual and discourse metaphors work together in order to conceal the very fact that 'open' itself is a metaphor. The open within open data rather acts as a strong illusion and the metaphors sustain the path of progress of which open data serves the means to come closer.

Ideally to become open data, the data have to be opened up according to open standards. In such manner that data; are machine-readable, provided with enough metadata that humans as well as non-humans are able to read, interpret and communicate the data as information to other human and non-human actors (interoperability). The telegraph, the telephone, the radio, the television, mobile telephony and now open data; "all were initially perceived as trying to bridge space and time to such extent that people would be able to

communicate without obstacles and without misunderstanding" (De Vries 2005, 13). This can be related to the following question that has been stated in the introduction: To what extent does ideal communication fit the discourse about open government data and to what extent does ideal information as a new concept help to better understand the open data discourse?

Open data thus seems to perfectly connect to the pattern of ideal communication which De Vries dissected throughout the evolution of media. Since open data indeed can be seen as yet another technological advent that brings us closer to the utopian dream where all humans communicate 'like angels'. Yet as has been outlined in section 2.5 there is another dimension at stake. To use another metaphor, ideal communication is about how something is being transmitted along the conduit, and it is less about what is inside the container that is sent along the conduit. For open data an opposite phenomenon can be detected. The underlying thought pattern is about the amount of 'new' or 'improved' information which will be generated, that will create new innovative services through opening up the container. The attitudes on open data within government policy documents therefore can be better understood as idealizations of information adding another dimension to ideal communication.

4.2 Final destination

In the introduction I referred to Flichy (2007) regarding his claim that myths might transform a utopia into a mask ideology, which occurs when aspects of reality are successfully concealed in order to promote the new technology. When this is done successfully and people comprehend a new technology to be offering complete solutions for society, such as framed within current open data policies, we might speak of a technological imaginary. According to new media scholar Ellen Roest (2005) the technological imaginary offers the possibility to uncover implicit suppositions related to technology and its supposed positive effect on society (Roest 2005, 21). Chapter two and three demonstrate that open data are presented as the next technological advent and in what manner this is done. A necessary fiction is created and a sublime situation is opposed to the current situation with government data being 'closed'. Open data seem to bring us closer to utopia as opposed to closed data and solve most if not all of current society's dissatisfactions.

But open data are no new technologies, basically the data already exist and they 'just have to be opened up'. Ideal information, as exemplified in section 2.5, underlies a thinking pattern that information should be liberated as much as possible in such a way that every bit of information is traceable and reusable. This implies that governments will become more knowledgeable since every (potential knowledge containing) information is open. Now, to be specific the issue is not about more or less openness and its dangers, but it is about the

presumption that more and better and wider disseminated *information* is *capable* of making governments *more open* in this case through open data. This is what ideal information is all about. Ideal information presumes that, if everyone has access to and is able to reuse information, information will become complete and knowledge will be universally organized. Ideal information thus is an addition to ideal communication, and both keep trying to eliminate misunderstanding. In this case the policy documents do not just create a technological imaginary but an information imaginary: an infomaginary. This is the case because the 'new' information is presumed to potentially solve most if not all of current society's dissatisfactions!

An infomaginary means that current dissatisfactions and desires of society are projected on the mythical powers of information sources and in this case open data. Moreover that those open data are capable of liberating information in such manner that society's dissatisfactions will be solved and desires are met. Within an infomaginary people are not provided any further knowledge which is necessary to dissect the plural nature and complexities of open data, for instance in processes that surround data collection, the opening up of data and their potential reuse. Within an infomaginary we will not be able to evince how open data, entailing those processes, will make society better, more transparent, democratic and economic prosperous whereas so many other factors are of influence. Just as with ideal communication, the dream of ideal information will never be fulfilled. Morozov asserts:

> The quest to organize the world's knowledge cannot proceed without doing at least some violence to the knowledge it seeks to organize; making knowledge "legible," to borrow James Scott's phrase, is tricky regardless of whether a totalitarian government or a Silicon Valley start-up does it. *(Morozov 2013)*

With every new requirement and 'open' standard, for instance through adding a data-format, we are inscribing our human-computational subjectivity into data. Open or closed, data are collected by means of *purpose* and in this process information is reduced and thus violence is done to knowledge since information is never complete. Instead of reaching infotopia it is more likely that collectively we will reside in an infomaginary that neglects not just the plurality of data but especially overestimates the positive power of open information and open knowledge.

The purpose of this thesis is not to say open government data are bad or good, neither am I an adversary of open data (and open government data). This thesis purpose, as mentioned in the introduction, is to provide more insight into the pattern with which language

is used within government open data policy and to characterise how meaning concerning open data is given direction. To state that we are currently living an imaginary regarding open data constructed by government policy would be too bold. Neither am I an adversary of open data. Innovative applications that might benefit particular aspects of our society are created by means of open data, yet to state that open data are the next solution in order to enhance our society's democracy, economy or in general our wellbeing is far too extreme. For example, think of applications that open up political discussions, record anything that politicians say during debates or visualise all kinds of relations of politicians. One such application is Politwoops.nl that collects all tweets of Dutch politicians that have been deleted after being published first. Sometimes a politician might regret something he published on Twitter and therefore decides to delete the tweet. This application prevents the tweets from being deleted and enables people to access the text that actually was not meant to be visible. Such applications can be an opportunity for the wider public to engage with politics and thus can be considered as an enhancement of democracy through transparency. Yet an application such as Politwoops could also result into politicians being more careful in what they say and therefore can lead to a deterioration of the political debate. It is in these instances that current policy primarily focus on merely one viewpoint, based on ideal information.

Therefore I argue that we have to be aware not to create a potential imaginaire: an infomaginary in which open data are considered natural resources and a one way ticket to innovation and success. Instead an infomaginary might result in a collective disappointment. Since open data bring no ideal information and infotopia does not exist. Currently government(s) and the organisations involved in creating and influencing the government(s) policies neglect the plural nature of open government data and the potential drawbacks of open data applications. This is a result of being succumbed to both a desire of ideal communication and information. Thus I propose an approach towards open data policy that includes the many implications that accompany the concept of open data without promising mountains of gold and includes the excitement of exploring and discovering things that were previously unknown. Instead of hoping to reach infotopia, I desire to remain distant from a real infomaginary for this will obstruct open data to remain open for critique…

Discussion

The topic for this thesis emerged when I started to investigate the current state of open data. In the governmental policy documents that I read and during the conversations with people at open data events I noticed a utopian tone of voice regarding open data. At that moement I did not yet know whether this voice was really present or if I merely interpreted it as such. Furthermore I noticed that the policy documents I analysed solely referred to research documents that are written from an economic or ICT perspective (cf. Rufus Pollock 2009; Huijboom & Vd Broek 2011; Kronenburg et al. 2012). Based on macro-economic research the policy documents claim that open data are of great economic value which is estimated to be millions of euros each year (for Netherlands only) but where is the empirical (or real) evidence? During this thesis I found out that such evidence does not exist. Furthermore I did not find any material from a new media studies perspective on open data, therefore this research from a new media studies perspective that inquires these presuppositions of Dutch government policy about open data became a necessity.

This thesis provides more insight into the present and prominent (governmental) presumed values regarding the concept of open data. This has been investigated through textual content analysis of Dutch governmental policy documents. Every selected document has been read, marked and dissected, resulting in a coding frame. The results of the coding frame have been juxtaposed to relevant academic literature which allowed this thesis to demonstrate how current Dutch policy documents succumb to ideas of ideal communication as well as ideal *information*. At the beginning of this research only five Dutch policy documents were available. In the meantime it is highly probable that much more policy documents about open data have been created, therefore I am convinced this thesis covers just the tip of the iceberg. Further research by means of textual content analysis might provide more insight in terms of how government policy regarding open data develops and to what extent ideal information remains to be present. More research into this area is necessary in order to find out if the levels, concerning benefits and drawbacks dissected and identified within this thesis correspond or deviate with findings of new research.

