Delft University of Technology

A comparison of open data policies in different countries

Lessons learned for an open data policy in Indonesia

Rininta Putri Nugroho (4181859)

Master Thesis

2013

Faculty of Technology, Policy and Management

A comparison of open data policies in different countries

Lessons learned for an open data policy in Indonesia

MASTER THESIS
Submitted in partial fulfillment of the requirements for the degree of
Master of Science in Engineering and Policy Analysis

Author : Rininta Putri Nugroho
Student number : 4181859

Graduation Committee:
Chairman: Prof. Dr. Yao-Hua Tan
Professor, Head of Section Information and Communication Technology

First supervisor: Prof. Dr. ir. M. F. W. H. A. Janssen
Section Information and Communication Technology

Second supervisor: Prof. Dr. W. M. de Jong
Section Policy, Organization, Law and Gaming

Third supervisor: Anneke Zuiderwijk-van Eijk, M. Sc
Section Information and Communication Technology

MSc Program Engineering and Policy Analysis
Faculty of Technology, Policy and Management
Delft University of Technology, the Netherlands

Abstract

Many countries around the world have joined the open data movement. Data is being published for a various number of reasons which include for the public to reuse, to create a more efficient government, and to increase transparency. Recent development in this field is that the data published is expected to be in machine readable format. In general, there is a lack in guidelines to regulate and help the process of opening data. Many countries are in different stages in developing these guidelines. Indonesia is an example of a country just beginning to join the open data movement. A field of study that is lacking is about how countries can learn from each other in developing the necessary guidelines. Being in the early stages of development, Indonesia can especially benefit from research in this area.

A complex comparison of open data policies is conducted in this research to provide a basis for drawing conclusions and recommendations for the open data policy in Indonesia. For this study, five different countries that are in different stages of development in their open data initiatives are explored which results in an extensive list of findings. These countries include the United States, United Kingdom, Netherlands, Kenya and Indonesia. For the design of the framework, literature and case studies are conducted. The case studies are in the form of interviews with eight respondents involved in open data in each country. First, it is identified what aspects influence the uneven development of open data. Second, lessons that are relevant for Indonesia based on the many similarities and differences are identified. As the scientific contribution of the research, this framework and comparison is given because there is currently lacking research in this area.

It was concluded that Indonesia can synthesize a number of lessons from the comparison that comprises a combination of elements that were presented as findings from each of the countries. The lessons that were developed include suggestions for a more robust legal framework, the creation of an ecosystem between data publishers and data users, the development of stronger IT and organizational support for open data, and the launch of initiatives that use open data at the district government levels. Interestingly, from the study, it suggests that the focus of the policies for countries in the developing stages are more related to the release of data from the publishers and less on the technical processes that are involved with opening the data. Other interesting results that were founded from the study suggest that the difference between the countries are influenced by specific forces and counter forces in the area of open government and also from the existence of individuals that highly advocate for the development of open data in that country. So, the practical contribution of this research is the lessons that each of the countries can derive from the comparison and also the specific lessons that are designed for Indonesia's open data policy both to reap the identified benefits of opening data.

Further to this research, there is a possibility to conduct research on further comparisons for a more robust and comprehensive learning process. Comparisons can be conducted on different countries about open data policies or about different aspects of open data itself. Another possibility is to create lessons for all countries involved in this current study.

Key words: *open government data, comparison framework, open data framework, open data policy, policy analysis, e-government*

Foreword

At last, the two year journey is complete and I would not be here without the help and support from a long list of people that have been with me from the start. First of all I would like to extend my gratitude towards my thesis committee, for their support, valuable insights, and patience in assisting me to finish this study. I would like to thank the chairman of my thesis committee, Prof. Dr. Yao-Hua Tan, for giving me the chance to complete this thesis project within the ICT section. A large thank you also to Prof. Marijn Janssen, whose patience and vast knowledge on the open data field has led me to gain genuine interest on the topic. I also want to give my special gratitude to my second supervisor, Prof. Martin de Jong, whose insight about different cultures and behaviors is astonishing and has helped greatly in providing the perceptions to keep my research going. I would also like to express my gratitude to Anneke Zuiderwijk-van Eijk, M.Sc that has given priceless guidance and provided me with a hand to hold throughout the process. The whole thesis committee has worked together with me in reaching this point and I am grateful to have them.

As part of the committee, although not officially, I would like to thank my thesis buddy Maaike Kaasenbrood that has gone through the exact same process and been my shoulder to lean on in times of doubt and my motivator in times of deadline-making. I am glad to have been able to spend the last five months working together with you. Special recognition also given to various individuals that have joined discussions and been sound boards for my brainstorming as other thesis buddies, thank you for your insights.

Recognition is also given to the Ministry of Communication and Information Technology of Indonesia that has given me the chance to come to the TU Delft and complete my studies under their scholarship. A big warm thank you goes out to all my friends that I have met along the way here in the Netherlands, especially fellow survivors of EPA, MoT, and SEPAM at TPM. Thank you for the extraordinary friendship and experience during the last two years. A special thank you goes out to my “Indonesian Mafia” that has showed unconditional love and support as my second family in Delft.

Finally I would like to show my love to my amazing family, without them I would not be here at all. They have shown unconditional love and support through my whole two year journey in the Netherlands. I owe them everything. Above all and the most importantly, thanks to God the Almighty for each of His blessings.

Delft, August 2013

Rininta Putri Nugroho

Executive summary

The field of open data can be seen as a new trend in the world of information and communication technology. Especially with the technological advancements to support the opening of data, more and more countries are interested in opening their governmental data to the public. Apart from the technological advancements, many benefits and positive impacts that have been identified also play a role in the eagerness of countries to implement open data programs. Some of the discovered benefits include transparency and accountability of the government, participation and self-empowerment to the citizens, economic growth and also stimulation of innovation through re-use of data (Janssen et al., 2012, Lundqvist, 2012, Tauberer, 2009, Huijboom and Broek, 2011, Zhang et al., 2005, Zuiderwijk et al., 2012). However there is an uneven development amongst the countries. Some countries like Australia, Denmark, Spain, the United Kingdom (UK), and the United States (US) are seen to be more advanced (Huijboom and Broek, 2011).

In relation to the difference of development of open data policies in the several countries, it is observed that Indonesia is a country that is still considerably behind. Although Indonesia has enacted a freedom of information law in 2008, there has been little progress since then to further regulate the opening of data. At the moment Indonesia only has a beta version of an open data platform (satupemerintah.net). However, Indonesia is widely supported by the international community to further progress in terms of open government and this is responded by action plans that are published on Open Government Indonesia. With the uneven progress of open data in different countries it is possible to learn from each other to develop a better program. Indonesia can benefit from these studies to develop their own open data policy. A method that can be applied to formulate better policies and address some of the current barriers to developing open data policies is by comparing the policies and creating a benchmark which is the problem that is addressed in this research.

The objective of the research in this thesis is to investigate what Indonesia can learn from the open data policies of the United Kingdom, the United States of America, the Netherlands, and Kenya in developing its own open data policies. By doing so, this research also presents a thorough comparison of the open data policies between the countries.

At the moment there is limited literature available on comparisons of open data policies especially in comparing different countries. This leads to the scientific contribution of the research, to provide a comparison framework that is used to compare different countries. The framework will include elements that are specifically relevant to include when comparing national open data policies such as the cultural context and also country demographics. The contribution for practice lies in the lessons that are derived from the comparison and also the findings from the comparison itself. The lessons presented in this research are explicitly targeted for the improvement of Indonesia's open data policy, but it contributes to the efforts of all involved countries to gain the benefits from opening government data.

In order to achieve the objectives and targeted contribution, the following research questions are formulated:

Main question:

What can Indonesia learn from the open data policies of the United Kingdom, the United States of America, the Netherlands, and Kenya in developing its own open data policy?

Sub-questions:

1. Which elements of open data policies should be compared to develop recommendations according to available literature?
2. What are the forces and counter-forces that drive and hinder the open data policies in the examined countries?
3. What are the differences and similarities between the countries when compared using the comparison framework that is based on the elements identified for the examined countries?

The methodology applied to answer the questions include a literature research to find previous comparison frameworks on both open data policies and other policy making processes and research on open data policies, a literature research is conducted to acquire knowledge about the various countries that are examined, lesson drawing from the results of the comparison based on Rose's (2002) ten steps, and interviews. A total of eight interviews were conducted for the purpose of the research.

For the purpose of the research, open data is defined as data that meets the criteria of being accessible at no more than the cost of reproduction, without limitations based on user identity or intent, in a digital format and free of restriction on use or redistribution in its licensing conditions. It is also important to examine the benefits and barriers of opening data and not only the definition. This is because they lead to the development of relevant policies that highlight the benefits and help mitigate the effects of the barriers. The formulation of policies is highly dependent on the existing benefits and barriers.

Based on the literature findings, the framework is sub-categorized into the elements that build the policy context and environment, policy content, and policy impact. The research was conducted to compare open data policies in the United States, United Kingdom, Netherlands, Kenya and Indonesia. The elements included in the framework can be seen in the following table:

Element category	Element	
Policy context	Level of government organization	
	Key motivations, policy objectives	
	Open data platform launch	
	Resource allocation and economic context:	Technology penetration
		ICT Infrastructure
		Source of Funding
	Legislation	
	Social and political contexts	Political constituencies
		Hofstede – Individualism and collectivism
		Minkov – Exclusionism and universalism
		Trompenaars – Particularism and universalism
		Cultural beliefs
	Drivers or forces for opening data	Event based
		Various forces
	Counter forces that hinder the opening	
Policy Content	Technical aspects of the open data process	Licensing (LP)
		Fees for access (FB)
		Data presentation (LP)
		Restricted data (LP)
		Contact with data users (CB)
		Amount of published data
		Processing of data before publishing (T)
		Costs for opening

Element category	Element	
	Technical aspects of the data	Types of data
		Data format and standards (T)
		Data quality (T)
		Provision of metadata (T)
		Interoperability with other data (T)
		Accessibility of data (T)
		Encouragement for data re-use (CB)
Policy impact	Re-use of published data	
	Possible predicted risks	
	Benefit alignment with motivation	
	Public values	

Furthermore findings from the interviews and cultural aspects showed different drivers for open data. Some of the findings that can be concluded as forces and counter forces include the need for more evidence of the benefits of opening data in order to stimulate further publishing of data, opening data as a big change in mindset in all of the observed countries, individuals that advocate for the development of open data are important in determining the development and progress of the open data movement, there needs to be a large diversity of data types that are published, data must be understandable for all segments even if it is already in the form of an application, a balance is needed between being completely open and regulations, and lastly a strong foundation for open data must be laid before implementation. At the moment it can be observed that one of the most important barriers for countries to publish data is the need for more evidence of the benefits. There already exists many studies that predict the benefits of open data but data publishers are seeking evidence of the effects before they are willing to participate further. On the other hand, an important driver that is observed to stimulate the process of open data in the countries is the existence of individuals that make an effort to see open data work in that country. Each country has a certain individual that consistently encourages the progress of open data. All of the findings that were categorized as these forces and counter forces are considered to be unique findings that hold importance to the reason why open data can or cannot flourish in a country.

From the comparison that was conducted certain aspects can also be summarized. In comparing the policy context it can be found that the UK, US and Netherlands, countries which have a stronger tendency to be open, have in common a supporting IT infrastructure, high technology penetration, government type (constitutional), higher GDPs, and cultural dimensions (individualists according to Hofstede (2010) and universalists according to Minkov (Hofstede et al., 2010) and Trompenaars (1998)) Furthermore the results of the policy content comparison suggest that, except for Indonesia, all the countries have already placed certain guidelines for the open data process even though not all of them are formalized as policies. In terms of the policy impact it can be concluded that all the observed countries are more focused on the process of releasing the data than the reuse of the data.

From the results of the comparison, certain lessons were developed for Indonesia's open data policy. These lessons were synthesized from the current practices that are regulated in the other countries and also the various forces and counter forces that were established from the previous findings. Because of the specific nature of these lessons for Indonesia, the following lessons were chosen in consideration of the

environmental context under which they will be implemented. Hence, based on the steps of lesson drawing, the lessons that are designed for Indonesia are the following:

1. A more robust legal framework that ensures the continuous release of data from the data publishers. The legal framework should include the minimum number of datasets that should be published on a regular basis and regulation of the privacy issues involved with opening data on a publicly accessibly platform.
2. Operational policies that cover more aspects of the open data process such as machine readability of the data and the accessibility of the datasets on an open data platform without registration.
3. Creating an ecosystem between data publishers and data users by maintaining contact.
4. Ensuring quality of the published data through specific information quality acts.
5. Creating an open data support system through organizational and infrastructural changes. Organizational changes include the designation of open data to a specific agency centrally or specific organizational entity within each data publisher's organization that is responsible for the published data. Infrastructural changes in this case are related to the development of a more supportive IT infrastructure and also building of an open data platform.
6. Creating initiatives at the district levels of the government to trigger more demand for data.

Other conclusions that can be drawn from the research conducted is that at the moment there is more urgency in developing legal frameworks that ensure the continuous release of data from the government compared to the policies regarding the requirements of the data itself. This holds true in the countries that are considered to be less advanced with their open data initiatives such as the Netherlands, Kenya and Indonesia. However, attention should still be paid to the machine readability of the data that is published which will pose an easier transition when the policies for the data are enacted as well. A recommendation that is suggested as further research is to identify the benefits that have materialized from the current published data and identify the impact of opening data thus far. This is especially relevant in countries that have been involved with open data for longer with enough data already published.

Existing research in the field of open data is often focused on the usage of open data or other activities that can benefit from the usage of open data. This research is focused on guidelines that enable the usage of open data which are identified to be open data policies. Comparisons of open data programs and open data policies have been completed in previous research (Zuiderwijk and Janssen, (to be published), Huijboom and Broek, 2011, Rothenberg, 2012) but on a different scale and complexity. Because the scope of the research compares open data policies in different countries, this creates an international setting to the analysis.

The findings in this thesis also have practical implications for the countries that are compared (US, UK, Netherlands, Kenya, and Indonesia) in terms of continuing the development of their open data programs. For each of the countries, this thesis can be used to draw lessons and also to observe the strengths and weaknesses of their current situation. The research that is presented in this thesis adds to the existing research that is available in guiding organizations, agencies, and countries in reaching their objectives of opening data to the public.

Table of Contents

List of figures

List of tables

Chapter 1
Introduction

Starting from the Freedom of Information Act (FoIA) that has existed since 1966 in the US and since 2000 in the UK, countries have tried to create a culture where the government publishes their data to the public. This act encourages governmental agencies to release their data to the public as requested. A newer concept called open data is a progression of this act that emphasizes the need for the data to be in formats that are easily re-usable and machine readable. Also with the advancement in infrastructure to support the opening of data, such as the World Wide Web and the Internet, the process of opening data has become the focus of some organizations and countries alike. Many can pinpoint the start of the open data trend in 2009 when President Barack Obama's administration announced the Open Government Directive (Obama, 2009) which motivated other countries to take more concrete actions to deliver similar directives. The European Union Public Sector Information directive in 2003 can also be called a start to the opening of data in Europe. On the global front, the Open Government Partnership (OGP) was launched in September 2011 as a global effort to make governments better (Open_Government_Partnership, 2011b). The eight founding governments include Brazil, Indonesia, Mexico, Norway, Philippines, South Africa, United Kingdom, and United States that endorsed the Open Government Declaration and announced their country actions plans. Since then the OGP has welcomed the commitment of 47 additional governments to join the Partnership. One of the commitments of the OGP is to increase the availability of governmental information that is collected on behalf of the public.

The field of open data can be seen as a new breakthrough in the world of information and communication technology. Open data is seen to be the concrete step towards the hopes of an open government. Research into the field of open government and open data has increased in the last years as many are looking for answers to how open data can help create a better government for the citizens. The reason why open data is seemingly an attractive practice is that many benefits and positive impacts have been identified. Some of the discovered benefits include transparency and accountability of the government, participation and self-empowerment to the citizens, economic growth and also stimulation of innovation through re-use of data (Zuiderwijk et al., 2012, Lundqvist, 2012, Janssen et al., 2012, Tauberer, 2009, Zhang et al., 2005, Huijboom and Broek, 2011). These benefits seem to draw more and more countries to be more actively involved in opening their data and in turn also being a more open government. For instance the Open Government Directive in US and the EU PSI directive which calls for attention to the principles of transparency, participation and collaboration (European_Commission, 2003). This allows citizens to monitor performance of their government and hold the government accountable for their actions. On the other hand, the aim of the OGP is to promote transparency, empower citizens, fight corruption, and harness new technologies to strengthen governance (Open_Government_Partnership, 2011b).

Concrete actions of open data from most countries include the launching of various open data portals as point of access for the citizens to retrieve the published data. In response to the Open Government Directive in the US, the data.gov portal was launched in 2009 by Federal Chief Information Officer as the national open data portal. In the UK, the data.gov.uk portal was officially launched in January 2010 and has since developed into a repository of 9000 unique datasets. Meanwhile a country that has recently joined the movement is the Netherlands by launching their national open data portal data.overheid.nl in September 2011. Slightly trailing behind these examples is Kenya (opendata.go.ke) in 2011 with their national open data

portal and Indonesia with a beta version (satupemerintah.net). Although the Kenyan open data portal already contains some datasets, they are not enough to stimulate reuse.

In light of the global movements towards a more open government this research is focused on the open data policies that help provide guidance for the opening of data in different countries. This research thesis identifies in detail the steps, approaches, and results of the research project in the following thesis report.

a. Problem description and research objective

As the release of open data becomes more of a common practice in some countries, open data policies have been developed to provide stimulation and guidance (Zuiderwijk and Janssen, 2012b). Countries that are considered to have an established open data policy include Australia, Denmark, Spain, the United Kingdom (UK), and the United States (US) although each country has different focuses for opening data (Huijboom and Broek, 2011). It can be observed that the countries indicated above are commonly known as western countries or more developed countries. There is no definite answer as to why the phenomenon occurs where policies are more developed in developed countries. Previous research on the openness of government, which was conducted on government websites, states that openness is highly correlated to the size of the country and also to national wealth (La Porte et al., 2002). Although it is also argued that these factors are not definitive in defining whether or not a government is open, the majority of countries follow this rule. In other previous research, it is stated that open data is still in its very early stages in developing countries with very limited data and access to the available data (Schwegmann, 2012). However, it is still acknowledged that developing countries can also benefit greatly from initiating open data programs and these countries are also enthusiastic about participating (Schwegmann, 2012). An example of a country in this situation is Indonesia. In 2008, Indonesia launched the policy to open public data which was the first step towards a more open government. As a next step to the policy, Indonesia recently launched the 'Open Government Indonesia' (OGI) website (opengovindonesia.org) in 2012 that contains information about the OGI initiative and its actions plans. Little is mentioned about open data specifically as it is more generalized. Even though progress has been made, the open government and open data initiative has yet to fully launch which is especially seen by the duration needed between the policies until the portal launch. The goal of the Indonesian government in launching this program is to initiate a change in the bureaucracy between sectors and institutions, to improve transparency for the public which in turn leads to identification of ineffective and inefficient processes or individuals, and also to improve the overall service that is provided to the public (OGI,2012). However, from initial study on the open data movement that has begun in Indonesia, it is evident that improvements could still be made in terms of their open data policies. This thesis will present lessons that can be used in the policy making process of Indonesia's open data policies that is derived from more established open data programs in leading countries.