As concluded in chapter three, to speak in terms of open with reference to open data is a metaphor itself. In chapter three it has been exemplified how conceptual metaphors match open data with the concept of progress and purpose while discourse metaphors tend to naturalise and ‘black box’ the constructed nature of open data. In other words, by means of metaphor analysis this thesis provides more insight into why particular claims and presuppositions will be easier taken for granted. Through metaphor analysis this thesis

exemplifies the role of language in constructing particular views that seem more natural and legitimate than others. The metaphors co-determine a particular view, in this case succumbed to ideal information, on reality and thereby foster the idea of a nearing infomaginary.

Are we really heading towards an infomaginary? To inquire and provide an answer to this question I first suggest to not just investigate more policy documents but also to widen the scope of research material regarding textual content analysis by adding more popular and academic articles into the corps of research material. Since this could assist in answering the question to what extent government policy correspond or deviate from public discourse and potentially offers an opportunity to further investigate if public discourse is affected by government policy or vice versa in order to provide more insight into a potential infomaginary being created. Secondly, I suggest to add more research methods in order to investigate the knowledge and views among citizens and government officials about open data, for example by means of survey. In this regard it is interesting to gauge their opinion concerning open data based on current government policy.

An interesting aspect of this thesis is that two methods have been used, one quantitative (textual content analysis) and one qualitative research method (metaphor analysis). Besides adding literature about open data from a new research discipline, new media studies, this thesis forms an addition to the debate between quantitative and qualitative research methods. Without the textual content analysis it would not have been able to discover patterns of ideal information and to determine the different overarching values of open data presumed by government policy. Subsequently the metaphor analysis provided more insight into how these values and presuppositions are framed within government policy legitimising particular viewpoints on open data while downplaying others. Therefore, taking a dual approach to inquire governmental open data policy seems to be a good method.

Nevertheless there are some remarks regarding the used methods within this research thesis. Firstly, I almost forgot to count the drawbacks during the coding process. Doing research the risk ‘to walk around with blinders on’ is high and something that has to be avoided. Sometimes you are so focused on discovering a particular kind of pattern that you tend to have only eyes for one part of the subject and forget the other. Janssen et al. (2012) reminded me that barriers and drawbacks are potentially part of government policies too. Furthermore at the start of the textual content analysis and in constructing the coding frame I started out with a three-level coding frame. This made the coding-process quite extensive yet difficult to fathom and it was not pragmatic in its use. The line between ‘which exact element acts as a requirement’ and ‘which as a benefit’ is very thin and depends on each different situation. From a pragmatic perspective, I decided to merge the second and third-level pieces

of content into one as I encountered this same structure within existing research executed by Janssen et al. (2012). For further research I recommend to use a two-level structure from the starting point since this approach proved to be more pragmatic.

Finally this thesis turned out to be an endless story about Dutch governmental framing of open data. Originally I received three months to write this thesis. Since I used these two methods I extended the three-month deadline by roughly another three months. And yet I am positive that both methods were necessary and indispensable. Therefore I believe that extending the deadline has improved the quality of this thesis. A thesis that is aimed to open up Dutch government policy about open data and the wider discourse, in order to make those who read it think about the thinking patterns that concern information and openness and what these values mean and hold for our society.

Bibliography

Barneveld, van, Dirk. 2013. 'Open data met beleid: Opendatabeleid bij het ministerie van Infrastructuur en Milieu.' *ibestuur online.* September 16th 2014: http://ibestuur.nl/podium/open-data-met-beleid

Bates, Jo. 2012. 'This is what modern deregulation looks like: co-optation and contestation in the shaping of the UK's Open Government Data Initiative'. *The Journal of Community Informatics* 8.2.

Blankena, Freek. 2011. 'Kroes: Open Data Is Een Goudmijn'. *BinnenlandsBestuur.* September 16th 2014: http://www.binnenlandsbestuur.nl/digitaal/nieuws/kroes-open-data-is-een-goudmijn.3097230.lynkx

Boomen, van den, Marianne. 2006. 'Networking by Metaphors.' Paper presented at 56th Annual Conference of *International Communication Association*, Dresden.

Boomen, van den, Marianne. 2014. *Transcoding the Digital: How Metaphors Matter in New Media.* Amsterdam, Institute of Network Cultures, 2014.

Broek, van den, Tijs, Mark Bastiaans. 2012. 'Geen woorden maar open data: voorbereiding op een hackathon'. *Frankwatching.* September 16th 2014: http://www.frankwatching.com/archive/2012/11/12/geen-woorden-maar-open-data-voorbereiding-op-een-hackathon/

Davies, Tim. 2010. 'Open Data, Democracy and Public Sector Reform. A look at open government data use from data.gov.uk.' Retrieved from: http://practicalparticipation.co.uk/odi/report/wp-content/uploads/2010/08/How-is-open-governmentdata-being-used-in-practice.pdf

Deacon, David, Michael Pickering, Peter Golding and Graham Murdock. 1999. *Researching Communications: A Practical Guide to Methods in Media and Cultural Analysis*. London, Arnold.

Dietrich, Daniel, Jonathan Gray, Tim McNamara, Antti Poikola, Rufus Pollock, Julian Tait, and Ton Zijlstra. "Open Data Handbook." Open Data Handbook. 2009. Accessed June 25, 2015. http://opendatahandbook.org/.

El Fassed, Arjan. 2010. 'Open data en ontwikkelingssamenwerking'. *Frankwatching.* September 16th 2014: http://www.frankwatching.com/archive/2010/11/05/open-data-en-ontwikkelingssamenwerking/

European Commision. 2011. 'Digitale Agenda: overheidsgegevens worden goudmijn'. *Press Release Database.* September 16th 2014: http://europa.eu/rapid/press-release_IP-11-1524_nl.htm

Flichy, Patrice. 1999. 'The construction of new digital media'. *New Media and Society* 1(1): 33-39.

Flichy, Patrice. 2007. *The internet imaginaire*. Massachusetts, The MIT Press.

Fung, Archon. 2013. 'Infotopia Unleashing the Democratic Power of Transparency.' *Politics & Society* 41 (2): 183-212.

Fung, Archon, and David Weil. 2010. 'Open Government and Open Society.' In D. Lathrop & L. Ruma, eds. *Open Government: Collaboration, Transparency, and Participation in Practice.* O'Reilly Media.

Gurstein, Michael B. 2011. 'Open data: Empowering the empowered or effective data use for everyone?' *First Monday* 16 (2).

Halonen, Antti. 2012. 'Being open about data. Analysis of the UK open data policies and applicability of data.' London, Finnish Institute. http://finnish-institute.org.uk/images/stories/pdf2012/being%20open%20about%20data.pdf.

Hansen, Henning Sten, Line Hvingel, and Lise Schrøder. 2013. 'Open Government Data–A Key Element in the Digital Society.' In A. Kö, C. Leitner, H. Leitold and A. Prosser (Eds.) *Technology-Enabled Innovation for Democracy, Government and Governance*. Heidelberg, Springer: 167-180.

Hellsten, Iina. 2002. *The politics of metaphor: Biotechnology and biodiversity in the media*. Tampere, Tampere University Press.

Hesp, Marc. 2012. 'Open data: de kansen voor overhead & bedrijfsleven'. *Frankwatching.* September 16th 2014: http://www.frankwatching.com/archive/2012/11/22/open-data-de-kansen-voor-overheid-en-bedrijfsleven/

Huijboom, Noor, and Tijs van den Broek. 2011. 'Open data: an international comparison of strategies.' *European Journal of ePractice* 12(1): 1-13.

Janssen, Katleen. 2012 'Open government data and the right to information: Opportunities and obstacles.' *The Journal of Community Informatics* 8 (2).