With the uneven progress of open data in different countries it is possible to learn from each other to develop a better program. A method that has been applied to formulate better policies and address some of the current barriers with developing open data policies is by comparing the policies and creating a benchmark. Countries with less developed open data policies can gain insight from the comparison and improve their own policies. The research objective of the research in this thesis is to investigate what Indonesia can learn from the open data policies of the United Kingdom, the United States of America, the Netherlands, and Kenya in developing its own open data policies.

At the moment this method has been applied, through research, in the open data field in different governmental institutions in the same country or on a very generic level (Zuiderwijk and Janssen 2012,

Huijboom and Broek, 2011). Even though it has been identified that it is beneficial to learn from comparisons of the policies, there are still many unaddressed issues when looking at a cross-country comparison of open data policies. The added value of this research is that it will be a cross-country comparison that will not only include important open data policy elements in the comparison but also country specific characteristics that influence the policy making process. These characteristics include cultural dimension, deep rooted beliefs, political culture, and also major forces and counter forces that motivate or hinder the opening of data. In further detail, the research will identify elements of open data policies that should be compared in the framework. This will be based on an explorative study into open data policies and also existing frameworks to derive the elements. These elements can be argued to be unchangeable variables as they are variables that have been implanted from cultural upbringing. However it is founded that most policy transplantations are only successful when these differences are taken into account (de Jong, 2009). This is especially true when the comparison is expected to result in a recommendation or policy change in a different country. Simply transplanting a successful policy of one country to another country does not guarantee success. Some studies have shown that forcing a policy may result in rejection by the society because it goes against cultural beliefs that exist (Dolowitz and Marsh, 2000).

In order to reach this objective, a framework is designed to compare the open data policies and derive lessons from the comparison. The lessons will then be formulated as recommendations to improve Indonesia's open data policy. Being in the early stages, Indonesia still has much room for improvement and development and will benefit from further research.

The outline of the thesis will be as follows:

- Chapter 1 – Introduction and research approach
    - ✓ This chapter will further detail the background of the problem that is researched. This includes the context of the problem, the societal and scientific relevance of the research that is conducted, a clear delineation of the problem that will be addressed and research questions that clearly stem from the problem identified. The methodologies that will be applied to answer the research questions will also be stated based on the appropriate approaches to reach the answer.
- Chapter 2 – Theoretical background
    - ✓ This section will give a more detailed explanation of open data and how open data has affected the way governmental agencies operate in the modern age. A definition of open data that the author has chosen for the research will be provided along with identified benefits and barriers of open data. Brief descriptions of each of the countries that are compared are also given from the theoretical standpoint. Following that a more detailed look specifically at open data policies and the existing comparison frameworks for open data policies and other relevant comparisons. This section will identify the initial elements to be included in the design of the framework.
- Chapter 3 – Analyzing the countries
    - ✓ This section will provide a more thorough overview of the examined countries based on the interviews that were conducted. Events that have influenced the open data process in each of the countries will also be identified. In addition this section will also address the cultural

aspects that influence open data policies that will also be compared. This section will then derive additional elements that are country relevant to be included in the design of the framework and also identify the forces and counter forces that strongly influence open data policies in each of the countries.

- Chapter 4 – Conducting the comparison of open data policies

  - ✓ Based on the findings from the previous sections, this section will list all the elements that will be included in the comparison framework and finalize the framework for the comparison. The next part of this section will apply the framework to the countries examined and create a comparison. Steps will then be applied to the results to derive lessons learned.

- Chapter 5 – Conclusions and recommendations

  - ✓ This section presents the main findings and conclusions to the research question on development of open data policies in the countries and also outlines the limitations of the research conducted and the designed framework. The conclusions and lesson drawn from previous sections will be applied to propose a set of recommendations to improve Indonesia's open data policies and also further research on the topic.

b. Practical and scientific relevance

The research is identified to be both relevant in practice and provide a contribution to science. First, as stated before, there are many identified benefits to open data that countries and governmental organizations can benefit from. These benefits will be further explained in the following chapter but overall improving a country's open data policy can help the country to reap these benefits. Opening data is not only for better transparency and accountability of the organizations that open their data but also for a better relationship between the citizens and the government. An example of the many identified benefits is the economic gain that is estimated in Europe as 30 billion euro per year (Lundqvist, 2012). Furthermore, public bodies hold a large number of data sets that may play a crucial role in innovation through the development of new applications, products and services (Janssen, 2011). This research will contribute to the effort of countries' to gain these benefits.

Second, this research will contribute to the current scientific studies on open data policies and comparison frameworks in the field of open data. Open data research still has a long way to go especially with the many barriers and challenges that have been identified during the process of implementation. At the moment there is still very little literature that discusses the issues of lesson learning using comparison frameworks in the field of open data. Most literature is focused on individual countries in their efforts to partake in the open data movement which also largely focus on lower levels of government. Research of open data policies on the national context is also limited.

Additionally, specific lesson drawing for open data policies in Indonesia is also a new field of study with many gaps to be filled. More studies are focused on the e-government and open government development with limited research into open data and related areas. These insights may also be relevant for other open data policies that are at the same level of maturity with Indonesia's policy.

c. Research question

From the problem description above it can be concluded that the problem addressed is a benchmarking study of open data policies. For the project itself the research will be focused on the following proposed research question and sub-questions:

Main question:

What can Indonesia learn from the open data policies of the United Kingdom, the United States of America, the Netherlands, and Kenya in developing its own open data policy?

Sub-questions:

1. Which elements of open data policies should be compared to develop recommendations according to available literature?
    a. Methodology: Literature on open data policies, general policies relevant to public sector information, comparison frameworks of open data policies, comparison frameworks of cross-country comparisons in policies or open government initiatives
    a. Expected results/outcomes: initial elements to be included in the comparison framework that is developed

2. What are the forces and counter-forces that drive and hinder the open data policies in the examined countries?
    a. Methodology: literature search on the drivers and barriers on open data policies from the political (from the political leaders and political class), cultural (from the citizen's acceptance and response), motivation and strategy (what the objectives to opening data are) point of view, contact with researchers of open data and policy making through the open data portals available from the countries as the initial point of contact, current news search on media and citizen responses to open data programs.
    b. Expected results/outcomes: additional elements that will be added to the framework as the drivers and barriers to open data in each of the examined countries.

3. What are the differences and similarities between the countries when compared using the comparison framework that is based on the elements identified for the examined countries?
    a. Methodology: using the elements identified in the previous sub-questions to design a comparison framework used to examine the differences and the similarities between the countries.
    b. Expected results/outcomes: a comparison framework of open data policies and relevant elements of open data policies of different countries.

d. Research methodology

In order to answer the proposed research questions, certain methodologies are applied to obtain and analyze the data. First of all, a literature research is conducted to find previous comparison frameworks on both open data policies and other policy making processes and research on open data policies. This point is more of an explorative research into the open data policy field to find which frameworks are suitable to adapt to compare different countries. Search terms include ‘open data policy’, ‘public sector information regulations’, ‘comparing policies’, ‘comparing open data policy’ to search databases such as Scopus, Science Direct, TU Delft repository, and Google Scholar. A previous framework designed by Zuiderwijk and Janssen ((to be published)) is used as the main framework to expand.

Second, a literature research is conducted to acquire knowledge about the various countries that are examined. The choice of countries stems from proof of an ongoing open data program, availability of initial information about the open data initiative and also possibility of contact with experts that have knowledge about open data in the countries. Further literature study is conducted for a more in depth look into the appropriateness of the countries for the comparison. The information gathered is based on policy documents and information that is stored on the various open data portals of the involved countries. Other criteria that are addressed are the cultural aspects of the countries. Hofstede's book on cultural dimensions (2010) is applied to analyze the similarities and differences of the countries that affect the comparison. Furthermore, literature that is more focused on dimensions developed by Minkov and Trompenaars in their respective publications are also studied. This desk research provides more elements that should be compared in the framework.

The previous desk research and interviews will provide knowledge of key elements that need to be included in the framework. In the design process of the framework, the approach of Bergmann (1957) is applied which includes observation, deduction, and induction. This approach is chosen because it provides logical steps to follow in the process of collecting and analyzing the data. Also it provides a certain level of flexibility which is needed when dealing with large amounts of data from different perspectives. The steps mentioned before are considered the observation step. Deduction comes from the interviews that validate the findings of the desk research and further refine the designed framework.

After applying the framework, inductive reasoning is used to derive recommendations. More specifically, from the results of the comparison, the next step to be conducted in the research is lesson drawing from the results of the comparison. For this purpose, a step-by-step process, that is outlined in a paper by Rose (2002), provides ten steps to draw lessons in the field of policy and policy transplantation that is followed. The lessons that are drawn from the comparison are intended to provide guidance for Indonesia's still developing open data policy hence the requirements of the lesson will include the applicability to Indonesia. According to Rose (2002), lesson drawing cannot be simply a description of what another country is doing but also the process (p. 5). The questions that should be answered in lesson drawing are under what circumstances and to what extent (p. 5). The following steps are developed by Rose (2002):

1. Diagnose the problem
2. Deciding where to look for a lesson
3. Investigating how a program works there
4. Abstracting a cause-and-effect model for export
5. Designing a lesson
6. Deciding whether to import
7. Dealing with resource requirements and constraints
8. Handling the problem of context
9. Bounding speculation through prospective evaluation
10. Using foreign countries as positive or negative symbols

Selected countries and interviews

To apply the framework that is designed, it was decided to analyze five countries that have open data programs in various stages and apply the framework to the existing situation. This can be considered as using a case study method as is described by Yin (2008). A case study method allows the investigators to retain the holistic and meaningful characteristics of real life events which include organizational and managerial processes and international relations (Yin, 2008). The choice of multiple case studies, by using

five different countries, is justified by the fact that Yin (2008) states that evidence from multiple cases is often considered more compelling and the overall study is therefore regarded as being more robust. Each selection of case or country in this research is done using the logic that the cases will (a) predict similar results (literal replication) or (b) predict contrasting results but for predictable reasons (theoretical replication). In this research it is done with a combination of the two predicted results, some cases are literal replications and a few other cases pursue different patterns of theoretical replications. The literal replications are between the countries that are in the same level of development in open data. The theoretical replications are between countries that are very much on different stages with open data but with reasons that can be foreseen. According to Yin (2008) an important step of these replications is that a rich theoretical framework is developed, which further justifies the design process of a framework to analyze the situation in the countries. This framework states the conditions under which the open data phenomenon starts.

The UK and the US are included in the comparison because of their influence on the open data movement globally. These countries can be seen as the literal replications that form the multiple-case study. Both countries are considered to have the most advanced national open data portals and have also been recognized as leading countries in the field of open government. Moreover the choice of the Netherlands to also be included in the comparison is because the Netherlands can be considered as a country that is in the middle of progress in the field of open data. It is not highly advanced as either the UK or US but has progressed immensely over the years. As for Kenya, the development of open data there is still relatively new. It was chosen because their current development with open data is still more advanced than the progress that has been made by Indonesia. Last of all, Indonesia is included in the comparison because of the early stage developments that Indonesia is going through. On top of that, because of personal interest in the development of Indonesia, this is the last country that is compared. As will be further deducted in the following research, more identifiers will establish the fitness of these countries to be compared.

Furthermore, targeted interviews about country specific questions are conducted for further details. Interviews are conducted with experts in the field of open government and/or open data in each of the compared countries. The experts vary from researchers that have conducted thorough case studies in the country, civil servants that are involved in opening data in the country, and people that work within organizations that are involved in the development of open government in the countries. Further details on the organizations or institutions that were approached for the interviews can be seen in Table 1. The choice of countries also depended on the willingness of the experts to participate in the research as interview sources. The full details and summaries of the 45 minute interview can be found in the appendix at the end of this report along with the list of questions that were asked. The first interview that was conducted was done in the early stages of the research and thus used a different approach compared to the rest. The purpose of this interview was explorative into the open data situation from the researcher's point of view.

Country	Field of work	Date of interview
United Kingdom	STFC Rutherford Appleton - Head e-science scientific applications	9 May 2013
United Kingdom	W3C on E-Government and Open Data Projects	7 June 2013
United States	National Science Foundation – Chief Technology Officer	7 May 2013
Netherlands	Delft University of Technology – PHD student	18 December 2012
Netherlands	Data.Overheid.NL – Project leader	17 May 2013
Kenya	Princeton University – Senior research specialist	22 June 2013
Kenya	Open Institute – Executive Director	4 July 2013
Indonesia	World Bank Jakarta – Public sector specialist	15 May 2013

Table 1 Interview sources

After examining the problem area of the research presented in this research, the following chapters will further elaborate the work that has been completed and also present the findings and analysis of the research. Using the research background, research question and presented methodology as a base of the research a thorough step-by-step plan to create a comparison framework is presented. Some drawbacks and constraints to these methods is the lack of access to people with sufficient knowledge to open data processes in different countries, very little previous research to draw from, lack of understanding of certain aspects that is researched by interview sources and also a lack of time to complete the extensive research. The steps that are taken to conduct the research are presented in Figure 1 as a flow chart to illustrate the process that is completed.

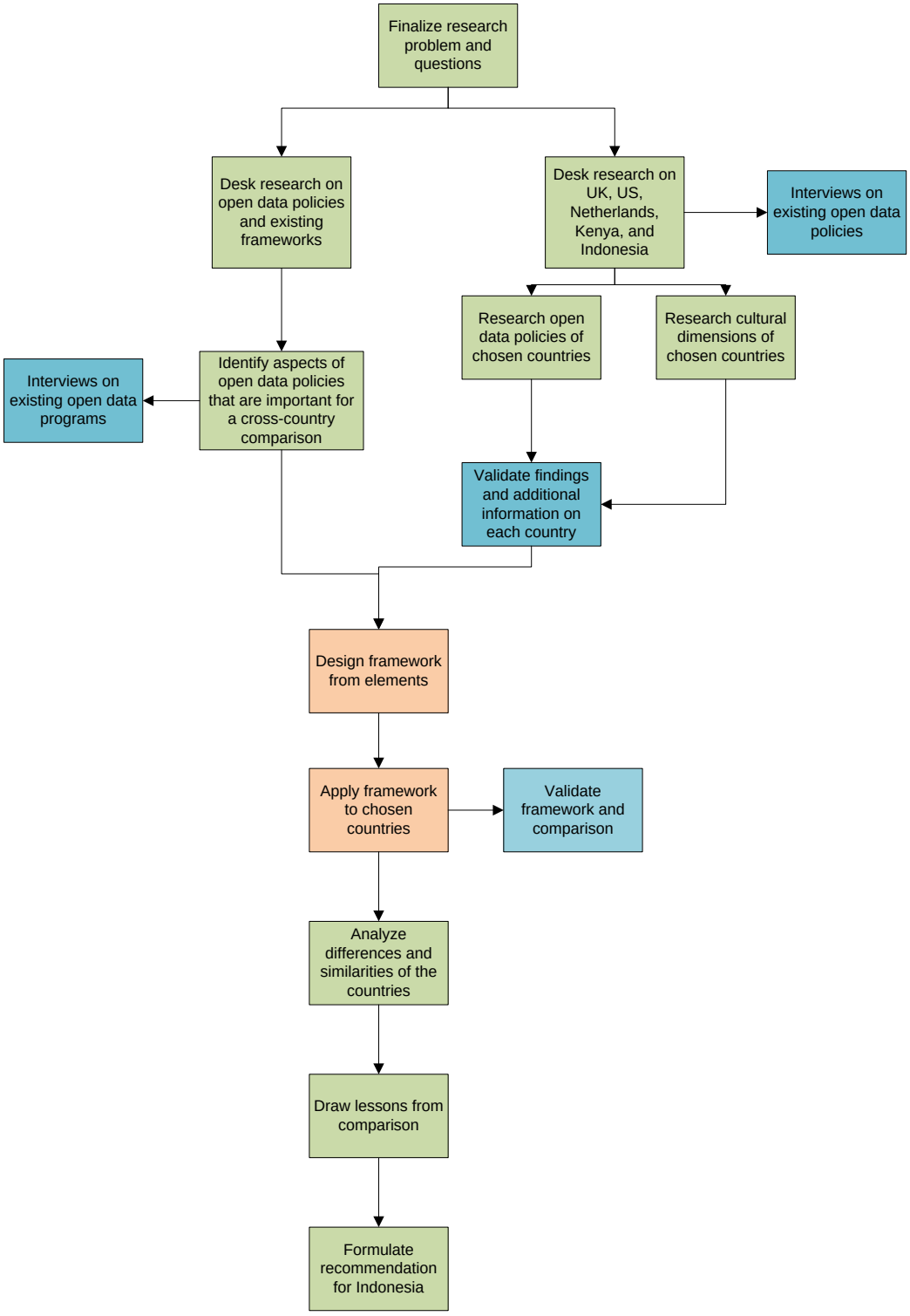

Figure 1 Research flow diagram

As can be seen, the diagram consists of different colors which represent different aspects of the research. Green indicates desk research or literature studies, blue activities are interviews that were conducted, and orange indicates the design activities. It should be mentioned here that 'validate' in the diagram means that the transcript and summary of the interviews were sent back to the sources of the interview to check and comment on. Additional aspects were mentioned and the findings were commented on before included.

Chapter 2
Theoretical background

This chapter presents the literature review about the various aspects that are analyzed in this research. It includes an overview of open data, the literature findings of the countries analyzed, and also a review of the existing comparison frameworks that are relevant to the field of open data and open data policy. This chapter also states the many assumptions that are made for the purpose of the research. The chapter first begins with an overview of open data which includes the chosen definition and the benefits and barriers to opening data. Next, is an overview of the facts about the observed countries which include demographics and also an overview of the current open data situation is provided. Following that is an overview of all relevant comparison frameworks to compare open data policies or other relevant cross country comparisons. Finally, this chapter presents the initial theoretical framework this is used in this research to compare the open data policies in the observed countries.

a. Open data

The following section provides a review of the basic knowledge in open data. As with any recent field of research there are still many contradictions between researchers on certain aspects of open data. This section clarifies the perspective taken on these contradictions which provides clear definitions of each element that is presented.

i. Defining open data

An example of this is the definition of open data. In previous research, there are many ways in which open data is defined. Before determining which definition best suits this research, an overview of the different definitions will be provided. The difference in definitions can be identified through the difference in data providers, type of data, users of data, format of data, and origin of data.

The Open Definition states open data as “a piece of data or content is open if anyone is free to use, reuse, and redistribute it – subject only, at most, to the requirement to attribute and/or share-alike” (opendefinition.org). It also further defines under what characteristics is data considered to be open through its accessibility, redistribution, reuse, absence of technological restriction, attribution, integrity, and non-discrimination. On the other hand according to data.gov, the official open data portal of the United States government, open data is considered as high value, machine readable datasets generated by the Federal government that is not private or restricted for national security reasons and is published to the public through the portal as the repository for all the information (data.gov). On the other hand, the Open Data White Paper published by the UK government defines open data as data that meets the criteria of being accessible at no more than the cost of reproduction, without limitations based on user identity or intent, in a digital, machine readable format for interoperation with other data and free of restriction on use or redistribution in its licensing conditions (HM_Government, 2012). A different definition can be found from the European Commission which defines open data as data that the public sector collects, produces, reproduces and disseminates (European_Commission, 2003). In Geiger and Lucke (2011) they present a very broad definition of open data as all stored data which could be made accessible in the public interest without any restrictions on usage and distribution. They go further in their article to define open government data which are all stored data of the public sector which could be made accessible by the

government in the public interest without any restrictions on usage and distribution. However it excludes the publication of all stored data of the public sector which must remain confidential.

Each of the above definitions has similarities and differences even though the main point is still evident. Both of the official open data documentation related to the US government and UK government respectively has similar definitions of open data which includes the variable of having the data machine readable and easily accessed. In contrast, the definition that is provided by the European Commission and Geiger and Lucke do not describe a lot about the specific requirements for the data.

However, because of the unique characteristic of the research that is presented in this report, a new definition that is based on the different definitions presented in the overview is used. The completeness of the definition provided by the UK government is the closest definition that is used as the reference point. The definition clearly states the accessibility requirements of the data, the format of the data as well as the licensing conditions that still bind it. To this point there are still limitations to what types of governmental data can be opened, a fact that is also clearly defined with this definition. However, it also provides the requirement to make the data linkable to other data thus ensuring the efficient re-use of the data that is opened. This part of the definition is adapted to fit the new definition needed for this research. Because of the different motivations in each country to open their data, creating linkable data is not included as a requirement for open data. Hence open data, for this research is defined as:

"Data that meets the criteria of being accessible at no more than the cost of reproduction, without limitations based on user identity or intent, in a digital format and free of restriction on use or redistribution in its licensing conditions".

ii. Benefits and barriers to opening data

Besides the many definitions of open data, there is also much research into the benefits and barriers that arise with the opening of governmental data. It can be said that the benefits and barriers to open data depend highly on the characteristics of the governmental agency that is opening their data. The research that is currently being conducted in this field hopes to identify more benefits of open data that can further stimulate the willingness of countries to join the open data and open government movement. However it must be noted that even though there are many benefits that have been identified, there is still lacking clear evidence that leads to proof of the impact (Huijboom and Broek, 2011, Schwegmann, 2012). This is because open data is still a relatively new movement which started around the year 2009. Most of the benefits are only realized in the long term which is yet to uncover itself at the moment. It is hoped that in the coming years it will become clearer. The following section will provide an overview of the relevant benefits and barriers for this thesis that is derived from the many existing ones. The importance of providing an overview of these benefits and barriers is because they lead to the development of relevant policies that highlight the benefits and help mitigate the effects of the barriers. The formulation of policies is highly dependent on the existing benefits and barriers. The explained benefits and barriers are examples of aspects that are taken into consideration when developing said policies.

First, the benefits of opening data will be explained. The main benefits include the improved accountability and transparency of the government (Zhang et al., 2005, Public_Accounts_Committee, 2012, Janssen, 2011) and also the increase of citizen participation (Schwegmann, 2012). Economic gains are also mentioned as a considerable benefit from opening data, it is said that the value of Public Sector Information in Europe is estimated at 30 billion euro per year (Lundqvist, 2012). On top of that *"public bodies hold a large number of data sets that may play a crucial role in innovation through the development of new applications, products*

and services" Janssen (2011, p. 446). Stakeholders that are involved in the open data process have also been discovered to have high hopes for the benefits of opening data even with knowledge about the barriers included (Zhang et al., 2005). In the report about Learnings from Kenya's Open Data Initiative (Kenei, 2012) some of the benefits that they have identified to motivate the opening of data in Kenya include transparency, accountability of government and donors to its citizens, empowerment of citizens, promoting inclusive development, and also economic benefits. In a nutshell, it covers all the benefits that were mentioned in previous research as well. Furthermore, open data can be offered to active citizens in order to facilitate innovation (Tauberer, 2009, Huijboom and Broek, 2011). An example is the utilization of open data to develop applications that are more widely used by the public such as traffic applications and updated weather forecasts.

Other benefits include the reduction of duplication of data collection and data handling, improvement of policy-making processes and sustainability of data. Reduction of duplication of data collection is best explained through an example of different governmental agencies needing the same data to perform public services which would be more efficient without having to collect the data multiple times but having it available for use by opening it. This would not only reduce the duplication of the data but also lower the costs that are involved. For the improvement of policy-making processes open data can aid policy makers by providing sufficient data that is needed so that they can better understand the problems that they are dealing with and thus take well-informed decisions (Arzberger et al., 2004). Open data is also a way to ensure that the data that is stored is maintained and protected from being lost thus ensuring the sustainability of the data.

Despite the many benefits to opening data, there still arise many impediments that need to be addressed. To create an effective use of open data there needs to be a balance between the benefits and barriers. Currently research is being carried out and policies are being developed to mitigate the barriers and promote more uses of open data. According to research conducted by Peled (2011), barriers such as the power play of the politicians and governmental agencies are factors that prevent the open data initiative from taking full potential. This is seen as a barrier because then the data that is published may be mistaken as data that has been previously doctored which eliminates the purpose of opening governmental data. Data manipulation of this type can also lead to false results if the data is further used in research by academics.

Upon examination of five countries worldwide, certain barriers for open data in each of the countries were derived and presented by Huijboom and Broek (2011). As each country has different motivations to open data, the barriers or impediments that are presented in their research differ accordingly and were placed in certain rankings based on how many countries agreed on a certain barrier. The top barrier was the closed government culture that currently exists making it difficult to change this frame of organization and have them be more open. In past years governmental practitioners have been more accustomed to working in secrecy and not in openness, for the data users this means that there will still be some data withheld by organizations (Australian_Government, 2012). The reason for this is there have previously been policies which actually ensure that governments keep the data private. Another mentioned barrier is the tension between the open data policy and the existing privacy legislations. Although open data is supposed to be as transparent as possible there are certain conflicting interests that arise from how open it can be without publishing information that may lead to the identification of the persons involved.

More barriers mentioned by Huijboom and Broek (2011) are about the quality of the data itself and how useable this data is. Because of the previously mentioned secrecy in the operations of governmental agencies in previous years, the data that is stored or kept is not always of highest quality making it difficult

to publish it for useable purposes. In relation to the quality, the data that is published has also been discovered to be in less user-friendly formats which are caused by the lack of standardization of the data that can be published. Commonly used formats for this type of data include excel files or .CSV files which should be standardized in the open data guidelines. Another impediment is that certain datasets require a certain payable fee before being able to gain access. This reduces the purpose of having the data available for public if fees are instilled it limits access to certain parties that may not have the ability to purchase such data. But looking at it from the data provider's perspective, most organizations gain income from selling data which means by providing it on open data platforms for free they are letting go of a source of income. One of the last barriers mentioned in this literature is the unequal access to the open data and network overload by presenting large amounts of data on the existing infrastructure. Unequal access to data is mainly discovered to be true in Spain and the US where there is a digital divide which causes this inequality of access. Unequal access in this case means that even though the data is published for everyone, the infrastructure does not allow some people to have easy access to the data. As for the network overload, this is mainly visible in the US where there are not enough available networks that have the capacity to hold this data because the US is publishing such large amounts of data.

A different perspective of analyzing the impediments or barriers to open data is to view it from the data user perspective. There are ten categories of impediments that are identified by Zuiderwijk et al. (2013) that is viewed as socio technical impediments of open data from the data user perspectives. The ten categories include 1) the availability and access, 2) find ability, 3) usability, 4) understand ability, 5) quality, 6) linking and combining data, 7) comparability and compatibility, 8) metadata, 9) interaction with data provider, and 10) opening and uploading. Most of the barriers identified relate to the ability of users to re-use the data and create value from the data that is published. Based on the research that was conducted, from literature, workshops and interviews, it is clear that there is a need for open data policies that not only puts pressure for the organizations to publish but to also publish data that is usable for the users (Zuiderwijk et al., 2013). At the moment there is still a certain reluctance from data providers to open their data because of the unknown impact (Zuiderwijk et al., 2013). One important barrier is the inability of users to handle the complexity of the data presented and actually use the data presented on the open data platforms. The more data is available it becomes more difficult to analyze and draw conclusions from it (Zurada and Karwowski, 2011). This barrier is also mentioned by Janssen et al. (2012) which creates the need for good structure and support in handling and using the available data. Making the data open and available to public may not be sufficient anymore. Additional support is needed for users to actually use this data in a meaningful way. Hence, there is a need for a uniform policy that ensures that the data published is not just published but also accessible for data users to understand and re-use (Zuiderwijk et al., 2013).

b. Examining the countries

As explained in the methodology, the countries that are examined for the purpose of the research are the United States, United Kingdom, Netherlands, Kenya and Indonesia. The reasons for choosing these countries are also explained in that section. Before explaining the current open data situation in each of the countries, an overview of how the countries differ from each other in terms of more general characteristics are provided in the following Table 2. The statistics provided in the table is taken from latest information that is provided in the 'The World Factbook' by the Central_Intelligence_Agency (2013) as the most up-to-date facts on each of the characteristics.

Country facts	Country				
	United States	United Kingdom	Netherlands	Kenya	Indonesia
Total area	9,826,675 sq km	243,610 sq km	41,543 sq km	580,367 sq km	1,904,569 sq km
Population	316,668,567	63,395,574	16,805,037	44,037,656	251,160,124
Age structure (25 – 54 years)	40.2%	41.1%	40.8%	32.4%	42.2%
Government type	Constitutional based Federal Republic	Constitutional monarchy	Constitutional monarchy	Republic	Republic
Legal system	Common law	Common law	Civil law system	Mixed legal system of English common law, Islamic law, and customary law	Civil law system based on Roman-Dutch model and influenced by customary law
GDP per capita	$ 50,700	$ 37,500	$42,900	$1,800	$5,100
Taxes	15.7% of GDP	40.8% of GDP	44.9% of GDP	17.9% of GDP	-1.9% of GDP

Table 2 Country facts and demographics (Central_Intelligence_Agency, 2013)

From Table 2, it can be observed that in terms of wealth, Kenya and Indonesia are quite different from the US, UK and Netherlands. As is previously mentioned in La Porte et al. (2002), characteristics of a country that have an influence on the openness of the government is dependent on national wealth and the size of the country. This is why these characteristics are to be included in the comparison that is conducted. The next table that is included contains indexes that are a result of the United Nations E-Government Survey in 2012 (United_Nations, 2012). This report also states the relation between highly developed e-government with high economy countries. For the survey, the included in the parameters are e-government development index, online services index, telecommunication infrastructure index, and human capital index. The e-government development index is a benchmark for numerical ranking of e-government development in countries that are in the United Nations. It is "*a weighted average of three normalized scores on the most important dimensions of e-government of e-government*" (United_Nations, 2012, p. 136). These dimensions are the other indexes that are included in the survey as well. The indexes can be seen in Table 3.

Parameters	Country				
	United States	United Kingdom	Netherlands	Kenya	Indonesia
E-Government Development Index	0.8687	0.8960	0.9125	0.4212	0.4949
Online Services Index	1.000	0.9739	0.9608	0.4314	0.4967
Telecommunication Infrastructure Index	0.6860	0.8135	0.8342	0.1212	0.1897
Internet users (per 100 inhabitants)	79	85	90.72	20.98	9.10
Main fixed telephone lines (per 100 inhabitants)	48.70	53.71	43.15	1.14	15.83
Mobile subscribers (per 100 inhabitants)	89.86	130.25	116.23	61.63	91.72
Fixed internet subscriptions (per 100 inhabitants)	26.63	31.14	37.02	0.08	0.73
Fixed broadband facilities (per 100 inhabitants)	26.43	31.38	37.97	0.01	0.79
Human capital index	0.9202	0.9007	0.9425	0.7109	0.7982

Table 3 E-government development indexes (based on UN E-Government Survey 2012)

The online services index is a measure of how much the country's governments put online. To reach this value each country's national website is assessed, including portals, e-services, e-participation and ministerial websites. The websites are assessed for content, features, and level of web content accessibility. The telecommunication infrastructure index contains the average of elements that are also included in Table 3 (internet users, main fixed telephone lines, mobile subscribers, fixed internet subscriptions, and fixed broadband facilities). Finally the human capital index is an average of adult literacy and gross enrollment ration which both show the education level of that country. Both Table 2 and Table 3 present more elements that are important to compare between countries when it comes to e-government programs such as open data. These elements are considered in the analysis of the cross country comparison that is in this report. These elements are identified in the two sources as elements that influence open data in the countries that are further observed in this report.

To further examine the countries under observation, the following sections include an overview of each country, in terms of open data, individually. Included in the overview are the actions that have been taken for open data, the current open data policy situation, and the goal and objectives for open data in each country.

i. United States of America

The government of the United States (US) is one of the first countries to opt for open data in the world. The US has been a long believer that citizens have the right to know about information the government possesses since the enactment of the Freedom of Information Act in 1966. The importance of the open government movement in the US is clearly seen by the fact that President Obama has made it a high priority since the first day of his administration (Open_Government_Partnership, 2011c). The Open Government memorandum was the first action that Obama signed during his administration on 21 January 2009. Shortly after was the launch of data.gov as the national open data portal in May 2009, in response to the Open Government Directive. The portal intends to increase public access to high value, machine readable datasets generated by the Executive Branch of the Federal Government. In a recent Executive Order in May 2013, President Obama signed the order to the opening of machine readable data as the new default for government information and published a new open data policy. Obama hopes that by doing so that it will help launch more start-ups, businesses, promote innovation and ingenuity that will transform the way many things are accomplished (Sinai and Dyck, 2013).

The Office of Management and Budget (OMB) was instructed by Obama to issue an Open Government Directive to implement the principles of transparency, participation, and collaboration (Obama, 2009). Transparency promotes accountability which can in turn improve the overall performance of the government and also to encourage participation of the public through information that is readily available through new technologies. In response to this memorandum, the director of the OMB issued the Open Government Directive to direct executive departments and agencies to take actions and implement the three principles (The_White_House, 2009). In the following months, the Freedom of Information Act (FoIA) was adapted to reflect the new commitment towards an open government in March 2009. The Open Government Directive states that executive departments and agencies are to take certain steps to create a more open government which includes publication of government information online, improvement of the quality of government information, creation and institutionalize a culture of open government, and creation of an enabling policy framework for open government.

In relation to publishing government information online, the memorandum states that it shall always be in favor to openness as is also specified in the newly updated FoIA. However, it also needs to be under the

extent that is permitted by law and subject to valid privacy, confidentiality, security and other restrictions. Data quality is also a step that is mentioned in the directive. This is conducted by ensuring that there is a process that is followed according to the Information Quality Act and is regularly maintained.

To create and institutionalize a culture of open government, openness should be underlying in any and all government work as is incorporated by senior leaders. This is encouraged by working together as a whole government and together achieving the end goal of an open government. In the memorandum itself, each agency is asked to develop and publish an Open Government Plan that details that agency's efforts in improving transparency. Another course of action is creating a working group that meets to discuss and help each other to create a more integrated open government effort that goes across all the agencies.

A policy framework should be created that takes into consideration the usage of emerging technologies to publish the data that is to be opened. Existing policies regarding the release of data that has already been regulated through the OMB will be updated to provide guidelines that will assist the opening of data through various available technologies. These are the current aspects that are mentioned in the policy document about the open data situation in order to reach the goal of transparency and accountability that is the main goal of open data in the US. From the initial findings about open data in the US, it is clear that it can be considered to be a strong leader in the field. The development of the policies that strongly encourage the opening of data is seen to elevate the level of openness within the government and other organizations. For the research that is presented, it was easy to access all the needed information about the open data process that is followed by Federal Agencies. This is also another reason why the US can be viewed as a leader because anyone can access information about how the data is managed which benefits both the publishers and the users. From this literature review of open data in the US, it is decided to include the US in the comparison that is conducted.

ii. United Kingdom

Following the US, the UK was next to join the open data movement. However, the beginning of open data in the United Kingdom (UK) began with the Freedom of Information Act in 2000 that stated that the public has the right to access public sector information (PSI). In this case, PSI is what we have defined so far as open data. Another campaign that ignited the awareness of open data in the UK was the "Free Our Data" campaign by a British daily national newspaper The Guardian in 2006. Even though it only reached a small target group of readers, it is still noticeable as one of the first movements towards open data in the UK. A few years later the open data portal of the UK, Data.Gov.UK, was made available since 30 September 2009 although it was officially launched in January 2010.