Janssen, Marijn, Yannis Charalabidis and Anneke Zuiderwijk. 2012. 'Benefits, Adoption Barriers and Myths of Open Data and Open Government.' *Information System Management* 29(4): 258-268.

Johnson, Jeff Alan. 2013. 'From Open Data To Information Justice.' Paper presented at the Annual Conference of *The Midwest Political Science Association,* April 13, 2013, Chicago, Illinois. Available at SSRN: http://ssrn.com/abstract=2241092

Katz, James E., and Mark Aakhus. 2002. *Perpetual Contact: Mobile Communication, Private Talk, Public Performance.* Cambridge, Cambridge University Press.

Kitchin, Rob. 2013. 'Four critiques of open data initiatives'. *The Programmable City.* September 14th 2014: http://www.maynoothuniversity.ie/progcity/2013/11/four-critiques-of-open-data-initiatives/

Kroes, Neelie. 2010. 'My vision for eGovernment, and how to make it real'. *Lift-Off towards Open Government" conference.* Brussels. September 16th 2014: http://europa.eu/rapid/press-release_SPEECH-10-752_en.htm

Kroes, Neelie. 2011. 'Data is the new gold'. *Press Conference on Open Data Strategy.* Brussels. September 16th 2014: http://europa.eu/rapid/press-release_SPEECH-11-872_en.htm

Kronenburg, Tom, Ton Monasso, Ellen Boschker and Marcel Thaens. 2012. 'De waarde van open data: Keuzes en effecten van open-datastrategieën voor publieke organisaties'. *Rijksoverheid.* September 16th 2014: http://www.rijksoverheid.nl/documenten-en-publicaties/kamerstukken/2012/06/08/kamerbrief-over-rapport-de-waarde-van-open-data.html

Kulk, Stefan, and Bastiaan van Loenen. 2012 'Brave new open data world.' *International Journal of Spatial Data Infrastructure Research* 7: 196-206.

Lakoff, George. 2008. *Don't think of an elephant: Know your values and frame the debate*. Chelsea Green Publishing.

Lakoff, George, and Mark Johnson. 1980. *Metaphors we live by*. Chicago, University of Chicago press.

Liew, Anthony. 2007. 'Understanding data, information, knowledge and their inter-relationships'. *Journal of Knowledge Management Practice* 8.2.

Lister, Martin, Jon Dovey, Seth Giddings, IainGrant, and Kieran Kelly. 2009. *New media: A critical introduction.* London, Routledge.

Morozov, Evgeny. 2013. *To Save Everything, Click Here: The Folly of Technological Solutionism*. London, Allen Lane.

Mosco, Vincent. 2004. *The Digital Sublime: Myth, Power, and Cyberspace*. Massachusetts, The MIT Press.

Nerlich, Brigitte, and Rusi Jaspal. 2012. 'Metaphors We Die By? Geoengineering, Metaphors and the Argument From Catastrophe.' *Metaphor and Symbol* 27(2): 131-147.

NOS OP 3. 2012. 'Er kan véél meer met open data dan naar de wc'. September 16th 2014: http://nos.nl/op3/artikel/334474-er-kan-veel-meer-met-open-data-dan-naar-de-wc.html

Peters, John Durham. 1999. *Speaking into the Air*. Chicago, University of Chicago Press.

Pollock, Rufus. 2009. *The economics of public sector information*. Cambridge, University of Cambridge.

Puschmann, Cornelius, and Jean Burgess. 2014. 'Metaphors of big data.' *International Journal of Communication*, 8: 1690-1709.

Roest, Ellen. 2006. *Foutloos, snel en natuurlijk; Een verkenning naar utopieën rond communicatie in adviesliteratuur*. Ma Thesis, Universiteit Utrecht.

RTL Nieuws. 2014. Open data, wat is het en wat heb ik er aan?'. September 16th 2014: http://www.rtlnieuws.nl/nieuws/binnenland/open-data-wat-het-en-wat-heb-ik-er-aan

Scherpenisse, Jorren, Josta de Hoog, Mark van Twist, Albert Meijer and Martijn van der Steen. 2012. 'Open data, open gevolgen: Kritische kanttekkeningen bij de nieuwe 'goudkoorts''. Paper presented at the conference *Rijk met Wetenschap: Bestuur is Informatie*, University Tilburg.

Shah, Nishant. 2013. 'Big Data, People's Lives, and the Importance of Openness'. *dmlcentral*. September 16th 2014: http://dmlcentral.net/blog/nishant-shah/big-data-peoples-lives-and-importance-openness

Stewart, Thomas A. 2002. *The Wealth of Knowledge: Intellectual Capital and the Twenty-First Century Organization.* London, Nicholas Brealey Publishing.

Vries, de, Imar. 2005. 'Mobile Telephony: Realising the Dream of Ideal Communication?'. Hamill, L. en A. Lasen (eds.) (2005), *Wireless World: Mobiles – Past, Present, Future.* London: Springer.

Vries, de, Imar. 2012. *Tantalisingly close: An archaeology of communication desires in discourses of mobile wireless media*. Vol. 7. Amsterdam, Amsterdam University Press.

Zinken, Jörg, Iina Hellsten, and Brigitte Nerlich. 2008. 'Discourse Metaphors'. In *Body, Language and Mind:Sociocultural Situatedness*, 363-86. Berlin: Mouton de Gruyter.

Government Policy Documents

(AOO): Ministerie van Binnenlandse Zaken en Koninkrijksrelaties. 2013. ‘Actieplan Open Overheid’. September 16th 2014: https://data.overheid.nl/sites/data.overheid.nl/files/actieplan-open-overheid.pdf

(Digitale Agenda.nl): Ministerie van Economische Zaken, Landbouw en Innovatie. 2011. ‘Digitale Agenda.nl: ICT voor innovatie en economische groei’. September 16th 2014: https://www.forumstandaardisatie.nl/fileadmin/os/publicaties/Digitale_Agenda_ELenI.pdf

(NHGODR): Rooy, de, Kasper, Ivonne Jansen-Dings, Miriam Reitenbach and Paul Suijkerbuijk. 2012. ‘Noord-Holland Geeft Open Data de Ruimte’. *Waag Society*. September 16th 2014: http://www.noord-holland.nl/web/Digitaal-loket/Kaarten-en-data-1/Open-data-1/Informatie-over-Open-data.htm

(RWOD): Crook, de, Roger, Boris Everwijn, Kier van Gijssel, Raymond Ingenhut, Natasja ’t Hart, Adri-Jan de Korte and Huibert-Jan Lekkerkerk. 2013. ‘Rapport Waterschappen en Open Databeleid. *Informatiehuis Water*. September 16th 2014: http://www.ihw.nl/informatiebeleid/open_data_beleid

(VOO): Ministerie van Binnenlandse Zaken en Koninkrijksrelaties. 2013. ‘Visie Open Overheid.’ September 16th 2014: https://data.overheid.nl/sites/data.overheid.nl/files/visie-open-overheid.pdf

Appendix

Coding Frame Benefits

Legend:

Between brackets (..): number of quotes.