One of the recent changes that have been published in regard with the existing FoIA is that several sections have been adapted to incorporate the free availability of data in machine readable format. This amendment to the FoIA also incorporates the new Code of Practice. These two policies further encourage the re-use of data by ensuring that the data published is in raw format. This allows a broader re-use of the data by users and also promotes the linking of data. Besides providing ways for data users to contact the data publishers in regard to the datasets that are made available or hoped to be made available, it is possible for citizens to participate in the process of designing these policies and guidelines to open the data. For instance the Code of Practice was open for consultation for twelve months before the draft was published. Consultation in this case means that data publishers and users that are active within the Data.Gov.UK communities can contribute to the decision making process by recommending certain courses of action that will ease the open data use and re-use from their perspective

The focus for open government in the UK has been on increasing public sector accountability, improving public services and more effectively managing public resources which is stated in two open letters from the Prime Minister to his cabinet. The letters established commitments towards opening data that is held within the governmental bodies such as spending data for the central and local governments, crime data, and data regarding the civil servants. Further to this, the second letter was described as one of the most ambitious open data commitments in the world because of the extensive list of data that was supposed to be opened which included health, education, criminal justice, transport and government financial information (Cameron, 2011). Additionally actions that were taken include the establishment of the Public Sector Transparency Board and the creation of the Open Government License. In terms of policies that regulate the opening of data there are several policy documents that exist. Some of the guidelines used to open data include the Public Data Principles which provides fourteen compulsory principles that were published by the Public Sector Transparency Board, Sir Tim Berners-Lee Five Star ranking system for the re-use process of the data, and the Government Principles for Open Standards especially for software interoperability, data and document formats.

In the beginning, the focus of the UK to open their data was also to increase transparency of the government but this has evolved as the open data initiative continues to grow. In the Open Data White Paper, the steps taken to unleash the full potential of open data in the UK are described and commitments of each participating department are provided as well (HM_Government, 2012). Some of the points in the White Paper that are worth mentioning include the commitment to use the Five Star Scheme by Tim Berners-Lee to measure the usability of open data. This will ensure the highest level of usability of the published datasets and in turn will promote the users to re-use the data in innovative ways. It also goes further to commit to include the participation of the data user communities in developing the databases further. This opens the traditionally closed government to have more interaction with the data users and gain feedback on how to improve the data that is disclosed. This also serves as a way to maintain the quality of data that is being published. As mentioned in the previous chapter, one of the concerns that arise from the release of data is the fact that personal information might be uncovered with the linking of several anonymous datasets. This challenge is also addressed in the White Paper by committing to the hiring of a privacy expert that will conduct Privacy Impact Assessments that will mitigate threats that could exist from releasing a dataset.

The UK is also considered a global leader in the field of open government and open data. After the initial research that was conducted this opinion is also shared for the purpose of this research. The findings through policy documents and website searches proved that the UK has indeed a strong foundation for their open data and it is also promoted strongly by the government, which is similar to the situation in the US. The published documents on open data also provide clear guidelines that are beneficial to both the publisher and the data user in order to manage the data that is opened. Even though, through the initial research that was conducted, no mention of having to open the data is found, it is evident that the UK have been actively opening data through the open data portal. The UK is also an important country to be included in the comparison based on the initial findings.

iii. The Netherlands

The Netherlands has been ranked second for ‘open government’ in the World Justice Project’s Rule of Law Index for 2010 (Open_Government_Partnership, 2011a) which shows its strong grasp on the principles of open government. This is also reflected on the long presence of the Government Information (Public Access) Act (Wet Openbaarheid van Bestuur; WOB) since 1980 which has been amended to fulfill the more specific requirements of open data. However, the open data movement started when the Obama administration

announced their open government directive and this created the initiative to follow in the Netherlands as well. Although the Netherlands is not considered one of the first countries to come on board with the open data movement, there has certainly been increasing progress in the field. In September 2011 the Minister of Interior and Kingdom Relations launched the national Open Data portal (data.overheid.nl) which is motivated by the need to get commercial value and contribute to economic growth after data enrichment that is done by the citizens. It is also said to lead to a more transparent government, stimulate economic activity, build an efficient government, improve the public services, and drive innovation. Under the Action Plan for the Open Government Partnership, there are certain areas which the Netherlands has committed to improve in order to move towards a more open government. These areas include amendments to the WOB, active publication policy, open data programs from each of the ministries, efforts in the area of integrity, increasing the number of civil society initiatives, and improvement of various public services. Each of these areas shows the structured plan of the Netherlands to prove their country to be as open as is ranked.

Through the WOB, citizens had the right to request for information about an administrative matter to an administrative authority. Most of the information that is produced by the government is made public under the WOB. Now, with the launch of Overheid.nl, this data is available for easier access which also aligns with the government's latest actions to promote the reuse of government information. The portal consists of no actual data but a reference index that provides access to sources of government information, an overview of the possibilities that data reuse can provide, news and background information about open data in the Netherlands, and also links to forums for citizens to actively participate in discussion about open data. As of now, the open data portal consists of over 5000 dataset links, the complete Dutch legislation and other deep links to publicly available government information which are all available for free reuse.

On the portal of Overheid.nl, there is much information about suggested guidelines to open data. However, these guidelines are not yet compiled in an official policy document. There is also not a law that forces any governmental bodies to publish their data. When observing the existing guidelines that are available on the portal, it can be seen that these guidelines are very detailed and provide usable guidance when a data owner wishes to open the data. The guidelines consist of five steps that are involved with the process of opening data. The steps start with deciding which data is suitable to publish. This step-by-step process is shown in a flow chart diagram that takes the publishers through the process. Another guideline that is included is about deciding which license is needed for the data or the legal check. The next step involves explaining how best to organize the data that is published and the tasks that are involved. The last step is to include identifiers to the data to make the data discoverable and accessible. This step-by-step process is only given as encouragement and best practice and is not obligated to be followed. To create these guidelines into formal policies it has proved to be a long process which could still take years to complete. Continuing efforts to increase the availability and provision of data, there are three areas that have been identified to stimulate the reuse of open government data which includes the usage to address specific societal issues, usage by the public sector itself and stimulating commercial reuse.

Because of the lack of policies, the overview of the Netherlands is unable to be as detailed as that under the US or the UK. It is interesting to note that even without policies of such, the guidelines that are made available are detailed and if followed can make opening data considerably easy. This is one of the reasons the Netherlands is included in the comparison because these guidelines are available even if they are not considered policies. The Netherlands is further considered as a country to be included in the comparison because in rankings and surveys (United_Nations, 2012, Open_Government_Partnership, 2011a) it remains one of the highest ranked countries for openness or e-government.

iv. Kenya

The start of open data in Kenya can be seen to have a longer story compared to the countries that are stated above. Different from the US, UK, or the Netherlands, Kenya had a sense of urgency to apply open government to help the development of the country as a whole. Under former President Daniel arap Moi, between years 1978-2002, the government restricted the free flow of information and hindered other forms of media networks. This was held under the Official Secrets Act that had existed from the colonial era. Under the influence of such an Act, it is understandable that the government culture was built to restrict access to information from the public and to closely guard all categories of information. It could be said that there was an pressing need for transparency and openness in order to rid the country of corruption. In 2005, Bitange Ndemo became the permanent secretary of Kenya's Ministry of Information and Communications and brought changes within the government that would enable a more open government. On 8 July 2011, President Kibaki officially launched opendata.go.ke with 200 datasets that were categorized into education, energy, health, population, poverty, and water and sanitation. The launching of the website also launched the Kenya Open Data Initiative which was an important step for the country.

When the portal launched, Ndemo wanted to avoid confrontation with public officials on publishing non-public data by launching the portal with data that was already categorized for the public but not yet published. However, no policies were enacted to ensure the government opened their non-published data. In 2010 there was an addition to the constitution which called for the government to "publish and publicize important information affecting the nation". This constitution amendment was the anchor to the open data efforts instead of waiting for the Freedom of Information law. Another document that contains statements about open data is the Vision 2030 Plan, which is a long term development blue print launched in 2008. This Plan provides the means towards a more open government through ICT infrastructural developments. However, because of the lack of Freedom of Information law there is no legal background or formal policies that enforce the government to open the data. To the extent of the document search that was conducted for this research, there were no mentions of formal policies that regulate the opening of data. There is however, mentions of requirements on the opendata.go.ke portal about the data that is to be published. Similar to the situation in the Netherlands but significantly less detailed. Issues such as licensing, privacy, metadata, and formats of data are mentioned very briefly and merely suggested as best practices. It is mentioned on the portal that there are mechanisms for users to present recommendations about the data. This is meant to provide input for the continuous evolvement of the portal.

On the global scale, Kenya has also made a statement to the world by joining the Open Government Partnership. Through the OGP, Kenya has formulated an action plan that is committed towards addressing certain areas of open government. The areas include the improving of public services, increasing public integrity, and more effectively managing public resources. The action plan also states many on-going initiatives that are targeted at solving these areas along with the target dates of accomplishing the plan (Open_Government_Partnership, 2012). The Freedom of Information law is currently still tied up in parliament and has yet to be announced as a new law. This law may be one of the turning points needed for the Kenya Open Data Initiative to be completely in action.

Kenya is said to be the first developing country to launch an open data portal (opendata.go.ke), which is one of the main reasons why it is included in this comparison. From a different perspective, the other countries in this research are all considered to be developed countries. Also, because the intention of the research is to create a recommendation for Indonesia, Kenya is a good example of how open data can be beneficial to a

country's reformation. To bring diversity to the comparison, Kenya is a country that has launched its open data initiative but is still considered to be in the very early stages.

v. Indonesia

In Indonesia the need for open government came in line with the 1998 riots and the fall of former President Soeharto. It became clear that the government needed to make drastic changes towards a more transparent government because of the heavily rooted corruption that had permeated into many areas of the government. This brought around changes in the way the public responded to governmental actions and also led to the start of the democratization process. In 2004, the devolution of power came to Indonesia through decentralization of the government. Since then, many efforts have been implemented to empower the citizens to be more engaged in public affairs. Openness in government activities is hoped to trigger the bureaucracy reform between sectors (Sekretariat_OGI, 2012b). This will lead to transparency and can improve un-effective and inefficient processes and procedures within the government. This will also clarify systems and procedures that are needed to provide high level of service to the public. These factors are considered to be the end goal or objective of open data in Indonesia at the moment.

The Open Government Indonesia (OGI) movement started in September 2011. In the effort to enforce the open government movement in Indonesia, a specific unit was appointed under the President called the UKP4. This unit is tasked with all the open government duties, also those related to the global Open Government Partnership of which Indonesia is co-chair in 2013. In 2008 the Freedom of Information law was decreed which is used as one of the legislative backgrounds for opening data in Indonesia. However there are not specific open data policies that are targeted at opening data on an online portal. The FoI law is already considered a big step towards a more transparent government. It gives citizens' rights to information related to public policy making, encourages active participation of the citizens, and improves the managing of public services. Because Indonesia is only at the very early stages of becoming an open government, many policies are still needed. Currently there only exists the 2008 FoI law in the field of open data which simply encourages the publishing of public information and does not force it. A specific law that regulates and guides the actual publishing is still lacking. The information that can be included in this section about Indonesian open data is very limited because of the lack of information that is available about the topic in general. Most of the information that is on the OGI website relates to the long term plans for openness that Indonesia wishes to accomplish. It is not focused on open data or the policies that have been enacted so far.

From the initial desk research that was conducted on Indonesia open data, it can be seen that open data is still in the very beginning stages of development in Indonesia. At the moment there is more focus on gathering awareness on the matter of open data and open government rather than formulating policies on the opening of data. From the desk search, there was not found to be any regulations that mention the need for specific data types, formats, licensing, or any other policy related aspects as were mentioned about the previous countries. However, the designation of a specific taskforce to ensure the development of open government is a sign of Indonesia's seriousness in joining the global movement. Also by participating in the Open Government Partnership, Indonesia has made an international statement of its commitment to the movement. This is why this research hopes to assist Indonesia in providing lessons that are needed to support enforcement of opening data nationally.

c. Current research on comparing policies

The Merriam-Webster dictionary defines policy as a high-level overall plan embracing the general goals and acceptable procedures especially of a governmental body. Another definition of policies is a purposive

course of action followed by an actor or set of actors in dealing with a problem or matter of concern (Anderson, 2010). When applied to the field of open data it can be said that open data policies provide guidelines that can help stimulate the continuous opening and reuse of data through a well thought plan. For the purpose of this research open data policies are seen as guidelines that regulate the continuous opening of data to the public that is not by request. Also the focus of this research is on national open data policies that are valid for the opening of data in the country and also specifically for publishing on the national open data portal. Some of these guidelines may not be formally known as national open data policies because they are not strongly reinforced but they are considered as the open data policies that are analyzed in this research.

As has been defined in the previous section in this chapter, open data is predicted to be highly important and valuable. This causes the appearance of some boundaries that need to be regulated through policies. These open data policies can be seen as a way in which a country can guarantee that their government will continue to open data and remain transparent and accountable for all their actions. Besides from ensuring the process of opening data, open data policies aim to achieve a certain impact on the society as does any policy.

Several sources have stated that a comparison of open data policies and implementation of the open data programs can assist the better formulation of the policies and overcome some of the current obstacles (Zuiderwijk and Janssen, 2012a, Zuiderwijk and Janssen, (to be published), Huijboom and Broek, 2011). In this section a brief explanation of the current frameworks that have been developed for this purpose will be given. Before explaining the existing frameworks, a look into the domains that need to be managed for efficient data access which is essential to open data will be explained. These domains are the building blocks on which the open data policies are developed then are translated into the elements that need to be analyzed when comparing different policies. The domains that need to be assessed for data access management include technological (T), institutional and managerial (IM), financial and budgetary (FB), legal and policy (LP), and cultural and behavioral (CB) considerations (Arzberger et al., 2004) which is seen in Figure 2. Arzberger et al. (2004, p.144) argues that these domains are in itself "*a framework for locating and analyzing where improvements to data access and sharing can be made"*

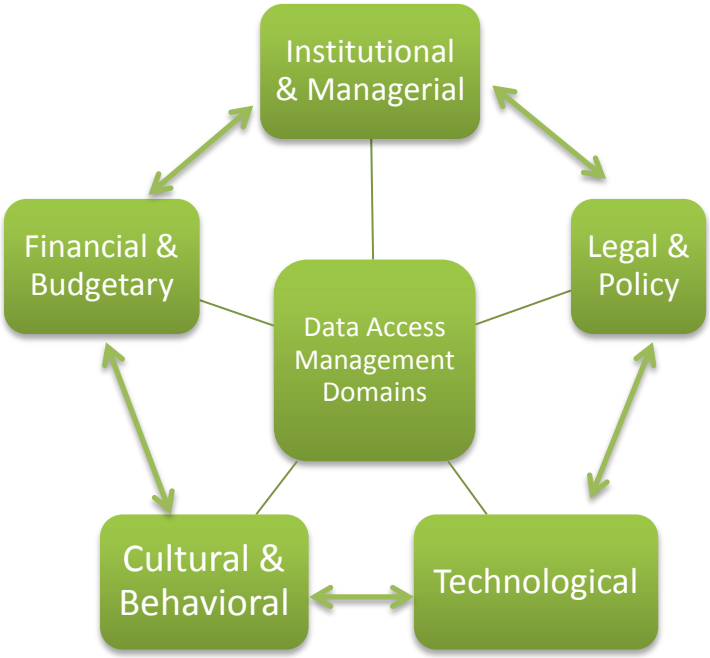

Figure 2 Data access management domains (Arzberger et al., 2004)

Moreover attempts have also been made to develop guidelines for the opening of government data (Zuiderwijk et al., 2012). Issues that need to be taken into account when opening data have been identified that are also part of formulating open data policies. According to research conducted by Zuiderwijk et al. (2012) there are general issues and dataset specific issues that need to be taken into account when opening the data. General issues include the confidentiality, deletion policy, embargo placement, organizational changes, ownership of data, privacy sensitivity, lack of metadata, use and reuse of data, policy sensitivity and unlawfulness. Dataset specific issues include the completeness and exhaustiveness, representation, validity, reliability, clearness, provision of additional reports, and overall data quality. These issues combined with the data access domains formulate the basis of the comparison frameworks that exist. In a further extension of this framework, Zuiderwijk and Janssen ((to be published)) separate their framework, based mainly on existing literature, according to environment and context elements as the input, policy content, policy evaluation using performance indicators, and policy evaluation by realizing public values. The domains that categorize the elements in the comparison framework are based on the policy making cycles of Stewart Jr et al. (2007) which include agenda setting, policy formulation, policy implementation, policy evaluation, and policy change or termination. The framework that was developed is used for comparing open data policies at different government levels within the Netherlands. The reason for this is that a closer look into the policies that exist at the lower levels of the government will affect the national policy too. The framework can be seen in Figure 3 below.

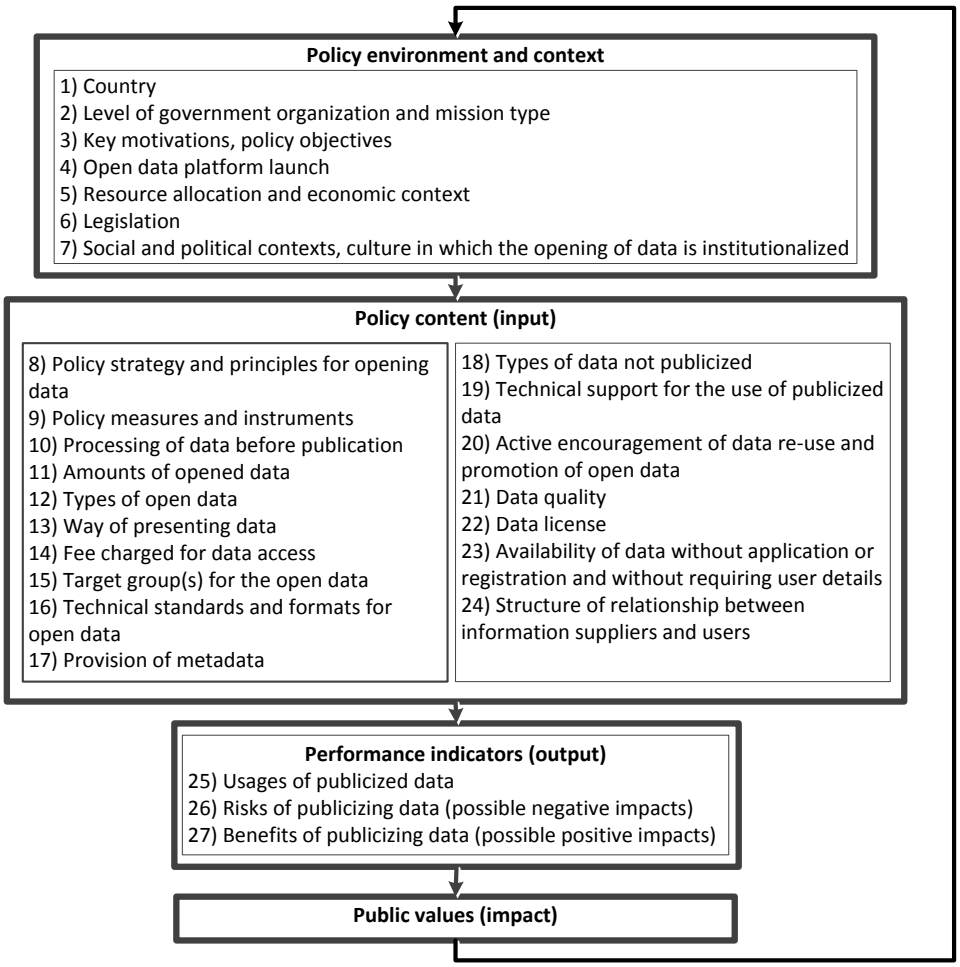

Figure 3 Comparison framework from Zuiderwijk and Janssen (to be published)

The framework in Figure 3 shows a detailed look at the various aspects that are considered when comparing open data policies. A limitation about this research is that the public value or impact of open data as part of the framework is near impossible to measure when comparing national open data policies of different countries. Also because this framework was designed to analyze open data policies at lower levels of the government specifically in the Netherlands, it is possible to gather more detailed information about the more technical aspects of open data. However because of its high level of detail into the elements compared in open data policies, this framework is used as the main reference.

In a different study by Gibbs et al. (2003), a framework for comparing the diffusion of e-commerce in different countries is presented. The use of this framework is to identify the additional elements that need to be inspected when performing a cross country comparison as opposed to a lower level comparison as conducted in the abovementioned frameworks. According to Gibbs et al. (2003), factors of global environment and national environment need to be taken into consideration when developing a comparison of this scale. The national environment factors that are mentioned include demographic factors, economic and financial resources, information infrastructure, organizational environment, public preferences, and national policies on legislation and the promotion of e-government. On the other hand, the mentioned global environment is not relevant for the study of open data or open data policies and is not taken into consideration for this research. From the study of these factors, many lessons can be drawn in conclusion about the adoption of e-commerce and e-government to some extent. Because this framework was

developed for a more global view on how to compare different countries not all of it is applicable to this thesis. Although a global view needs to be analyzed when conducting a cross-country comparison, for the case of open data policies a more detailed overview of the technical aspects that are involved with open data should also be considered. However, this framework provides the elements from the national environment that need to be compared which is included in the framework presented here. A closer look at the conceptual framework by Gibbs et al. (2003) can be seen in Figure 4.

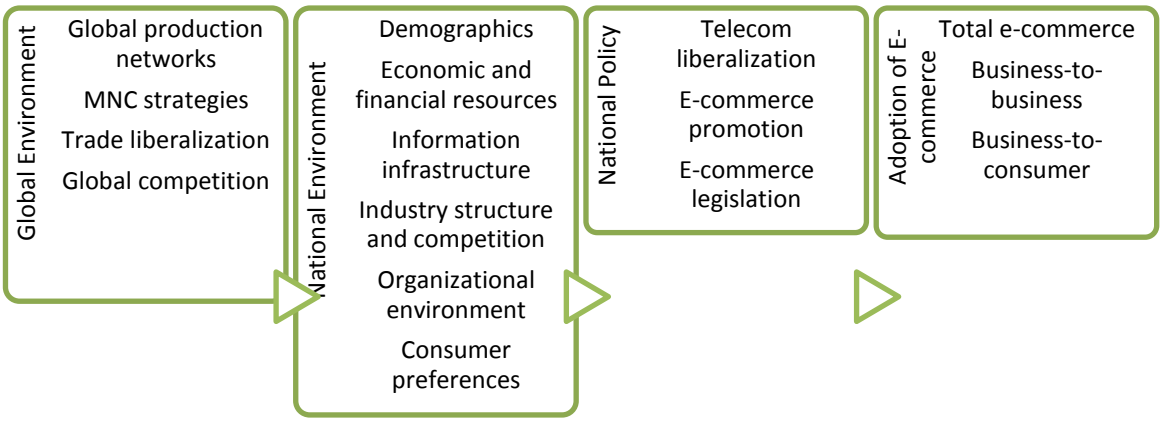

Figure 4 Conceptual framework by Gibbs, et. al (2003)

Another source that compares open data programs and the implementation of open data policies across different countries is provided by Huijboom and Broek (2011). The countries that are evaluated in this literature are the five countries that are considered to have an established open data initiative (Australia, Denmark, Spain, UK, US). This comparison is mainly focused on the implementation of open data policies, instruments used to implement their open data strategy, and how they are motivated and hindered in each country. However it can also be seen through this comparison that there are similar factors that are evaluated compared to the previous framework by Gibbs (2003) such as the economic instruments that the country owns to execute the policy which is considered the national environment. The factors compared include the education, voluntary approaches, economic instruments, and legislation and control. The framework that was developed can be seen in Table 4. A limitation from this framework and also to the previous framework by Gibbs is that they do not take into consideration that cultural aspects or even specific forces or counter forces might be the underlying cause as to how open data is progressing in those countries at a different pace. This is how the research presented here provides additional value.

	Broad category			
	Education and training	***Voluntary approaches***	***Economic instruments***	***Legislation and control***
Instruments	Knowledge exchange platforms	Overall strategies and programs	Competitions, app contests, and camps	Freedom of information act
	Guidelines	General recommendations	Financing of open data portals	Technical standards
	Conferences, sessions, workshops	Public voluntary schemes		Monitoring

Table 4 Framework by Huijboom and Broek (2011)

d. Proposed theoretical framework

From this overview of the current research that has been conducted on the various comparisons of policies related to open data and also from the overview of the countries, an initial framework can be developed for the purpose of the benchmark in this thesis. The initial elements that are included in the framework are based on the available literature in the field and are formulated to answer the first sub-question of the research questions. It should be noted that the comparison that is conducted in this thesis is an extension of the research conducted by Zuiderwijk and Janssen ((to be published)). The difference is that this research compares national open data policies of different countries instead of different government levels within one country. The basic overview of the framework is highly similar to the framework by Zuiderwijk and Janssen ((to be published)). The framework is separated into the following categories:

1. Policy context

This category evaluates the external factors that exist in each country that affect the formulation of national policies. The elements included here form the context in which the policy is formulated. It includes the aspects that form the circumstances under which the policy is enacted. This includes the objectives, allocation of resources, existing legislations, social and political contexts, and the forces and counter forces that affect the opening of data. These elements were deducted based on the existing framework by Zuiderwijk and Janssen (to be published) and also by the overview of the countries and other frameworks that was conducted in the beginning of this chapter. Based on literature provided by Gibbs et al. (2003) and the domains of Arzberger et al. (2004) the policy context also compares the country demographics, economic resources, information (IT) infrastructure and cultural influences. The complete elements and description of how they are identified are provided in the following Table 5:

Element		Literature	*Description*
Level of government organization		Huijboom and Broek (2011), Gibbs et al. (2003)	*What level is the policy observed?*
Key motivations, policy objectives		Huijboom and Broek (2011)	*What are the motivations to opening data?*
Open data platform launch		Huijboom and Broek (2011)	*When was the national open data platform launched for public use?*
Resource allocation and economic context:	Technology penetration	Huijboom and Broek (2011), Gibbs et al. (2003), Eskelinen et al. (2008)	*According to statistics, how has technology permeated in the society?*
	ICT Infrastructure		*Does the current ICT infrastructure support open data?*
	Source of Funding		*Where is the main source of funding of the open data program?*
Legislation		Eskelinen et al. (2008), Gibbs et al. (2003)	*Under what legislation is the opening of data based on?*
Social and political context:	Political constituencies	Eskelinen et al. (2008), Schwegmann (2012), Gibbs et al. (2003), La Porte et al. (2002)	*Do the existing political constituencies influence open data for personal gain?*
	Cultural dimensions		
	Cultural beliefs		*What beliefs exist about information sharing?*
Drivers or forces for opening data	Event based		*What event/s triggered the open data movement?*

Element	Literature	*Description*
Various forces		*What other forces influenced the opening of data?*
Counter forces that hinder the opening?		*What non-technical issues create a barrier to open data?*

Table 5 Policy context elements

2. Policy content

For the policy content category, it is further refined into sub-categories which are more relevant in terms of national policies. It includes technical aspects of the process and technical aspects of the data itself. The difference with the framework that was developed by Zuiderwijk and Janssen ((to be published)) is that the national policies is at more of a high level policy compared to the various governmental agencies that were examined. High level, in this case, means that it is from a national level instead of from specific governmental agencies. This results in a more high level comparison of the policies as well, which requires an adjustment to the elements that are included. Moreover, the list of compared elements must also be derived from the definition of open data that was taken for the purpose of the thesis. Another aspect that is taken into account when formulating the elements in this framework is the previously mentioned data access management domains that are presented by Arzberger et al. (2004). The elements are further categorized into these domains which are shown by the abbreviation for each domain in brackets. The elements to be included in the framework are:

Category	Elements	Literature	*Description*
Technical aspects of the open data process	Licensing (LP)	Rothenberg (2012), Jaeger and Bertot (2010), Sunlight_Foundation (2010), Arzberger et al. (2004)	*What are the conditions that must be fulfilled in terms of ownership of the data?*
	Fees for access (FB)	Sunlight_Foundation (2010), Arzberger et al. (2004)	*Do users pay to gain access to data?*
	Data presentation (LP)	The_White_House (2009), Huijboom and Broek (2011), Arzberger et al. (2004)	*Is the data presented in a user friendly way?*
	Restricted data (LP)	Jaeger and Bertot (2010), Arzberger et al. (2004), Rothenberg (2012)	*What data is not allowed to be open?*
	Contact with data users (IM)	Rothenberg (2012), The_White_House (2009), Arzberger et al. (2004)	*What mechanisms are available to create a feedback loop with the users?*
	Amount of published data	Zuiderwijk and Janssen ((to be published))	*How many datasets are currently published for re-use?*
	Processing of data before publishing (T)	Rothenberg (2012), Jaeger and Bertot (2010), The White House (2009), Arzberger et al. (2004)	*What processing is needed before it is published on the portal?*
	Costs for opening (FB)	Sunlight Foundation (2010)	*What costs are involved in opening data?*
Technical aspects of the data	Types of data (LP)	Zuiderwijk and Janssen (to be published), Arzberger et al. (2004), The White House (2009)	*What categories of data are made open?*

Category	Elements	Literature	*Description*
	Data format and standards (T)	Rothenberg (2012), Huijboom and Broek (2011), Sunlight Foundation (2010), The White House (2009)	*What formats and standards need to be followed when publishing data?*
	Data quality (T)	Rothenberg (2012)	*How is the quality of the data maintained?*
	Provision of metadata (T)	Rothenberg (2012), Huijboom and Broek (2011), Zuiderwijk and Janssen (to be published), Arzberger et al. (2004)	*Is the provision of metadata regulated?*
	Interoperability with other data (T)	Rothenberg (2012), Huijboom and Broek (2011), Zuiderwijk and Janssen (to be published), Arzberger et al. (2004)	*Is interoperability with other data regulated in the policy?*
	Accessibility of data (T)	The White House (2009), Huijboom and Broek (2011), Jaeger (2007)	*How can users gain access to the data?*
	Encouragement for data re-use (CB)	Rothenberg (2012), Huijboom and Broek (2011), Zuiderwijk and Janssen (to be published), Arzberger et al. (2004)	*Is data re-using strongly encouraged by the data providers?*

Table 6 Policy content elements (technological (T), institutional and managerial (IM), financial and budgetary (FB), legal and policy (LP), cultural and behavioral (CB))

3. Policy impact

As explained through the benefits of opening data, it is hoped that these policies have the capability to realize some of the benefits and bring a significant impact to the society. However, because of the fact that open data is a relatively new field that is being done the clear impact from opening data is still considered un-measureable. Keeping that in mind, there are a few elements that are included in the framework to compare what the impact has been so far in the different countries. The performance indicators taken here are similar to those stated in the framework designed by Zuiderwijk and Janssen (to be published).

Element		Description
Re-use of published data	Huijboom and Broek (2011), Zuiderwijk and Janssen (to be published)	*How is the data being actively used for re-use by the targeted group of users?*
Possible predicted risks	Huijboom and Broek (2011)	*What risks are predicted to happen from the release of open data?*
Benefit alignment with motivation	Huijboom and Broek (2011)	*Are the current benefits of open data aligned with the original motivation to open?*
Public values	Zuiderwijk and Janssen (to be published)	*n/a*

Table 7 Policy impact elements

An important factor to consider when analyzing policies is the impact of those policies on the public, especially when the intent of open data is frequently stated to benefit the public ((Huijboom and Broek, 2011, Zuiderwijk and Janssen, (to be published), Schwegmann, 2012). However, because open data policies are newly developed and have not yet been implemented long, the impact and value are still yet to be seen.

Moreover, because the impact of open data can be seen when there is sufficient re-use of the data that has been published, the tangible impact is still unforeseen. Many can make estimates as to how big the influence of open data will be, but to prove these estimates can only be done in the long term. As stated in the research by Zuiderwijk and Janssen ((to be published)), “the only possible measure of policy impact is over time” (p. 10). Because of this, although public value of policies is highly important, they are disregarded in this research but are recommended to be studied in future research.

Chapter 3
Analyzing the countries

The following chapter contains the analysis that was conducted on the countries that are observed for the comparison. Different from the information that was given in the previous chapter, this chapter will focus on the results that were obtained from the interviews that were conducted and also on the analysis based on cultural dimension literature and the influence on open data. The summary of the interviews that were conducted can be read in the appendix of this report. This chapter contains only findings that are relevant for the comparison. The contents of this chapter will be further analyzed in the comparison in relation to the other countries in the following chapter using the proposed framework. First, the findings from the interviews are explained followed by the cross cultural analysis. Second, the forces and counter forces that influence open data in each country are presented.

a. Findings from conducted interviews

Literature and theoretical studies are not enough to gain knowledge about a certain topic. Additional knowledge from people that are active in the field can be used as another perspective and to gain more information that is not founded in literature. Because of its unique nature, interviews and the knowledge gained can be the added value that a research can offer. In the case of the research presented in this report the interviews were conducted with sources that have an involvement in the development of open data in specific countries and also globally. As can be seen in Table 1, a total of eight interviews were conducted with sources from all five countries that are compared. With the exception of the first interview, the interviews focused on verifying the findings about each country and also to gain more information about the various elements of the framework that were not founded in literature. Elements of the framework were questioned without showing the framework itself to the interviewees. The questions were formulated as open questions to stimulate more information instead of closed questions. For each country there were three main themes that categorized the questions throughout the interview. Each interview was also recorded to create a more accurate transcription of the conversation. Because of the global scale under which the interview was conducted, most of them were conducted via Skype. The following is the list of themes and questions of the interviews:

I. General about open data

1. Can you tell me how you are involved with open data?
2. In this country what is the main goal or strategy of open data?
3. Why is it beneficial for the country at this point?
4. What event started the open data/open government initiative in this country?
5. Is all data subject to publishing on the open data portal?
6. What is the official legislation for open data?

II. Open data policies

7. What policies exist on a national level that helps regulate the opening of data?
8. What areas do these policies regulate?
9. What are the licensing conditions?
10. What is required from the users to access the datasets?
11. What dataset formats is required?

12. How do the datasets receive feedback?
13. What type of preprocessing is required before publishing the data?
14. What are the estimated costs of opening data? What do the costs cover?
15. How is the quality of the data maintained?
16. How is the inclusion of identifiers regulated?
17. How is the interoperability with other data ensured?
18. How effective have the policies been in ensuring the opening of data?
19. Would you agree that in practice these elements are efficiently regulated? What challenges do they create?

III. Open data impact

20. By nature, is it easy for the involved organizations/agencies to open their data?
21. What underlying beliefs or habits create a barrier?
22. How is the impact of opening data checked or measured?
23. In your opinion are the efforts/policies enough? What is being done to improve or expand them?
24. Are there any future improvements in store for this country in terms of open data or open government?

After conducting the interview and writing the transcript and summary, each interview source was given the chance to reread the results of the interview. Some of the sources provided follow-up answers or information and also corrections to the written summaries. The following sections will further detail the findings that were uncovered from each of the interviews. The complete details of the interviews can be found in the Appendices.

General open data

Under the general theme of open data, an interview was conducted initially to explore the possible research areas that still need to be conducted. This interview was the starting point of the research which assisted in narrowing down the scope of the research to the comparison of open data policies. The results of the interview further encouraged the continuation of research that is presented here.

It can be seen that there are many different aspects that need to be considered when launching a project of this kind. Even with an open data platform that is targeted for a certain group of researchers and citizens that narrows down the scope of the platform, this is still considered to be difficult and complex to handle. There is observed to be a problem with incorporating all the user requirements for use of the platform. Because of the target groups of this platform, there are also specific user requirements that are expected to be fulfilled. From experience with the platform, it can be seen that on a national level, the user requirements would also be as complex if not even more. This sheds some light into the fact that launching a portal, that is ready and capable for data reuse, is not an easy task. To achieve the degree of reuse that open data expects, the portals and the data must be ready for other abilities than just minimum visualization.

In terms of policies, an aspect that is considered highly important in publishing data is the inclusion of proper metadata to explain the data. Without metadata a dataset will be unsearchable or might be misused. This observation, along with other sources that state the same importance, puts the inclusion of metadata as an element that is required for publishing data. Another observation that is made is that there needs to be stronger connection between the open data process and the policies that guide it. A link that can be made is,

for example, the inclusion of regulation about metadata. Although at the moment there are many policies involved with open data, they are not always related to the actual process.

The last analysis that is included from the interview is that "less advanced countries in open data can learn from the frontrunners" (appendix A). The frontrunners that were discussed in the interview include US, UK, and Australia because they are able to make improvements in a relatively short period of time. The next step that is important in the process is to create the platform. From here it seems logical that there needs to be a strong foundation to build a complex infrastructure such as an open data portal. A portal which contains government data without any regulations or guidelines could become useless because the data might be unfit for reuse. This then led to the conclusion to conduct the comparison on the open data policies instead of more general open data or open government. The need for the policies to relate better to the actual process further creates a stronger reasoning for the policies to be well developed as the first step.