Red: Indicators which are more frequently quantified

Blue: Environmental value, merged with societal value

Value	Benefit category	Requirement	Quote
POLITICAL/ DEMOCRATIC (15 quotes)	More transparency (7)	Equal access to open data (6)	"*De overheid laat zien wat ze wel en niet doet en hoe ze werkt*. De structuur en de rollen van het overheidsapparaat dienen daarvoor *helder te zijn*. *Daarom wordt informatie* over activiteiten, besluitvorming en financiële informatie *actief open en beschikbaar gesteld voor het publiek.* … in formaten die voldoen aan open standaarden" (Actieplan open overheid 12-13). "Een laatste belangrijke baat is dat het openstellen van data de transparantie van het beleid vergroot, immers zowel het publiek als de waterbeheerder kan over dezelfde informatie beschikken. Op deze manier worden democratische besluitvormingsprocessen met betrekking tot waterbeheer ondersteund" (Rapport Waterschappen en Open Databeleid 6).
		Development of new products and services (1) Equal access to open data (3e)	"De provincie wil zich structureel inzetten voor de ontsluiting van Open Data… Het komende jaar ligt de focus daarbij op de Noord-Hollandse gemeenten, want zij hebben veel te bieden. Gezamenlijk hebben we een enorm aanbod aan openbare informatie, dat ontwikkelaars de kans geeft leuke, handige en praktische toepassingen te ontwikkelen. Dit draagt bij aan een transparante overheid en stimuleert innovatie en economische

			activiteit" (Noord-Holland geeft open data de ruimte 3).
	More transparency (4)	Active Open data policy (2)	"Vergroting transparantie" (Rapport Waterschappen en Open Databeleid 23).
	Democratic accountability (6)	more participation and self-empowerment of citizens (1) Governmental collaboration (1)	"Het toegankelijk maken en beschikbaar stellen van overheidsinformatie kunnen leiden tot economische impulsen. … Het met de samenleving beleid ontwikkelen, uitvoeren en evalueren; Openstaan voor initiatieven in de samenleving; Openheid van zaken geven over haar handelen; Stimuleren om gegevens te gebruiken of te hergebruiken voor nieuwe diensten en producten" (Rapport Waterschappen en Opendatabeleid 12).
	More transparency (5e) Lowering the thresholds to information (1)	Easier access to data and discovery of data (4) Active open data policy (2e)	"De definitie zoals hierboven beschreven gaat in op actief open databeleid. … Het is het verschil tussen een bestemmingsplan dat in de krant kenbaar is gemaakt, maar slechts op kantoor is in te zien, of dat deze via de website vrijelijk beschikbaar is. Letterlijk en figuurlijk worden de drempels tot informatie verlaagd" (Rapport Waterschappen en Open Databeleid 9).
	More transparency (6-7e) Democratic accountability (3-4e)	Equal access to data (4-5e) Easier access to data and discovery of data (2-3)	"Ook veel websites zoals het Deense folketsting.dk houden de activiteiten en het proces van wetten maken in het parlement bij, zodat men kan zien wat er precies gebeurt en welke parlementariërs erbij betrokken zijn" (Noord-Holland geeft open data de ruimte 8). "Websites als Argumenten.nl en Politiekinzicht.nl maken handig gebruik van deze data en bieden inzicht in het democratische proces. Op deze manier is het niet alleen eenvoudiger volksvertegenwoordigers te volgen en te controleren, maar krijgen meer mensen toegang tot informatie die voor hen doorgaans verborgen blijft" (Noord-Holland geeft open data de ruimte 15).
	Democratic Accountability (5-7)	Transparency (5)	"*Burgers willen steeds meer meebeslissen en meedenken* over het beleid en daarom is een *gelijkwaardige informatiepositie* belangrijk. Tenslotte gaat het bij goed en verantwoord overheidsbestuur ook om financiële transparantie. Daarbij zal worden

		Equality (6)	onderzocht (of en) hoe budgetten en uitgaven van de overheid beter en actiever beschikbaar kunnen worden gemaakt" (Actieplan open overheid 13). "Bovendien laat de overheid aan burgers zien welke gegevens worden gebruikt in welke processen en geeft hen de gelegenheid deze gegevens te corrigeren wanneer nodig" (Actieplan Open overheid 15). OKFN is een internationale non-profit organisatie die open kennis bevordert, door middel van open content en open data. … Budgetmonitoring is een middel waardoor burgers, gemeenschappen en andere organisaties toegang krijgen tot financiële informatie en aldus zicht krijgen op begrotingsprocessen en de besteding van middelen van de overheid. *Met behulp van deze methodiek worden overheidsbestedingen gecontroleerd en kan een dialoog plaatsvinden tussen burgers onderling en tussen organisaties en overheid* over prioritering, behoeftes en aanpak van problemen" (Actieplan Open Overheid 24). "Transparantie en openheid van de overheid is een voorwaarde voor burgers om de overheid en haar activiteiten in de publieke sector te kunnen bewaken en te controleren. De Amerikaan Louis Brandeis refereerde aan het voorkomen van corruptie en fraude toen hij stelde dat "Sunlight is the best disinfectant. Openbaarheid heeft een reinigende werking" (Visie Open Overheid 9). "Bovendien laat de overheid aan burgers zien welke gegevens worden gebruikt in welke processen en geeft hen de gelegenheid deze gegevens te corrigeren wanneer nodig" (Visie Open Overheid 16).
	Democratic Accountability (9e)	Equality (6e) Interoperability (1)	"Om (technische) belemmeringen bij hergebruik te voorkomen en om de maatschappelijke waarde van de data te maximaliseren wordt de overheidsdata zo veel als mogelijk gepubliceerd in formaten die voldoen aan open standaarden. … Burgers willen steeds meer meebeslissen en meedenken over het beleid en daarom is een

			gelijkwaardige informatiepositie belangrijk" (Visie Open Overheid 15).
	Equality	Easier access to data and discovery of data (4e) Equal access to data (6e)	Voor een moderne democratie is het nodig dat burgers en andere partijen goede en snelle toegang hebben tot openbare overheidsinformatie. In de netwerk- en informatiesamenleving is sprake van meer gelijkwaardigheid tussen overheid, burger en bedrijf. (Visie Open Overheid 9).

Value	Benefit category	Requirement	Quote
SOCIETAL VALUE (16 citations)	Public Engagement (4)	Active promotion of data reuse (2) (open data portal) Equality (1)	"Op de nieuwe portal www.jijmaaktutrecht.nl zijn de eerste gemeentelijke datasets openbaar gemaakt: straatnamen, straatmeubilair en de grenzen van stad, wijken en buurten. *Samen met andere partijen* in de stad/regio en andere Europese steden *wordt het concept open data verder doorontwikkeld*. Samen met de stad wil de gemeente de portal www.jijmaaktutrecht.nl stapsgewijs verder ontwikkelen. *Iedereen* kan hier aandacht vragen voor een mooi initiatief en in contact komen met mensen die hierbij kunnen helpen" (Actieplan Open overheid 10).
	Public Engagement (2e)	Equal access to open data (10)	"Voor gestructureerde informatie (open data) kan de prioritering worden bepaald op basis van de economische en maatschappelijke waarde (financiële data, lokatiegebonden data, data over mobiliteit). Aan de hand van maatschappelijke vraagstukken wordt informatie proactief beschikbaar gesteld door de overheidsorganisatie zodat burgers en bedrijven actief kunnen meedenken over de oplossing van vraagstukken" (Actieplan Open Overheid 17).
	Public Engagement (3e) Innovation (10)	Equal access to open data (2e)	"Openheid is ook belangrijk voor maatschappelijke initiatieven waarmee (groepen) burgers maatschappelijke vraagstukken oppakken en sociale ondernemingen starten" (Visie Open Overheid 11). [in context van] "Door inzet van open data kunnen burgers en bedrijven innovatieve toepassingen ontwikkelen met overheidsdata" .
	Innovation (2-4)	Equal access to open data (3-5e)	"Dit hergebruik kan daardoor veel positieve economische en maatschappelijke effecten opleveren. Dit kan alleen als gekozen wordt voor een actieve in plaats van passieve