United States of America

A finding that is worth mentioning about the US is that it is considered a common activity within the governmental agencies to publish data which did not impose extra burdens or resistance from the publishers of the data. It is stated that "… research paid by tax payers should be available to the tax payers with little or no cost or delay" (Blatecky, 2013) which underlies the various agency's objective in ensuring that their data is available to the public. The open data movement in the US is seen to be a continuation of the existing FoIA. The extension involves making it readily available and in machine readable format. Also, the act of making data readily available for the public is not a strange or new thing for the government to do and there have not been considerable issues with changing this mindset. This is also a point where the US has a lead over the other countries. Because the US has had the FoIA for a longer period, making data public is not a new concept. Even though the FoIA can be cumbersome, it is clear that it has helped the opening of data as specified by the open government directive (OGD). It can be concluded that having a legal background for opening data helps the process as it moves towards digital data like open data.

Seemingly the open government and open data movement in US has started strong and now the continuation of the process is being planned by encouraging the innovative reuse of the already published data that is available through the portal. This is a finding that is noted by the interviewee, that open data is now encouraged for the linking and innovative reuse which has slightly shifted from the key motivation that is stated in data.gov. However the OGD does focus on the creation of economic opportunities as well. A key motivation that is seen to be achieved is the collaboration that is formed between the agencies because of the openness. It is easier to form collaborations between agencies because each agency is more transparent about the information that they own and it can be used together.

From the encouraged usage of open data for innovation, this leads to a related finding that can be considered to be unusual. Even though there is a strong intention for the data to be used as linked data, there are currently no guidelines that make metadata a necessary attribute to be included with the published data. As mentioned in the previous interview (appendix A), the inclusion of metadata is vital in ensuring proper reuse of the published data. This finding is also considered to be unusual to the interviewee because it is recognized that metadata on published datasets are important. It is suggested that metadata might be embedded in the process of opening data in the US. Hence, there is no need to formalize it in policies. It may be argued that this is only applicable from a certain agency's point of view. However, when other agencies conduct their monthly meetings, the issue of metadata is never discussed in the formulation of policies. This can either be seen as a strong or weak point in the US. It can be said strong in open data because it is so embedded that the inclusion of metadata does not need to be regulated. But this can also be

a weak point because there is a possibility that important datasets are fall through this loophole and fail to be discoverable for reuse.

Even though the US as a country can be considered to be a frontrunner in the field of open data, within the US itself there is still uneven development amongst the agencies that publish their data. This depends on the mission type and also the level of that agency. This in turn also affects the costs that are involved with publishing the data. Because the open data process is not only about publishing the data, but updating it periodically and maintaining it as necessary. The feedback channels that are open to the public are used as a mechanism to maintain the data and ensure the interoperability. Users are encouraged to write to the publishers in the case that the data that is available is not suitable for use. Besides that, data that has been requested through the FoIA should also be made openly available in machine readable form and available on data.gov. This shows there is a strong enforcement from the existing memorandums and OGD about the way the data is to be presented and this makes these two documents categorized as the existing policy documents. It was an interesting finding that both interviewees state that there is currently no existing national open data policies in the US. Even though according to the definition that is provided in the theoretical section (p. 20), the memorandums that currently exist on open government and open data fall under the category of policies. It must be noted here that at the time of the interview the newly published memorandum on Open Data Policy in the US (The_White_House, 2013) had not yet been published and was only mentioned during the re-evaluation of the interview summary.

As for the impact of opening the data, the measures being done are very general and only see how many times a dataset has been viewed or downloaded. Another way is by seeing how many new industries (cottage industries) arise that use the data that is published. There is still the need to find the balance between keeping the data as open as possible which means that no registration to use the data is needed but also to keep track on who uses the data and how the data is used. No registration to use the data means it is harder to trace who uses it and for what purpose. One mechanism that is carried out in the US is the monthly meeting between all the agencies that are involved in the open government plan to discuss the challenges and to form new collaborations that use the data that is available. These meetings ensure that there is continuous progress and development in the field.

United Kingdom

In relation to the objectives of opening data it can be seen that there exists different views on what the open data objectives of the UK are. This relates to the type of data that is considered. For research data that is opened under the research councils, the goal is more towards validation of the research that is conducted by peers of the researchers. However, it also looks for ways that the data can be reused for further research. This relates to one objective was also mentioned, to support innovation. This objective is also stated on the data.gov.uk portal which highlights the development of services that use the information. From the government data perspective there is a need for open data for transparency and efficiency in the government.

In the UK it seems that there is also an uneven development of how open the involved organizations and agencies are. It is considered to be a long standing concept for the research councils to make their data available to the public. The only change in the recent years is that it has become more reinforced as a practice when before it was merely a concept that the councils followed. This is an interesting finding because from the perspective of government data, the FoIA was only legalized in 2000 which is different from the US which has had the act for longer. It is also observed that the opening of government data was seen as a huge change within the government and even until now the concept of open data is not a well-

known concept. This is an interesting finding because on one side, the research councils have been making their data available to the public but on the other hand for the government it is seen to be a huge change.

As mentioned previously one of the similarities between opening data within the research councils and the government is that both hope to promote the development of innovative discoveries using the data. According to findings from the interview, this is something that has shifted in the past years. When open data first came to the UK in 2009, it was more promoted by the government for the innovation. However the next year when there was a change of government it shifted to be promoted for transparency. Although there was no real shift in what actually happened or how the data was allowed to be used, it was merely the language the government used to promote the concept of open data. This finding shows that there is a possible link between political views and the openness of a country because different political views can have a different reason to support openness. From this discussion another finding can also be derived, even though there was a change in government within the first year, the open data program in the UK still survived. This shows that the concept of open data was strongly supported by the government even if the promotion of it shifted. The fact that it survived in its early stages also shows that open data was, to some degree, embedded in the processes of the government within that year.

As stated before, there seems to still be an uneven development of open data within the involved agencies across the UK. Another example of uneven development can also be seen across the many cities in the UK. Some cities within the UK participate more in the national effort to open data. It is suggested that how each city or even country implements open data is highly related to the individuals that are involved. Individuals that are enthusiastic about open data go extra lengths to implement it in the area they are residing.

Another finding that can be concluded is that from the many different datasets that can be opened each type have different requirements to be fulfilled. For instance, for scientific data, privacy is something important because research data is collected for a specific research which might not be eligible for different research when it is opened. Another is that scientific data is hard to be available in a standard format because of the diversity. These reasons are also applicable to all data that is published on government portals. Different ministries have different data requirements that need to be considered. Connected to the standardized format, it is important that the data be machine readable but the degree of usability of the data depends on the included identifiers. A solution that might be the way to solve the standardized formatting of data is by adding an extra layer of interpretation as is being developed by the W3C for CSV files. This comes to the opinion that machine readable data is crucial and is more valuable with the interpretation that is also included. At the moment the UK is encouraging data publishers to publish raw data which is also a practice that is being encouraged by the European Commission. This motivates the publishing of data and also makes it easier to reuse. However the extra interpretation is also very important when the data published is raw.

Overall the UK is seen to be flourishing in terms of open data. However, there are currently no mechanisms being done to see what the impact of opening that data. There is a research group that carries out this study to discover the impact which is the Open Data for Developing Countries (ODDC). However, it is still under research and has not completed its findings. In relation to the effect that open data has had, only certain communities understand what open data is and how to use the data that is opened. For scientific data it is targeted for the academic communities at the moment although it has been proven to accelerate the human genome project. On the other hand for government data many citizens in the UK will probably have no idea what open data is. Even in countries that are considered frontrunners there is still a need to create the awareness of what open data is and what the benefits are.

Netherlands

As one of the leading countries in open government (Open_Government_Partnership, 2011a), opening information to the public in the Netherlands is not a problem. However, for now, it is dependent on the FOI law that exists which controls the government to respond to the public's request for information. This law has existed since 1978 and this practice has been relatively embedded because the release of data by request is part of the task. However, demands from open data specify that the data must be publicly available even when not requested. This has proven to be a difficult change in mindset of the civil servants. Making the data readily available constantly is a big change in the process and many concerns are brought to light for this case. It seems that in the Netherlands there is a certain enthusiasm from the data publishers about the concept of open data. Conversely, the uncertainty that lies beyond the opening process creates a barrier. These uncertainties on what happens to the data and what the benefits are is what needs to be proven in order to stimulate the opening of data. The main question that is raised by the publishers of the data is what is the impact and why is it beneficial to open all this data. This is also considered to be the question that is raised by countries globally about open data. As is mentioned in the formulation of the framework (p. 25), impact is something that is very important to measure but open data has not been around long enough to see proof of the benefits.

At the moment the focus that is taken from the Netherlands is to publish as much of the Dutch government data online. There is an issue of creating a balance between what the publishers expect and what is needed to fulfill it. On one hand the publishers of data expect to see what they can gain from opening data but on the other hand this is only possible to prove if there is enough data provided to reuse. An effort that is done is to start with small community projects and try to solve issues that exist in communities using publicly available data. Under the circumstances and also aligned with the objective of open data in the Netherlands, the project using the data is a good example of how to provide the needed proof. The project aims to solve real life situations that are in need of innovative solutions and uses the data that is readily available. This can also provide an example to citizens that they now also have the means to solve problems and may further stimulate the development of applications that are based on open data. When the projects have been realized this can be the proof that the publishers of data need to open more data. The more creative and innovative the solutions that are developed can also show that there is a need for a diverse type of datasets to be able to reach that solution. The more diverse the datasets then there can also be a more diverse solution space to solve the existing issues. An interesting issue that was brought to the table during the meeting about open data was that there is little reporting being done about the government spending data in the Netherlands. This came to a surprise because usually when opening data, one of the first datasets that are required is data about the spending habits of the government. Another finding that seems to be a common issue with the previous countries is that there is uneven development between the cities and municipalities. Most of the big cities within the Netherlands have launched their individual open data portals but other cities have not yet understood the concept of open data.

As mentioned previously, the Netherlands already has a FOI law that regulates the citizen's right to obtain information that is held by the government. However, at the moment there are no official policies that regulate the opening of data in terms of the open data definition for this thesis (p. 10). There are no policies about the format of the data or any other guidelines about the process of opening data itself. There are many guidelines suggested as best practices and provide enough detail to make the process of opening easy. The guidelines that are provided on the portal hope to clarify the many technological issues that exist in regard to opening data. The issues that are faced at the moment are mostly non-technological issues that will take longer to solve. The format of data that is encouraged is mainly raw data. There is currently only a

low level of standardization of the data formats. It is important for the data to have the ability to be linked to other data and also be easy to reuse which is why standardization is also an important issue. However, there is also an persistence that the data be completely open. Completely open means that there are no barriers that may cause the publishers or users to be reluctant to join. Various issues such as certain data formats, inclusion of metadata, insurance that the data is linkable, and data quality mechanisms are only provided as suggestions or best practices. But there is not a strong regulation that enforces it in order to guarantee the complete openness of the data. There are currently changes being done to develop a new FOI law that incorporates the current open data requirements and to further develop the policies.

This shows that there is another dilemma that is involved with the opening of data. Besides the needing of publishers to see the benefits of opening data before publishing more data, there is also a dilemma on the amount of regulation and guidelines that are needed. There needs to be certain policies to enforce the opening of data. As was discussed before, having a FOI law supports the opening of data and guidelines for opening eases the process. But it can also cause some barriers to the process of opening. Having to ensure the readiness of the data is an extra task for the publishers of data. Until the mindset of publishing data is already embedded, adding the extra tasks might cause unwillingness to open. The change in mindset of the civil servants is something that is certainly needed to encourage further opening of data. It needs to be in the mindset that open data is a way for civil servants to achieve their goals and tasks. At the moment the mindset is that opening data is an extra burden.

In the Netherlands there is the willingness to make the data public. But there is also still a lot to be done in order to make the data readily available and not just by request. Policies are still being discussed, which in the case of the policy making process in the Netherlands, will take a significant amount of time. However it can be seen that there are fast developments being done in the area of open data and open government. With the community based projects that are being carried out, there is soon to be proof of the benefits that will encourage more organizations to open data. Besides those projects there are also a lot of initiatives that have started from open data. Although not all of them are successful there are a few that have gained considerable notice with their successful business plans. A difference that the Netherlands has compared to the previous US and UK is that the current data and guidelines are all made available in Dutch. This also questions the openness of the data that is provided. If it is only available in the local Dutch it limits the users of data considerably. The academic community or developers do not only consist of Dutch speakers. From the findings that were discovered based on literature and also the interviews the Netherlands can be said to be in the middle of development but is still quite advanced.

Kenya

In Kenya, the intention for opening data was to stimulate the development of applications that would create more jobs. The case that was presented to President Kibaki at the time was that opening data would not only fuel the development of applications and create jobs but also to enable the government to deliver better services and run more efficiently. This is a finding from Kenya that is considered to be interesting because from the historical events that occurred in Kenya it would be more logical for open data to be implemented for the sake of transparency and accountability of the government. However the focus for knowledge innovation economy is caused by the existing foundation that Kenya had at the time. The country had in place a supportive ICT infrastructure with good fiber bandwidth, a large mobile community and also a large developer community. The developers were calling for data from which they could develop the applications that they had in mind. Based on this the Permanent Secretary at the time Bitange Ndemo committed to provide this data. That is why within the government the angle that is taken for the open data

movement leans towards the innovation stimulation. As is with any country the concept of open data is quite new and is something that needs adjusting to, not only from the government but also from the other segments. It is revealed just 48 hours before the launch of the Kenya open data portal, there was a meeting to discuss what data is intended to be published. Several officials were worried that opening the data would mean the publicizing of government secrets and lead to a “wiki-leak” situation. This is still a common issue within the open data movement, that not many understand what open data means. It can be concluded that the culture of opening data has yet to be embedded within the processes. On one hand the advantage that Kenya has is that there is a clear demand from the developers for this data which creates a certain urgency to publish the data.

Because of these concerns from the data publishers there is still limited data being published on the portal. As far as the research had been done there were some disappointments about the situation of open data, namely (1) the government was not putting enough data online, and (2) not enough developers were using the datasets on the site to create useful apps (appendix F). The research was done in 2012 around a year after the launch of the portal. This may not be enough to fully explain the open data situation at the moment because it is still at a very young stage. At the moment there are many changes happening in the government. Under the new constitution there are no longer provinces but counties. Also there has been a coalition of the ministries. Before there were 42 ministries and now they have been merged to only 18 ministries. These changes affect the way the government works at the moment and needs adaptation before more progress can be made in certain areas. There is uncertainty because from the changes within the government it is not sure where open data will fall under. There needs to be a structure in place for it to launch. These factors can be a contributing factor to the reason as to why there has been little data uploaded or published on the portal. The lack of laws and regulation that insist the government to open data also serve as a reason why there has been no significant development. The reason for the lack of laws is also partly because of the many changes in the government. Based on these findings it can also be concluded that, according to the situation in Kenya, even if the proper infrastructure is in place there still needs to be a certain structure that enforces it.

Another important initiative that has started in response to the open data movement is Code4Kenya. An issue that exists also in the previously analyzed countries is that there is little awareness from various segments as to what open data is and how to benefit from it. Code4Kenya is an initiative to build awareness and embeds it not from the government angle but from the developers, media and, civil society. It aims to help other organizations as well to see what data they themselves own and how they can use and benefit from it on top of the data from the government. Code4Kenya is considered to be on the same track as the type of projects that are being done in the Netherlands. However there is a slight difference in emphasis. Code4Kenya is seen to be more directed towards making the whole of the country more open and more aware of the concepts of open government by also involving organizations and media to open their data instead of just for the government. It can be seen that both initiatives try to show more of the potential benefits by starting on smaller scales and approaching certain groups.

At the moment Kenya does not have any open data policies or legislations that insist on the publishing of data. The FOI law has yet to be passed by parliament since for the past two years. Currently there is not yet a draft policy submitted and the policy itself still needs to be designed and implemented. It is suggested that the development of the policies should be both about the release of the data and also about the guidelines for the process itself. Another point that is important to be included in the process of making the policy is to take into account the many perspectives of the involved stakeholders and also to open the policy making

process for consultation. In the UK, opening the process of policy making for consultation is also done. This is a way to incorporate the many different perspectives that need to be included. For any policy making process, each involved stakeholder will have different concerns and different hopes from the process. With a law in place it will help stimulate the process of opening. Especially from officials that are not opening their data at the moment. There needs to be policy in place that makes it law for Ministries to release data on a regular basis. “Again it is envisioned that once the legislation is passed there will be better release of data that will lead to better uses of the data and that will translate to advantages such as greater transparency more applications being built and greater data use by citizens” is an opinion that is expressed by the interviewee (appendix G).

As mentioned before, there are certain changes happening in the Kenyan government. However, open data is high on the agenda of the government but higher than that is the placing of the proper infrastructure, systems and policies before implementation. In the beginning, the process of publishing data was just to publish it quickly before people could change their minds. There was no focus on implementing the correct structures which is what the focus needs to be now.

Whether or not the citizens of Kenya were influenced by the open data movement is not measured. According to the interviewee, data needs to be disseminated to the citizens in an understandable way. So that it is clear what the benefits are for the citizens. So the data needs to be presented in a way of which citizens can access and understand it. This can be in applications, intermediaries, and community media. The average citizens do not really understand about the actual data but cares more about the services that can be gained. This can be a way to show the citizens the impact of opening data that has a meaningful influence on their lives. This interviewee’s opinion is logical because to the average citizens most data that is published can seem just numbers or statistics that hold little meaning. The same opinion was also founded in literature that stated that it has come to the attention of those involved in the Kenya Open Data Initiative that the data that is published must also make sense to the citizens if they are to use it (Hunja and Estefan, 2013). However it must be taken into consideration that data that is needed by the application developers and also for the purpose of innovation will be raw data that has not yet been processed. From this data it is more possible to create a wide variety of applications. It is still important to include information that can explain what the data is for both giving understanding to the average citizen and also to the data re-users but whether or not the data that is given to the citizens should be in a different format is still questionable.

Indonesia

At the moment, it is clear that Indonesia is still in the beginning phases of open data and open government. There is still a large amount of development that is needed before the program in Indonesia can officially launch. There are many action plans in relation to open government in Indonesia but the things that are pushed are transparency innovation and participation. The implementation of open data concepts are spread out on the many levels of government. There have been many district governments that have made significant advances in improving the e-government services that are available to the citizens. This started from the post-Soeharto era (former Indonesian President 1967-1998) where there was a lot of decentralization in the district levels. The decentralization of the district levels are considered to be one of the events that triggered the open government movement in Indonesia. Besides that, the 2008 FOI law also made it easier to coax the organizations to make their data available. However in the case of Indonesia the FOI law itself has not been enough to push all elements of the government to publish their data. A factor that contributes to this is the current situation of the ICT infrastructure in Indonesia that is still a barrier.

Because publishing data is a big change from the usual everyday processes within the government, a lot still has to be improved in terms of the factors that support the data opening.

One of the reasons why open government was seen to be important to Indonesia is that there needs to be bureaucracy reform, especially in the post-Soeharto era. The opening of data will lead to a more transparent, accountable and efficient government in providing services to the citizens (Sekretariat_OGI, 2012a). This goal is stated under the open government Indonesia website as an initiative for open government in Indonesia. However, bureaucracy reform has yet reached that stage. At the moment bureaucracy reform involves getting points and allowances for fulfilling documentation claims. This can be seen as a way in which the officials can ease into the idea. The current situation may not be the ideal situation that needs to be achieved however there is still a process that needs to be followed before the ideal situation can be reached. From the situation in Indonesia there is a need for opening data but not exactly from the reuse side but to encourage transparency and accountability of the government. There is still a silo mentality for civil servants to keep their data or just not yet an enthusiasm to participate. It is understandable that Indonesia can benefit from the concept of open government and also from open data. Over the past years there have been a lot of cover-ups from the government in relation to the processes that are carried out. The government was seen to be a secretive group and citizens were expected to receive the results blindly. However in the post-Soeharto era a new mindset was taken by the citizens. They expected the government to be more transparent in order to banish corrupted practices that weakened the country. It can be said that there is a demand for open data but whether or not the citizens will be able to reuse this published data is another matter.

In terms of policies and regulations, the current legal framework is deemed to be robust. “The 2008 law on freedom of information to all levels of the government and its entities prevents the withholding of information from the public” states the interviewee. Although there is such a legal framework in place for the opening of data, there is not yet an Indonesian open data portal that is launched for use. There is some uncertainty as to which agency will take the lead and launch a national portal of such. There are beta-formats of portals that include some datasets such as SatuPemerintah.net (open government portal) and SatuLayanan.net (for e-government public services) but the data that is available is only in PDF format. The current FOI law does not regulate any technological aspects of the datasets and only encourages the opening by request. Again this is an area that still needs further development. Considering the stage of the open government program in Indonesia it is understandable that detailed aspects such as data standards, metadata, interoperability standards, and other technological issues are not of the main concern at the moment. There is still a need to actually obtain the data from the government that is publishable. At the moment, the beta-versions of the existing portals only contain a few datasets from selected ministries. For the statistical projects with the National Statistics Agency (BPS) there is developed some metadata based on the Data Documentation Initiative (DDI). Besides that there is no further emphasis to implement such measures. The minimum amount of information that is available about policies and how to open data in Indonesia is because the development of open data has not yet reached that point. It is still in the defining stages of what open data is and what it means to make the data available.

From the interview, it was clear that there is still a high need to further develop the efforts that are being done. There is still a high uncertainty for what is in store for open data in Indonesia. Especially with the upcoming elections for office there is even more uncertainty for the fate of open government programs. Unlike the UK, there is uncertainty whether or not Indonesia’s open government effort will survive the change of government at a young stage. It will be highly dependent on the political agenda of the new

administration. At the moment, the UKP4 seems to be the taskforce that is in the lead to rise and take charge of the open government efforts. Whether or not this taskforce will still exist in the next cabinet is still uncertain. For now Indonesia has made available a beta version of an open data portal and also of an e-government one stop shop. Both of these efforts are noticeable steps forward but have not been launched officially hence no traffic is seen to be generated. These concerns might become less worrisome if there is a demand for open data from the citizens themselves. The existence of the FOI law gives the power to the citizens to be more actively involved in the processes of the government. There is seen to be some demand rising from the smaller communities. Because Indonesia is a country that consists of several small communities, on occasion these communities work together in small community development programs. This is also because of the decentralization of the government; these small communities rise up and take charge of the development in their own communities. At some point these communities will begin to demand data from the government to be able to move forward with their development programs. This is similar to the project in the Netherlands aimed to start opening and using government data on smaller scale projects rather than focusing on the national picture first. Another citizen participation program that was seen to be successful is a program called “Lapor” (in English is translated to Report). This is a portal which enables citizens to report anything related to the government or public services. The program has received considerable traffic from the citizens which brings confidence that programs such as open data will be well received in Indonesia as well.

Being a highly complex country with a wide geographic spread it is difficult to completely open everything straight away. Many measures still need to be taken from the central and local governments. But because Indonesia has also taken a strategic advantage on the international level by being a founder of the Open Government Partnership, there are many stakeholders that are interested in developing the situation in Indonesia. Other interviewees also mention that Indonesia is one of the countries of interests that have many supporters that intend to develop the open government initiative. These factors support the decision to formulate recommendations for Indonesia in terms of developing an open data policy based on the findings in this research.

Summary of findings

From the previous summary and analysis of all the interviews that were conducted for the research there are some aspects that can be concluded as findings that are relevant when comparing open data in these countries. These are also the findings that are relevant when developing the open data policies. Without the intention of generalizing the countries, these findings are seen to be relevant across the countries and shows tendencies of practices or behavior that has been adopted.

1. *There needs to be more reasons to open by seeing first hand for the benefits to realize*
   Both the UK and US are at a stage in their open data initiatives where they are asking the question “now what?” which is when they have opened the data that is needed and are now trying to further stimulate the reuse. On the other hand, countries that are developing their open data initiatives also need to see these benefits in order to have more reasons to open the data. In relation to the policies, there needs to be legislation that makes the government and agencies release the data. Another important aspect is to make the data released in machine readable format for easier reuse.

2. *Opening data is a big change in mindset*
   It is observed that even though in some of the countries (US and Netherlands) the concept of providing citizens with some government data is a common approach, the newer concept of

publishing it online constantly is a big change in mindset. It can be imagined that for countries that are not already accustomed with the concept of freedom of information this is an even bigger change in mindset. As for the policy related to this finding is that with the proper regulation of the licensing conditions, restrictions for the users, what types of data are not supposed to be open, and also about pre-processing of the data to avoid publishing of personal information.

3. *Individuals are important in determining the development and progress of the open data movement*
   A similarity that can be found is that each country has individuals that are strong advocates of open data. Also a similarity is that each interview also states that those individuals are who make a difference in whether an open data initiative can be implemented. This is not something that can be formulated as a policy but it can be considered to be a force that drives the opening of data that can be compared between the countries.

4. *Data that is published must be highly diverse to stimulate reuse*
   An issue that is mentioned with the countries that are still developing their open data initiatives (Netherlands, Kenya, and Indonesia) is that there needs to be a certain critical mass of datasets published before it can be reused in the way that it is intended. But it also cannot be just one type of dataset. For instance there is a high emphasis on publishing spending data, but if only spending data is published then it cannot be developed into innovative applications. A policy that can be included for this is about what types of data are expected to be published.

5. *Making the data understandable to all segments is important*
   Many measures are taken to ensure that data is not just understood by the publishers of the data but also to those that have access to the dataset. The people that have this access can vary widely from application developers, other government officials, and even the average citizen. This is why in each country a similar finding is that the datasets that are published needs to include a set of identifiers. In relation to policies it is important that the data includes metadata that makes the dataset understandable and searchable. Another approach that is founded through the interview is that the data should be given to citizens in formats that are already usable.

6. *There needs to be a balance between having the data completely open and regulating the necessary aspects*
   As with many other regulations there is a question on how much regulation is enough and until what point is it too much. This is also an issue that is faced by the UK and US that already have policies in place and also the Netherlands that is starting to develop their policy. The issues that these countries have in common, at the moment, is not about regulating the release of data but more detail about the process itself. There is a common concern that too much regulation will further create a barrier from the publishers to open the data because the process becomes more complicated. The approach that is observed is that the policies progress at certain pace. The first regulations are more focused to stimulate the release of data. Even in the US, the policy for the data to be in machine readable format is only recent.

7. *A strong foundation for opening data must be laid before it can be implemented*
   The foundation in this case means that there needs to be specific ICT infrastructure, organizational structures, and regulations that are in place before open data can be completely implemented. In

the UK and US there are no concerns mentioned in the interviews in terms of ICT infrastructure and organizational structures because they are already in place. However in the other three countries this is mentioned as an issue that needs to be resolved in order for the process to run smoothly.

b. Cross cultural analysis

In this section the cultural and behavioral analysis of the countries will be explained. An analysis will be presented based on cultural dimensions of the countries that are seen to be relevant to open data and the development of open data policies. Besides that, further analysis into the cultural aspect is also described based on the interviews that were conducted as part of the study.

When identifying the difference between countries, besides the demographics and other facts, cultural differences are also interesting to compare. The comparison of cultural aspects is incorporated because they form the context and the environment of the policy. It can be seen in previous frameworks that culture is also compared, although not specifically in terms of cultural dimensions, in frameworks that compare initiatives or programs (Arzberger et al., 2004, Gibbs et al., 2003, Zuiderwijk and Janssen, (to be published)). In order to compare the cultures present in different countries, the cultural values are needed which are represented by cultural dimensions. A dimension is an aspect of a culture that can be measured relative to other cultures (Hofstede et al., 2010). In the book by Hofstede et al. (2010), there are four dimensions which are called power distance, collectivism versus individualism, femininity versus masculinity, and uncertainty avoidance. The comparison uses these cultural dimensions because it is viewed to provide an aggregate view on how the nation responds to certain events and also these dimensions hold meaning in a comparison because they are relative to other societies. Through previous research it was concluded that the relevant dimension when analyzing open data is collectivism versus individualism (La Porte et al., 2002). The other dimensions were founded to have very little relevance to the openness of a country and thus were excluded from the comparison.

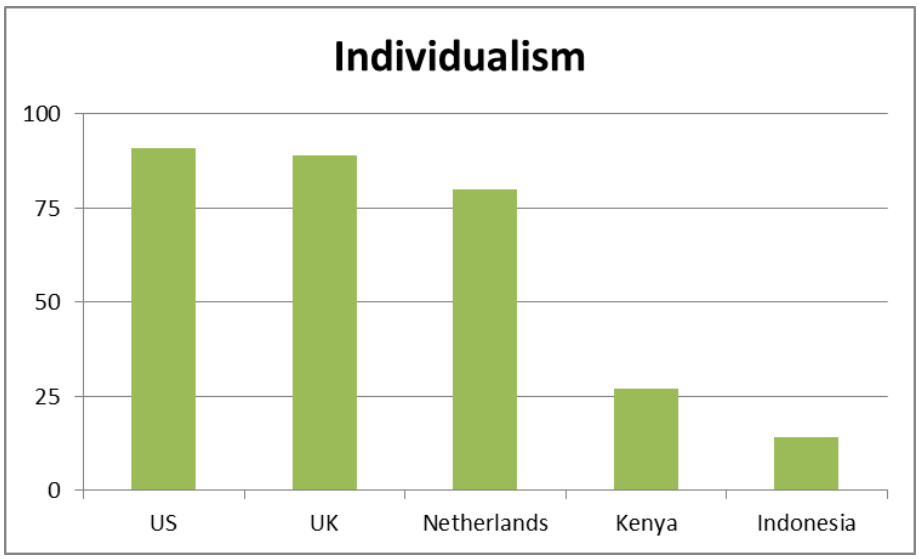

Figure 5 Individualism indexes of the countries

In **Figure 5** the individualism indexes of the US, UK, Netherlands, Kenya, and Indonesia are presented. From the initial look at the indexes, it can be seen that the countries that are categorized as more advanced with open data or openness in general are countries with high individualism indexes such as the US, UK and the Netherlands. A high individualism index means that the country is a highly individualistic society which focuses more on individual needs than the collective needs. On the other hand a low individualism index

means that the country is a collectivistic society where strong cohesive groups of people are formed. In individualistic societies the tendency is individuals think as themselves as “I” that are distinct from other “I”s. From young ages children in individualistic societies are encouraged to be independent and move away from their families when they are able to support themselves. Whereas in collectivistic societies individuals will think of themselves as “we” group instead of singular beings. They form strong relationships that create a mutual dependence within a cohesive group. The bond within this group promises unquestionable loyalty.

To explain why the countries that are seemingly more open have higher individualism indexes, the characteristics that come into play is that in collectivistic countries there is a concept of shame or losing face which differs in individualistic societies. In relation to the opening of data, there is a chance that the data that is published contains errors or estimates that are not accurate and if this is traced back to the publisher of data there is a fear to lose face in the social environment. It becomes more difficult when this losing of face happens to a high ranking official or civil servant that is viewed as a person of power within the environment. On the other hand, in individualistic societies the concept that is used is guilt. Guilt is more of an individual burden which does not influence the appearance in front of a social environment. From the previous evidence given about the countries, the US, UK and the Netherlands are seemingly more advanced with their open data and they are also categorized in the more individualistic cultures. On the other hand, Kenya and Indonesia are more collectivistic cultures which have the tendency to be less open.

Another characteristic that differentiates the two is that there is a strong relationship between national wealth and the degree of individualism. Moreover according to La Porte et al. (2002), there is a tendency for openness to be highly correlated to national wealth. This also argues that individualistic countries are usually more open about their data. The relation between individualism and national wealth has also been presented by Hofstede et al. (2010) which says that national wealth causing individualism is more plausible than vice versa. When a country’s wealth increases, citizens retain access to more resources that allow them to individually conduct their affairs. This however creates vagueness on whether it is the individualistic culture that influences the openness or only the national wealth that then also influences individualism.

Also stated in Hofstede et al. (2010) is Minkov’s formulation of exclusionism versus universalism which is closely tied to individualism versus collectivism. Exclusionism can be defined as the cultural tendency to treat people on the basis of their group affiliation and to reserve favors and services for friends, relatives and other groups that are identifiable. This is close to the characteristics of a collectivistic culture which is identified in Kenya and Indonesia. Outsiders are considered undeserving of these privileges. Universalism on the other hand has the cultural tendency to treat people on the basis of who they are as individuals and not considering their group affiliations.

In relation to open data and the openness of different countries, it can be analyzed that Universalist countries have a tendency to be more open. Because they do not separate into groups they have fewer issues with providing information to people that wish to have it. By having the view that everyone in a country has the same rights to information it is seen logical for information to be published and openly available to the public. This is something that can also be observed in countries that are considered Universalists such as the US, UK and Netherlands. It is observed that these countries consider the process of making data public part of the government process. Requests for this data are handled well even though maybe the release of the data online is not yet 100% optimal. As mentioned before, the Netherlands already has a strong lead on being transparent but is just hindered by the concept of publishing online. On the other hand, exclusionist societies have a tendency to reserve the data within certain groups that they deem as worthy of owning the data. Within the government this can mean that the data is kept within a certain pay

wall of access. This also can be observed through Kenya and Indonesia where the process of making data available to the public has been more difficult. In both countries there was a previous tendency to keep government data a secret from the public. This creates a bigger change in mindset to the current open government and transparency that is hoped to be achieved. In these countries there is also a tendency to view information as a certain power that is only reserved for those of high ranks. This can relate to exclusionism as well as the tendency for officials to keep data inside the groups that produced the data.

Further to these dimensions, another dimension that can be tied to openness is Trompenaars and Hampden-Turner (1998) dimension of universalism and particularism. This dimension can explain the process of formulating policies that guide and support the opening of data in certain countries. Universalistic societies have the tendency to hold rules, codes, values, and standards above the needs and claims of friends and personal relationships. In this case universalistic societies are countries such as the UK, US, and the Netherlands. On the other hand particularistic societies focus more on human friendships and personal relationship rather than formal rules and laws. The countries that fall under this society are Kenya and Indonesia. From the findings in the previous sections it can be seen that universalistic societies have better established policies in place that encourage the release of data. In the case of the Netherlands, at the moment there are no policies in place yet about open data. However there is already a draft being submitted and is in the process of development. There is just a different process of policy making in the Netherlands which takes time. On the other hand in Kenya and Indonesia there are no policies at the moment in place that regulate the opening of data and neither are there drafts being submitted on the matter. In Indonesia the FOI law is also something that is quite recent and has yet to have an influence on the process. Whereas in Kenya there has been a draft for the FOI law submitted to parliament but is still tied up and has not shown progress. Also through the interviews, in Kenya and Indonesia there is a tendency of concern from the government officials about what data is to be released and whether it will affect the perception of the citizens towards the government. This shows that there is a bigger concern about the relationship between the government and the citizens. Similar findings were not discovered from the UK, US or Netherlands where the concerns lay more with how the data will be reused more efficiently.

c. Forces and counter-forces to opening data

In conclusion for this chapter a look into the forces and counter forces that are seen to influence the development of open data policies will also be presented. These forces and counter forces are based on findings from interviews, news reports, and blog posts of people that are involved with open data globally or locally in the countries under observation. The findings here are focused in answering the sub-question of the research question of this research. The forces and counter-forces for each country are presented in the following figures.

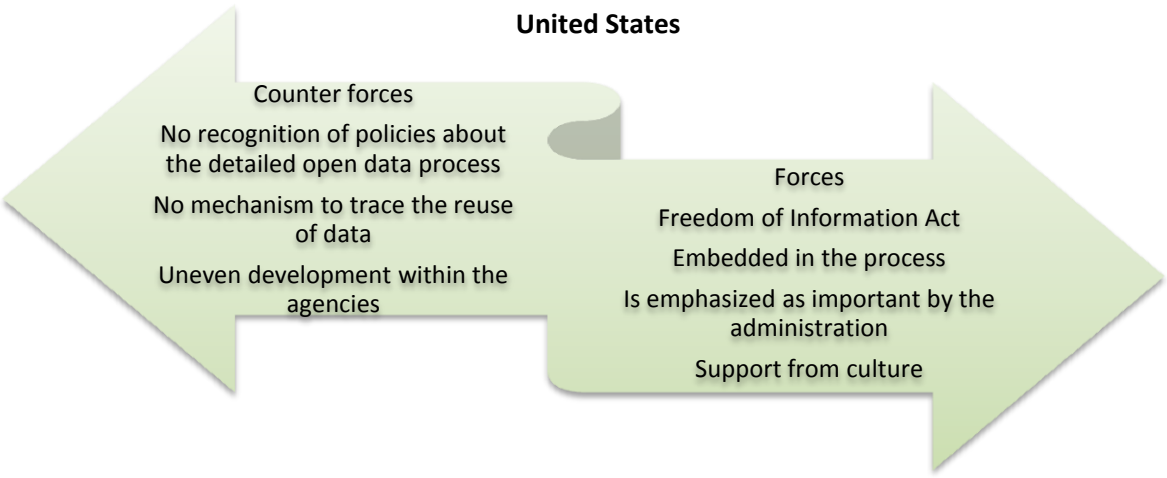

Figure 6 United States forces and counter forces

Figure 6 contains the forces and counter-forces that influence the open data process in the United States. The reason it is said "emphasized as important by administration" is based on the fact that it is the first acts that Obama signed when he started his administration. Besides that there are continuous efforts to improve the initiative by the regular meetings that are conducted between officials of each agency to discuss the further possibilities. Support from culture and the fact that it is embedded in the process are two factors that are related. This means that opening data comes natural.