			open data verstrekking" (Rapport Waterschappen en Open Databeleid 4). "Op die manier kunnen andere partijen deze informatie gebruiken, verrijken en (op maat) aanbieden aan de maatschappij" (Rapport Waterschappen en Open Databeleid 13). "Op die manier kunnen andere partijen deze informatie gebruiken, verrijken en (op maat) aanbieden aan de maatschappij. Overheidsdata is makkelijk vindbaar via het open data portaal" (Visie Open Overheid 14).
	Public Engagement (4e) More participation and self-empowerment of citizens	Development of new products and services (1) Equal access to open data (6e)	"Kapotte lantaarnpalen, vieze bushokjes of losliggende stoeptegels: applicaties als BuitenBeter en Verbeter de Buurt bieden de mogelijkheid melding van het probleem te maken. Dergelijke applicaties zouden echter niet mogelijk zijn zonder Open Data" (Noord-Holland geeft open data de ruimte 10).
	Innovation (5)	Active promotion of data reuse (2e) & Government collaboration	"Burgers en bedrijven worden actief uitgenodigd en gefaciliteerd om samen met of los van de overheid te werken aan oplossingen voor maatschappelijke vraagstukken en aan de ontwikkeling van innovatieve toepassingen" (Visie Open Overheid 15).
	Improvement of citizen services (2)	Availability of information for investors and companies (3)	"Door als gemeente data vrij te geven over de locatie van bushokjes, de routes van de vuilophaaldienst en zelfs de onderhoudsplanning voor objecten in de publieke ruimte, kunnen dienstverleners steeds beter de informatie op een gebruiksvriendelijke manier aanbieden" (Noord-Holland geeft open data de ruimte 10).
	Improvement of citizen services (2e)	The ability to merge, integrate, and mesh data (2)	*Bedrijven kunnen bijvoorbeeld meer bijdragen aan de innovatie van publieke diensten. Als wij onze gemeentelijke gegevens beschikbaar stellen zijn zij in staat om deze te combineren met andere informatie en daardoor nieuwe, nuttige diensten te leveren*

		Innovation (1) Availability of information for investors and companies (2e)	(Actieplan Open Overheid 8).
	Innovation (6-9e) Development of new products and services (4)	Equal access to open data (7-10e)	"De overheid stelt gegevens beschikbaar die door hergebruik kunnen bijdragen aan ontwikkeling van nieuwe diensten" (Actieplan open overheid 13). "De overheid stelt gegevens beschikbaar die door hergebruik kunnen bijdragen aan ontwikkeling van nieuwe diensten" (Visie Open Overheid 15). "Bovendien opent de *vrije beschikbaarheid* van geodata de mogelijkheid te komen tot andere toepassingen" (Digitale Agenda 18). "Met deze open data kunnen organisaties en bedrijven in de stad apps of andere nieuwe toepassingen maken. Aan de gemeenteraad is voorgesteld dat alle datasets in principe openbaar zijn, tenzij er juridische risico's of privacyaspecten aan vast zitten" (Actieplan Open overheid 10).
	Creation of new insights in the public sector (1)	The ability to merge, integrate, and mesh data (2e)	"Terwijl er eindeloos veel gevallen zijn van de manieren waarop Open Data reeds zowel sociale als economische waarde creëert, weten we nog niet welke nieuwe mogelijkheden er gaan komen. Nieuwe combinaties van gegevens kunnen nieuwe kennis en inzichten genereren, die weer kunnen leiden tot geheel nieuwe vormen van toepassing. Dit onaangebroken potentieel kan ontketend worden als we openbare overheidsgegevens in Open Data veranderen" (Noord-Holland geeft open data de ruimte 7).
ENVIRONMENTAL	Innovation (10e)	Availability of information for	"*Beschikbaarheid van geo-informatie*, zoals gedetailleerde plaatsgebonden informatie

(1) **In societal value**		investors and companies (3e)	over bodemgesteldheid en gewasontwikkeling, is een belangrijke randvoorwaarde voor een grootschalige doorbraak van precisielandbouw. Daarom wil de overheid geo-informatie als open data beschikbaar stellen zodat toepassingen kunnen worden ontwikkeld voor precisielandbouw. Hierdoor kan worden bespaard op gebruik van milieubelastende stoffen, brandstofgebruik en uitstoot van broeikasgassen, met behoud van een goede opbrengst" (Digitale Agenda 18).

Value	Benefit category	Requirement	Quote
ECONOMIC (16)	Economic Growth and stimulation of competitiveness (8)	Innovation (6) Development of new products and services (7) Availability of information for investors and companies (7)	"Naast dat ondernemers efficiënter kunnen werken [open standaarden], moeten ze kunnen groeien. Daarom gaat het kabinet overheidsdata als grondstof beschikbaar stellen voor nieuwe innovatieve diensten. Creatieve ondernemers kunnen deze open data verwerken, zoals bijvoorbeeld Buienradar doet met data van het KNMI" (Digitale Agenda 6). (Brackets by Svb). "Buiten het op een efficiënte manier zakendoen willen ondernemers kunnen groeien. De overheid creëert ruimte voor groei door overheidsdata als grondstof beschikbaar te stellen aan ondernemers. Creatieve ondernemers kunnen deze open data verwerken tot nieuwe innovatieve diensten" (Digitale Agenda 9). "Binnen de overheid zijn veel data beschikbaar die relevant kunnen zijn voor ondernemers om te worden hergebruikt. Verschillende studies laten zien dat de economische waarde van de innovatie en bedrijvigheid die ontstaat op basis van open data groot is. Met open data als grondstof kunnen nieuwe toepassingen en diensten ontwikkeld en vermarkt worden" (Digitale Agenda 14). "Door inzet van open data kunnen bedrijven innovatieve toepassingen ontwikkelen met overheidsdata. … Het is nog onduidelijk hoe groot de markt zal zijn die hierdoor ontstaat, maar duidelijk is wel dat het vrij beschikbaar stellen van data voor

			economische activiteit zorgt" (Visie Open Overheid 11).
	Economic growth and stimulation of competitiveness (5-6)	Innovation (5-6)	"Stimulatie van innovaties en nieuwe bedrijvigheid" (Rapport Waterschappen en Open Databeleid 23). "Een belangrijk deel van deze informatie is openbaar op grond van de WOB. Deze overheidsdata kunnen – meestal na verrijking door marktpartijen – commerciële waarde krijgen en daarmee bijdragen aan economische groei" (Noord-Holland geeft open data de ruimte 7).
	Economic growth and stimulation of competitiveness (7-8e) Innovation (6)	The ability to merge, integrate, and mesh data (2) Development of new products and services (5e)	"Terwijl er eindeloos veel gevallen zijn van de manieren waarop Open Data reeds zowel sociale als economische waarde creëert, weten we nog niet welke nieuwe mogelijkheden er gaan komen. Nieuwe combinaties van gegevens kunnen nieuwe kennis en inzichten genereren, die weer kunnen leiden tot geheel nieuwe vormen van toepassing. Dit onaangebroken potentieel kan ontketend worden als we openbare overheidsgegevens in Open Data veranderen" (Noord-Holland geeft open data de ruimte 7).
		Development of new products and services (6e) Availability of information for investors and companies (5e)	"De provincie wil zich structureel inzetten voor de ontsluiting van Open Data… Het komende jaar ligt de focus daarbij op de Noord-Hollandse gemeenten, want zij hebben veel te bieden. Gezamenlijk hebben we een enorm aanbod aan openbare informatie, dat ontwikkelaars de kans geeft leuke, handige en praktische toepassingen te ontwikkelen. Dit draagt bij aan een transparante overheid en stimuleert innovatie en economische activiteit" (Noord-Holland geeft open data de ruimte 3).
	Innovation (4e)	Equal access to open data (2)	"Dit hergebruik kan *daardoor* veel positieve economische en maatschappelijke effecten opleveren. Dit kan alleen als gekozen wordt voor een actieve in plaats van passieve open data verstrekking" (Rapport Waterschappen en Open Databeleid 4).
	Innovation (4-5)	Equal access to open data (2e)	"Door data beschikbaar te stellen aan hergebruikers, ontstaan nieuwe toepassingen welke niet worden gerealiseerd en geëxploiteerd door de bronhouder"(Rapport Waterschappen en Open Databeleid 23).

		Development of new products and services (7e)	
		Availability of information for investors and companies (6e) Active promotion of data reuse (1) Use of the wisdom of the crowds: tapping into the intelligence of the collective (1)	"Het toegankelijk maken en beschikbaar stellen van overheidsinformatie kunnen leiden tot economische impulsen. … Het met de samenleving beleid ontwikkelen, uitvoeren en evalueren; Openstaan voor initiatieven in de samenleving; Openheid van zaken geven over haar handelen; Stimuleren om gegevens te gebruiken of te hergebruiken voor nieuwe diensten en producten" (Rapport Waterschappen en Open Databeleid 12).
	Innovation (6e)	The ability to merge, integrate and mesh data (2e)	"De economische waarde van Open Data wordt door de Europese Commissie geschat op 140 miljard euro jaarlijks. Voor Nederland is Open Data met name voor de dienstensector van groot belang. … Het ontsluiten, combineren en visualiseren van Open Data stimuleert innovatie" (Noord-Holland geeft open data de ruimte 6).
	Financial benefits (3)	Government Efficiency (1)	"Besparingen ten aanzien van het voldoen aan informatieverzoeken, besparingen op het aantal in te zetten FTE" (Rapport Waterschappen en Open Databeleid 23).
	Financial benefits (2e) Organisational benefits (1)	The ability to reuse data/not having to collect the same data again and counteracting unnecessary duplication and associated costs (also by other public institutions) (1)	"Open Data leidt tot positieve effecten in de organisatie, zowel op financieel als organisatorisch gebied. Organisaties als DUO en het PGR besparen bijvoorbeeld op het aantal uren dat medewerkers besteden aan het vervaardigen van informatieproducten" (Noord-Holland geeft open data de ruimte 9).
	Financial benefits (3e)	Transparency (1)	"Daarnaast is het de verwachting dat een grotere openheid leidt tot besparingen. Immers vreemde ogen dwingen. Openheid biedt de mogelijkheid kritisch te kijken naar de eigen

		Democratic Accountability (1)	werkwijze en bestedingen" (Visie Open Overheid 11).
	Contribution toward the improvement of processes, products, and/or services (1)	Availability of information for investors and companies (7e) Public Engagement (1)	"Aan de hand van maatschappelijke vraagstukken wordt informatie proactief beschikbaar gesteld door de overheidsorganisatie zodat burgers en bedrijven actief kunnen meedenken over de oplossing van vraagstukken" (Actieplan Open Overheid 17).

Value	Benefit category	Requirement	Quote
GOVERNMENT Value (16) (Operational & Technical)	Government Efficiency (3) Optimization of administrative processes (2)	Digitization government service and public sector information (open data) (1)	"Digitaliseren van haar dienstverlening waar toepasbaar … Deze openbare overheidsinformatie wordt voor meerdere doelen en doelgroepen gebruikt. Bijvoorbeeld door de overheid zelf om publieke taken efficiënter en effectiever uit te voeren" (Rapport Waterschappen en Open Databeleid 12-13).
	Government Efficiency (2e) Financial Benefits (2)	Transparency (3)	"Daarnaast is het de verwachting dat een grotere openheid leidt tot besparingen. Immers vreemde ogen dwingen. Openheid biedt de mogelijkheid kritisch te kijken naar de eigen werkwijze en bestedingen" (Visie Open Overheid 11).
	Government Efficiency (3e) Optimization of administrative	Use of the wisdom of the crowds: tapping into the intelligence of the collective (1) More participation and self-	"Open data biedt de kans om een kleinere overheid te realiseren waarbij de markt zelf diensten kan optuigen die goed aansluiten op de informatievragen vanuit de maatschappij zonder dat de overheid hierin hoeft te sturen of financieel hoeft bij te dragen" (Rapport Waterschappen en Open Databeleid 21).

	processes (2e) Smaller government (1)	empowerment of citizens/ users (2)	
	Financial benefits (2e)	Government efficiency (1)	"Het is ook waardevol voor de overheid zelf. Het kan bijvoorbeeld de efficiëntie van de overheid vergroten. Het kan de efficiëntie en de effectiviteit van de overheid vergroten, wat op termijn geldbesparing oplevert" (Noord-Holland geeft open data de ruimte 7).
	Legitimacy (4)	Sustainability of data or Increase in data quality (3)	"Versterking van de legitimiteit doordat de informatie die de organisatie in beheer heeft vaker wordt gebruikt" (Rapport Waterschappen en Open Databeleid 23). "De legitimiteit van de organisatie wordt versterkt doordat de informatie die de organisatie in beheer heeft vaker wordt gebruikt. Het effect van haar bestaan heeft vaker impact op de samenleving" (Noord-Holland geeft open data de ruimte 9).
		Transparency (2e) Improvement of policy-making processes (1)	"Open overheidsdata kan ook helpen om beter onderbouwde beslissingen te maken in het leven, of maakt het mogelijk actiever betrokken te zijn bij de samenleving" (Noord-Holland geeft open data de ruimte 8).
	Transparency (1)	Sustainability of data or Increase in data quality (3e)	"Door Open Data wordt niet alleen economische waarde gecreëerd, maar ook de openheid en transparantie van overheidsorganisaties vergroot. De RCE zag een enorme verbetering in de kwaliteit van hun data, door de feedback die ze na openstellingen ontvingen" (Noord-Holland geeft open data de ruimte 9).
	Increase External orientation (2)	Collaboration with chain partners (2)	"Vergroting externe oriëntatie, door het kiezen van een Open Databeleid worden de waterbeheerders gedwongen om zich meer naar buiten te richten en meer samen te werken met ketenpartners" (Rapport Waterschappen en Open Databeleid 23). "Door voor een Open Data strategie te kiezen, kan een organisatie gedwongen worden om naar buiten te kijken en meer samen te werken met ketenpartners" (Noord-Holland

			geeft open data de ruimte 9).
	Improvement or new government services for citizens and companies (1)	More participation and self-empowerment of citizens (2e) Equal access to data (1)	"Betere dienstverlening aan burgers en bedrijven. Op basis van de vrijgegeven data kunnen toepassingen ontwikkeld worden die een uitbreiding of verbetering zijn ten aanzien van de al bestaande dienstverlening" (Rapport Waterschappen en Open Databeleid 23).
	Easier to reuse data by other parties (1)	Interoperability (1)	"Wanneer meer data volgens (open) standaarden ontsloten worden, kunnen andere overheidspartijen deze data ook eenvoudiger hergebruiken. De interoperabiliteit van de overheid wordt hierdoor vergroot" (Noord-Holland geeft open data de ruimte 9).
	Creation of trust in government (1) Legitimacy (4e)	Transparency (3e)	De overheid dient niet transparant te zijn 'omdat het moet', maar omdat zij dat in de context van de democratische rechtstaat van nature is. Transparantie is daarmee een wezenskenmerk van de overheid, een conditio sine qua non om het vertrouwen in de overheid te bevorderen (Visie Open Overheid 9).
	Increase Data Quality (2)	External quality checks of data (2) Scrutinization of data(2)	"Van belang is te beseffen dat door de data beschikbaar te maken en je een goede terugmeldfaciliteit inbouwt, de kwaliteit van de data eenvoudig kan worden vergroot" (Rapport Waterschappen en Open Databeleid 5). "Verhoging van de datakwaliteit, de waterbeheerders worden gedwongen om de eigen datakwaliteit te verhogen alvorens data naar buiten wordt gebracht" (Rapport Waterschappen en Open Databeleid 23).
	Easier to reuse data by other parties (2)	Active Open data policy (1)	"Eenvoudiger hergebruik van data door andere partijen" (Rapport Waterschappen en Open Databeleid 23)

Value	Benefit category	Requirement	Quote
PROGRESS/ POSITIVISTIC (4 CITATEN)	Positive effects (2)	Active promotion of data reuse (4)	"Vanwege de positieve effecten die het hergebruik van overheidsinformatie kan opleveren, is de Nederlandse overheid er toe overgegaan het hergebruik van openbare overheidsinformatie actief te bevorderen" (Rapport Waterschappen en Open Databeleid 13). "Vanwege de positieve effecten die het hergebruik van overheidsinformatie kan opleveren is de Nederlandse overheid er toe overgegaan het hergebruik van openbare overheidsinformatie actief te bevorderen" (Visie Open Overheid 14).
	Improve competitiveness (2)	Active promotion of data reuse (3-4) Connect with OGD movement (2)	"De ontwikkelingen gaan dusdanig snel dat Nederland op relatief korte termijn tot de achterhoede kan gaan behoren wat betreft het breder benutten van overheidsdata, als ze daar niet in mee gaat. Daarom is de Nederlandse overheid er toe overgegaan het hergebruik van openbare overheidsinformatie actief te bevorderen" (Rapport Waterschappen en Open Databeleid 4). "De vorming van een beweging als het Open Government Partnership en de snelle groei van dit partnerschap (binnen een jaar naar 55 landen) laat zien dat wereldwijd besef groeit dat overheden aansluiting moeten zoeken bij de beweging naar meer openheid in de samenleving" (Visie Open Overheid 7).

Coding Frame Barriers

Legend:

Between brackets (..): number of quotes.

Red: Indicators which are more frequently quantified

Green: Categories or indicators deviating from Janssen et al. (2012).

Level	Barriers	Quotes[30]
INSTITUTIONAL (15)	Risk averse culture (no entrepreneurship/ recalcitrance to change) (3)	"Overheden zijn in de kern verticaal gerichte organisaties: ambtenaren werken voor de samenleving, maar in opdracht van de minister. Vanwege de ministeriële verantwoordelijkheid kan er terughoudendheid bestaan om al in een vroeg stadium van gedachten te wisselen met belanghebbenden. Om de trend tot meer horizontalisering door te laten zetten, is een cultuurverandering nodig die door de politieke en ambtelijke top wordt ondersteund of in gang gezet, en openlijk wordt gewaardeerd" (Visie Open Overheid 9). "Stel geen voorwaarden aan diegenen die de data gaan hergebruiken" (Noord-Holland geeft Open data de ruimte 16). "Omarm initiatieven die in de buitenwereld ontstaan. Diensten en apps die door startups worden ontwikkeld kunnen ook van meerwaarde zijn binnen de eigen organisatie, en de efficiency en effectiviteit vergroten" (Noord-Holland geeft Open data de ruimte 19).
	Financial costs too high (4)	"De huidige informatiesystemen maken nog onvoldoende gebruik van de technologische mogelijkheden, maar aanpassingen hebben forse consequenties voor de werkwijze en brengen grote investeringen met zich mee" (Visie Open Overheid 9).

[30] The barriers as mentioned within the document "Noord-Holland geeft open data de ruimte" become clear by applying a 'reversal strategy'. Within this document, the authors have devised a plan of *"how to get started with Open Data"*. The authors devised several imperatives that act as a guidance to get started. Reversing these imperatives means that the 'hidden' barriers to 'start with Open Data' are disclosed.

		"Het realiseren van openheid zal echter ook investeringen en kosten met zich meebrengen die in tijden van bezuinigingen niet vanzelfsprekend zijn. Daarom is het belangrijk om stapsgewijs toe te werken naar meer openheid" (Visie Open Overheid 11). *"Maar ook hierbij geldt, niet alles wat mogelijk is moeten we willen. Zo moeten we alert zijn op de gevolgen voor de bedrijfsvoering en met name de kosten. In het huidige participatiebeleid ligt de nadruk op transparantie. Maar als we transparantie als extra service bieden, bovenop onze normale bedrijfsvoering, dan verhogen we onze kosten en creëren we een omgeving die moeilijk te borgen is"* (Actieplan Open Overheid 8). "Opvallend tijdens de verschillende contactmomenten is dat enerzijds men terughoudend is ten aanzien van een Open Databeleid. Deze terughoudendheid richt zich met name op het werk en de kosten van Open Data en het niet voldoende inzicht hebben van de voordelen van Open Data" (Rapport Waterschappen en Open Databeleid 5).
	No uniform policy for publicizing and maintaining data (3)	"Vaststellen dat de beschikbare technologie voorhanden is om informatie pro actief beschikbaar te maken betekent niet dat die direct wordt toegepast… Daarom moet nu worden uitgewerkt welke ontwerpprincipes in de toekomst moeten worden toegepast om bij geplande vernieuwing of aanpassing van informatiesystemen actieve openbaarmaking eenvoudiger te maken (Visie Open Overheid 10). "Breng in kaart welke data jouw organisatie ter beschikking heeft en inventariseer welke data ontsloten kan worden" (Noord-Holland geeft Open data de ruimte 16). "Zorg dat de continuïteit van de data gewaarborgd is. Het onderhoud en ter beschikking stellen van de data dient structureel belegd te zijn binnen de organisatie" (Noord-Holland geeft Open data de ruimte 19).

	Unclear trade-off between public values (transparency vs. privacy values) (1)	*"Maar ook hierbij geldt, niet alles wat mogelijk is moeten we willen. Zo moeten we alert zijn op de gevolgen voor de bedrijfsvoering en met name de kosten. In het huidige participatiebeleid ligt de nadruk op transparantie. Maar als we transparantie als extra service bieden, bovenop onze normale bedrijfsvoering, dan verhogen we onze kosten en creëren we een omgeving die moeilijk te borgen is"* (Actieplan Open Overheid 8).
	Loss of public image (1)	"Data zijn beschikbaar gesteld voor hergebruik terwijl dat niet mocht" "Er is wel een gevaar van bestuurlijk imagoverlies dat nadelige gevolgen kan hebben" (Rapport Waterschappen en Open Databeleid 18-19).
	No internal support for publishing open data (3)	"Creëer intern draagvlak op zowel bestuurlijk als ambtelijk niveau" (Noord-Holland geeft Open data de ruimte 16).
	Unclear value: data* may appear to be irrelevant or benign when viewed in isolation, but when linked and analysed collectively it can result in new insights. (1) No internal support for publishing open data (2-3[e]) * (Janssen et al. 2012 use the word information, however data is not necessary information for everybody)	Kijk welke datasets eenvoudig open te stellen zijn. Soms kunnen simpele bestanden als katalysator fungeren; Kijk goed om je heen. Welke datasets worden door andere (vergelijkbare) organisaties vrijgesteld? Dit kan nieuwe ideeën en inspiratie opleveren (Noord-Holland geeft Open data de ruimte 16). "Opvallend tijdens de verschillende contactmomenten is dat enerzijds men terughoudend is ten aanzien van een Open Databeleid. Deze terughoudendheid richt zich met name op het werk en de kosten van Open Data en het niet voldoende inzicht hebben van de voordelen van Open Data" (Rapport Waterschappen en Open Databeleid 5).

LEGISLATION (8)	Privacy violation (3)	Vaststellen dat de beschikbare technologie voorhanden is om informatie pro actief beschikbaar te maken betekent niet dat die direct wordt toegepast. Naast infrastructurele veranderingen, vergt dit mogelijk ook verandering in wetgeving en cultuur. Soms is openbare informatie als het ware verweven met niet-openbare informatie (Visie Open Overheid 9). "Privacy moet ten alle tijden worden beschermd. Het is belangrijk de data op zo'n niveau te aggregeren dat deze niet herleid kunnen worden tot personen" (Noord-Holland geeft Open data de ruimte 18).

		"Data zijn beschikbaar gesteld voor hergebruik terwijl dat niet mocht" (Rapport Waterschappen en Open Databeleid 19).
	No License for using data (1) Limited conditions for using data (1)	Stel geen voorwaarden aan diegenen die de data gaan hergebruiken (Noord-Holland geeft Open data de ruimte 16).
	Dispute and litigations (4)	"Stel zeker dat u eigenaar bent van de data en er geen sprake is van rechten van Derden" (Noord-Holland geeft Open data de ruimte 16). "In bepaalde situaties kunnen overheden aansprakelijk gehouden worden voor de door hen geleverde producten en diensten, waaronder Open Data. De aansprakelijkheid ten aanzien van Open Data wordt beschreven in het onderzoek van Marc de Vries 'Van Erik Engerd naar J. J. de Bom'. Gebleken is dat door een goede beschrijving en uitleg van de data de aansprakelijkheidsrisico's beperkt zijn" (Noord-Holland geeft Open data de ruimte 18). "Data zijn beschikbaar gesteld voor hergebruik terwijl dat niet mocht" (Rapport Waterschappen en Open Databeleid 19). "Koud watervrees om aansprakelijk te worden gesteld ten aanzien van de data die wordt vrijgegeven is ook bij de waterbeheerders aanwezig. Uit recent onderzoek zoals genoemd in paragraaf 3.7 Aansprakelijkheid en Open Data blijkt dat de risico's ten aanzien van aansprakelijkheid gering zijn wel zijn voorzorgsmaatregelen noodzakelijk" (Rapport Waterschappen en Open Databeleid 21).

TASK COMPLEXITY (7)	No explanation of the meaning of data (1) No information about the quality of the open data (see category "Information Quality") (2)	"Bij de informatie is het van belang dat deze niet alleen "open" is, maar dat juist de samenhang in de informatie voor publiek duidelijk wordt" (Actieplan Open Overheid 17). "Leg goed uit met welk doel de data verzameld zijn; Leg de betekenis van de data goed uit" (Noord-Holland geeft Open data de ruimte 16).

	No access to the original data (only processed data) (1)	"Geef data in een zo'n ruw formaat als mogelijk vrij, zonder bewerkingen" (Noord-Holland geeft Open data de ruimte 17).
	No Tools or support helpdesk (2)	"Zorg ervoor dat er een duidelijk contactpersoon is voor vragen over de datasets" (Noord-Holland geeft Open data de ruimte 17). "Het vrijgeven van Open Data kan leiden tot extra vragen aan de waterbeheerders door eenieder die de data gebruikt. Is de organisatie hierop ingericht?" (Rapport Waterschappen en Open Databeleid 21).
	Difficulty in searching and browsing due to no index or other means to ensure easy search for finding the right data (2)	"Met open data als grondstof kunnen nieuwe toepassingen en diensten ontwikkeld en vermarkt worden. Overheidsdata zijn vooralsnog echter vaak niet 'open' omdat ze bijvoorbeeld niet vindbaar zijn of niet door computers kunnen worden gelezen" (Digitale Agenda.nl 14). "Data die de waterbeheerders kunnen gaan vrijgeven als Open Data is verzameld in het kader van een overheidstaak en dus binnen een context en met een betekenis. Derhalve moet de publicatie van data altijd vergezeld gaan van een beschrijving van deze context en de betekenis. Het meegeven van meta-data zorgt er voor dat de data ook vindbaar is" (Rapport Waterschappen en Open Databeleid 21).

TECHNICAL (6)	Data must be in a well-defined format that is easily accessible: while the format of data is arbitrary, the format of data definitions needs to be rigorously defined Absence of standards	Met open data als grondstof kunnen nieuwe toepassingen en diensten ontwikkeld en vermarkt worden. Overheidsdata zijn vooralsnog echter vaak niet 'open' omdat ze bijvoorbeeld niet vindbaar zijn of niet door computers kunnen worden gelezen (Digitale Agenda.nl 14). Maak gebruik van open standaarden, zoals XML en CSV, in plaats van PDF en XLS. Hiermee wordt de data 'machineleesbaar' en kan er eenvoudig mee gewerkt worden; (Noord-Holland geeft Open data de ruimte 17). "Vind aansluiting bij de reguliere ICT-infrastructuur. Hiermee wordt de datavoorziening een onderdeel van de reguliere processen" (Noord-Holland geeft Open data de ruimte 19).
	Lack of meta standards	"Maak een duidelijk beschrijving per dataset, waarin bijvoorbeeld de betekenis van rijen en

		kolommen wordt toegelicht en hoe vaak de data vernieuwd worden (metadata)" (Noord-Holland geeft Open data de ruimte 17). "Data die de waterbeheerders kunnen gaan vrijgeven als Open Data is verzameld in het kader van een overheidstaak en dus binnen een context en met een betekenis. Derhalve moet de publicatie van data altijd vergezeld gaan van een beschrijving van deze context en de betekenis. Het meegeven van meta-data zorgt er voor dat de data ook vindbaar is" (Rapport Waterschappen en Open Databeleid 21).
	No central portal or architecture, No support for making data available	"Kies een geschikte plek om de data te ontsluiten, zoals de centrale website, of via een API" (Noord-Holland geeft Open data de ruimte 17).

USE AND PARTICIPATION (5)	Public organizations do not react to user input (2)	"Zorg ervoor dat er een duidelijk contactpersoon is voor vragen over de datasets" (Noord-Holland geeft Open data de ruimte 17). "Wanneer er interessante en waardevolle data ontsloten zijn, zijn er verschillende manieren waarop startups, bedrijven, ondernemers en studenten hierop geattendeerd kunnen worden" (Noord-Holland geeft Open data de ruimte 20).
	No incentives for the user (2)	"Omarm initiatieven die in de buitenwereld ontstaan. Diensten en apps die door startups worden ontwikkeld kunnen ook van meerwaarde zijn binnen de eigen organisatie, en de efficiency en effectiviteit vergroten" (Noord-Holland geeft Open data de ruimte 19). "Wanneer er interessante en waardevolle data ontsloten zijn, zijn er verschillende manieren waarop startups, bedrijven, ondernemers en studenten hierop geattendeerd kunnen worden" (Noord-Holland geeft Open data de ruimte 20).
	Threshold to access data: (Having to pay a fee for the data, Registration required before being able to download the data) (1)	"Denk eraan geen onnodige drempels op te werpen om toegang te krijgen tot een dataset, zoals registratie of inlogprocedure" (Noord-Holland geeft Open data de ruimte 17).

DATA QUALITY (4)	[Essential] data is missing Lack of data Lack of accuracy of the data	"Stel zeker dat de datasets zo compleet mogelijk zijn" (Noord-Holland geeft Open data de ruimte 17). "Een belangrijk risico dat rondom het open maken van data vaak genoemd wordt is de- soms door de waterschappen zelf als ontoereikend beschouwde - kwaliteit van de data" (Rapport Waterschappen en Open Databeleid 5). "De voor hergebruik beschikbaar gestelde data zijn gebrekkig" (Rapport Waterschappen en Open Databeleid 20-21). "De data moet in ieder geval voldoen aan de volgende aspecten: Volledigheid; Correct, data moet juist zijn; Actueel; Geldigheid, in tijd; Consistent, gebruik standaardisatie; Betrouwbaarheid; Tijdigheid, hoe snel is de data op te vragen/ te benaderen?" (Rapport Waterschappen en Open Databeleid 20-21).
INFORMATION QUALITY (2)	Incomplete information, only part of the total picture shown or only a certain range [Essential] Information is missing Lack of information	"Leg goed uit met welk doel de data verzameld zijn; Leg de betekenis van de data goed uit" (Noord-Holland geeft Open data de ruimte 16). "Overheden gebruiken en genereren data voor de uitvoering van hun publieke taak. Wanneer deze data voor andere doeleinden worden gebruikt, kan het zijn dat de data minder geschikt zijn. Dat kan vanwege kwaliteit of volledigheid zijn. Het is belangrijk om goed aan te geven met welk doel de data binnen de overheid gebruikt worden. Daarmee wordt voor de hergebruiker ook aangegeven welke nauwkeurigheid van de data verwacht mag worden" (Noord-Holland geeft Open data de ruimte 18).
ECONOMIC (1)	Unfair competition (1)	"De gratis verstrekking veroorzaakt oneerlijke concurrentie met marktpartijen" (Rapport Waterschappen en Open Databeleid 18).