Figure 7 United Kingdom forces and counter forces

Furthermore, Figure 7 contains the forces and counter forces that influence the opening of data in the UK. The important individuals that are mentioned as a force in this case are David Cameron (the current Prime

Minister) and also the involvement of Sir Tim Berners-Lee. David Cameron is seen to be a strong advocate of transparency that can be reached by opening data.

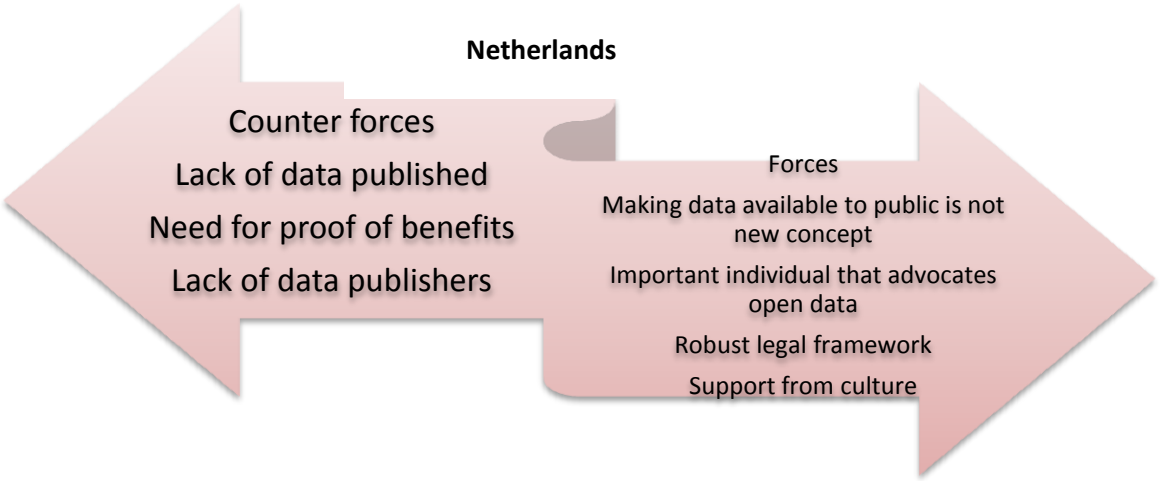

Figure 8 Netherlands forces and counter forces

The forces and counter forces in the Netherlands are shown in Figure 8. It can be seen that there are many forces that support the opening of data but in this case the counter forces are observed to have a strong impact because the Netherlands is not yet at the same level with the US and UK in terms of open data initiatives. A strong force that is also mentioned by Mr. Archer in his interview (Appendix D, p. h) is that Mr. Suijkerbuijk is a strong individual in the Netherlands that strongly pushes for open data.

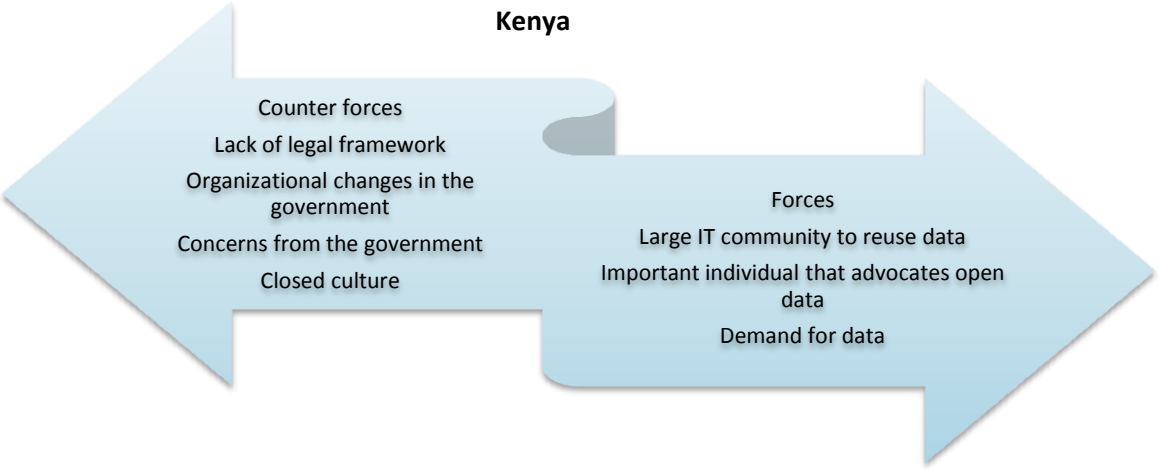

Figure 9 Kenya forces and counter forces

Figure 9 shows the forces and counter forces that influence the situation in Kenya. Even though the individual that is mentioned to have influence is no longer serving in office, he is still considered as an important force in Kenya to start the open data situation at that time. He jumpstarted it and made it possible. Another force that is important to mention in Kenya is that there is obvious demand for data from the developer community.

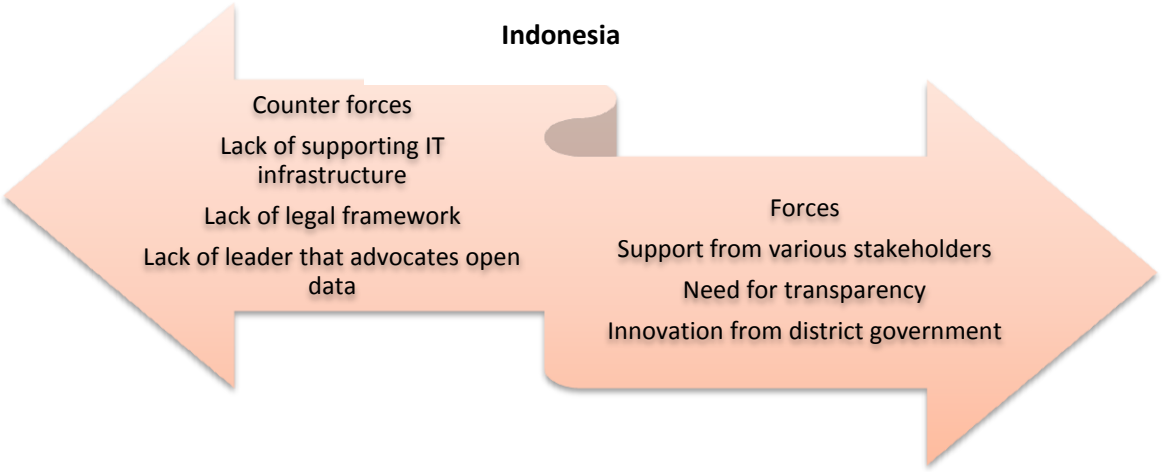

Figure 10 Indonesia forces and counter forces

The last figure shown is Figure 10 shows the forces and counter forces for open data in Indonesia. As mentioned, there are current efforts from various stakeholders around the world that support Indonesia in its effort to have a more open government. Also there are various district governments that show strong tendencies to be able to use the existing IT infrastructures to develop e-government services. However there are several counter forces that need to be addressed before open data can be fully implemented in Indonesia.

The analysis presented in this chapter has led to the addition of elements within the theoretical framework and also provided more information to be compared using the framework. The elements that are added are the different cultural dimensions that are identified to have an influence in the open data process namely Hofstede's individualism versus collectivism, Minkov's exclusionism versus universalism, and Trompenaars' universalism versus particularism.

Chapter 4
Conducting the comparison of open data policies

a. Elements of the designed framework

Based on the theoretical research and also the interviews that were done, it is now possible to finalize the framework. Not many changes were made from the theoretical framework except for the inclusion of the specific cultural dimensions that are seen to have an influence on the open data process. Therefore, the final framework used to compare the countries' open data policies are presented in Table 8 below:

Element category	Element		*Description*
Policy context	Level of government organization		*What level is the policy observed?*
	Key motivations, policy objectives		*What are the motivations to opening data?*
	Open data platform launch		*When was the national open data platform launched for public use?*
	Resource allocation and economic context:	Technology penetration	*According to statistics, how has technology permeated in the society?*
		ICT Infrastructure	*Does the current ICT infrastructure support open data?*
		Source of Funding	*Where is the main source of funding of the open data program?*
	Legislation		*Under what legislation is the opening of data based on?*
	Social and political contexts	Political constituencies	*Do the existing political constituencies influence open data for personal gain?*
		Hofstede – Individualism and collectivism	*Is the country individualist/collectivist?*
		Minkov – Exclusionism and universalism	*Is the country exclusionist/universalist?*
		Trompenaars – Particularism and universalism	*Is the country particularist/universalist?*
		Cultural beliefs	*What beliefs exist about information sharing?*
	Drivers or forces for opening data	Event based	*What event triggered the open data movement?*
		Various forces	*What other forces influence the opening of data?*
	Counter forces that hinder the opening		*What non-technical issues create a barrier to open data?*
Policy Content	Technical aspects of the open data process	Licensing (LP)	*What are the conditions that must be fulfilled in terms of ownership of the data?*
		Fees for access (FB)	*Do users pay to gain access to data?*
		Data presentation (LP)	*Is the data presented in a user friendly way?*
		Restricted data (LP)	*What data is not allowed to be open?*
		Contact with data users (CB)	*What mechanisms are available to create a feedback loop with the users?*

Element category	Element		*Description*
		Amount of published data	*How many datasets are currently published for re-use?*
		Processing of data before publishing (T)	*What processing is needed before it is published on the portal?*
		Costs for opening	*What costs are involved in opening data?*
	Technical aspects of the data	Types of data	*What categories of data are made open?*
		Data format and standards (T)	*What formats and standards need to be followed when publishing data?*
		Data quality (T)	*How is the quality of the data maintained?*
		Provision of metadata (T)	*Is the provision of metadata regulated?*
		Interoperability with other data (T)	*Is interoperability with other data regulated in the policy?*
		Accessibility of data (T)	*How can users gain access to the data?*
		Encouragement for data re-use (CB)	*Is data re-using strongly encouraged by the data providers?*
Policy impact	Re-use of published data		*How is the data being actively used for re-use by the targeted group of users?*
	Possible predicted risks		*What risks are predicted to happen from the release of open data?*
	Benefit alignment with motivation		*Are the current benefits of open data aligned with the original motivation to open?*
	Public values		*n/a*

Table 8 Framework for comparing open data policies

b. Comparing the open data policies

When conducting the comparison of the countries, there are several elements that are already previously compared in the theoretical chapter (p. 13) thus only included partly in this table of results. The elements that are already compared relates to the technology penetration that happens in the country that is shown in Table 3 along with the e-government development index. Both Table 2 and Table 3 that are presented in the theoretical chapter are also considered in the comparison. The results that are presented in the table answers the last sub-question that is in the research question. After Table 9 details each of the elements for each of the countries an analysis based on each of the categories will be provided after as further results of the research conducted. Below is the table of the comparison results:

Framework element	USA	UK	Netherlands	Kenya	Indonesia
Policy context					
Level of government organization	National	National	National	National	National
Key motivations, policy objectives	To implement principles of transparency, participation, and collaboration. Also to create economic opportunities and accountability	Increase the use of linked data standards, kick start the development of services that find novel ways to make use of the information, transparency, and efficiency in the government	Have a more transparent government, stimulate activity, efficient government, create improved services and innovation	Social and economic innovation and data driven decision making but also improved transparency and accountability	Give citizen rights to access PSI, increase public participation in the policy making and decision making process, develop a transparent, efficient, effective and accountable government, develop knowledge and increase public education, create high quality public services and bureaucracy reform
Open data platform launch	data.gov in 2009	data.gov.uk in 2010	data.overheid.nl in September 2011	opendata.go.ke on July 2011	satupemerintah.net is still in beta version
Resource allocation and economic context:					
-Technology penetration (telecommunication infrastructure index from Table 3)	0.6860	0.8135	0.8342	0.1212	0.1897
-ICT infrastructure	Yes	Yes	Yes	No	No
-Funding	Government funding	Government funding	Public funds	World Bank through ICT infrastructure grant	Unclear

Framework element	USA	UK	Netherlands	Kenya	Indonesia
Legislation	The Freedom of Information Act that was adapted in 2009 and the Open Government Directive in 2009 which both state in the event of doubt should always favor openness. Open data policy in which data should be in machine readable format	Secretary of State for Justice's Code of Practice (datasets) on the discharge of public authorities functions under Part 1 of the Freedom of Information Act 2000, issued under Section 45 of the Act in April 2013.	Chapter IV of the Open Government Act (WOB - Wet openbaarheid en bestuur)	Currently no law. KODI is based on a new constitution in 2010 that gives citizens the right to the data and for the state to publish data that affects the nation	UU No. 14 tahun 2008 about openness of public sector information
Social and political contexts, culture in which the opening of data is institutionalized including:					
-Political constituencies	No	No	No	Yes	Yes
Hofstede – Individualism and collectivism	Individualistic	Individualistic	Individualistic	Collectivistic	Collectivistic
Minkov – Exclusionism and universalism	Universalistic	Universalistic	Universalistic	Exclusionist	Exclusionist
Trompenaars – Particularism and universalism	Universalistic	Universalistic	Universalistic	Particularistic	Particularistic
-Cultural beliefs	Already accustomed to opening sharing data	A change in mindset but is becoming more embedded	There is a need to be certain what will happen with the data but is otherwise accustomed to opening data	Previously the government accustomed to keeping information secret	Previously accustomed to a closed government culture

Framework element	USA	UK	Netherlands	Kenya	Indonesia
Drivers or forces that initiated the opening of data, including:					
-Event based	The issuing of the Open Government Directive by Obama	A lunch between the Prime Minister Gordon Brown with Sir Tim Berners-Lee that initiated the launch of data.gov.uk and the technology section of The Guardian started a "Free Our Data" campaign.	After Obama's administration announced the opening of data	Bitange Ndemo became permanent secretary of Kenya's Ministry of Information and Communication in 2005, constitution came about in the wake of the 2007-2008 election which resulted in violence	Since 1998 it started after the decentralization of the government and because of Soeharto's long presidency and the information that he closed from the public to cover his many corruption activities, then 2008 the KIP legislation, 2011 joining of OGP
-Various forces	Existing FOIA, high importance within the government, being embedded in the processes	High level of ambition, high importance within the government, involvement of enthusiasts of open data	Opening data is not new concept, individuals that are enthusiastic, robust legal framework	Demand for data, large IT community for data reuse, individual that are enthusiastic	Support from the international stage, need for transparency, innovation from district government
Counter forces that hinder the opening of data	No recognition of policies that details the open data process, no mechanism to trace the reuse of data, uneven development within the agencies	Segmented to certain communities, high concern of privacy, change in mindset	Uncertainty of the impact of opening data stops most agencies from opening, lack of published data	Civil servants in ministries and departments had a silo mentality and closely guarded all kinds of information, conflict with political interest, lack of legal framework	Upcoming elections bring uncertainty because there has not been a leader that steps up, lack of legal framework
Policy content					
a. Technical aspects of the open data process					

Framework element	USA	UK	Netherlands	Kenya	Indonesia
Licensing conditions	Data is allowed to the extent permitted by law and subject to valid privacy, confidentiality, security, or other restrictions	Encouraged to use the UK open government license (wide use and re-use of data including commercial usage), Non-commercial Government License (for non-commercial use), Charged License in Beta (for commercial purposes that involves payment of a fee and/or royalties)	The data has no copyright or other rights of third parties. If the rights belong to the government then it is available for reuse with the government held liable. Can only be re-used if it falls under the category public as stated in the WOB. No formal license.	Based on the UK open license	Straight forward use and reuse
Fees for access	None	None. In some cases the publisher of data has the right to charge a fee for the re-use of the data	None	None	Minimum fees
Data presentation	Machine readable format	Re-usable, machine readable form, of raw or source data	Register dataset to data.overheid.nl in links or deep links, open on central site or a data API	Advised for raw data	In a user friendly format
Restricted data	Data that is privacy sensitive and has relations with the security of the nation. Also data that provides unfair competition.	Data that is published needs to fall under Freedom of Information, guide to information under publication scheme, data that is exempt from disclosure according to the Act.	There are absolute grounds for refusal and relative grounds for refusal. Absolute grounds include data that might create a threat to the crown, harm the security of the state, company and manufacturing data that is confidential, particular personal data. Relative grounds include international relations, economic/financial	All datasets accessed through www.opendata.go.ke are confined to public information and must not contain information prohibited by the Constitution or the laws of Kenya. The supplying Ministry/Department is required to maintain currency with public disclosure requirements. All information accessed	Information that endangers the nation, that creates unfair competition between businesses, related to personal rights, data that is not yet documented

Framework element	USA	UK	Netherlands	Kenya	Indonesia
			interests of the government, the investigation and prosecution of criminal offenses, if the information is collected first, and respect for private space privacy	through www.opendata.go.ke must be in compliance with current privacy requirements provided for by the Constitution and other relevant statutes.	
Contact with data users	Provide means of contact for the users to provide feedback, input for which information to use and also input for the open government plans	Several active forums and feedback points	There are ways that data publishers can gather information from the users and receive feedback on which they improve the datasets, but it depends on user willingness	Mechanism for users to suggest datasets, recommendations from individuals, groups and organizations regarding the presentation of data, data types, and metadata will contribute to the evolution of open data www.opendata.go.ke	At the moment data is given by request
Amount of data published	73.623	9441	5000	5376	n/a
Pre-processing before publication	-	Anonymous datasets remain stripped of identifiers and also personal data remains personal. If not yet presented in re-usable format, not obliged to change it. Under the definition of datasets it is supposed to be raw data.	No processing before publication, if possible make it in raw data format	-	-

Framework element	USA	UK	Netherlands	Kenya	Indonesia
Costs to open	Depends on the agency because the amount of data published varies	Costs to the public authority of complying with dataset requests, cost of providing it in re-usable format that can be charged to the applicant for making a dataset.	Related to the existing infrastructure if already supports published data	Depends on the ICT infrastructure in place	-
b. Technical aspects of the data					
Types of data	Public data funded by the taxpayers	Health, education, criminal justice, transportation and government financial information	Public data that is collected and produced by the government (environmental and demographic data, data on incidents, mobility, education, connectivity, infrastructure, planning, and the use of public space)	Public information from various government sources.	Information that is related to public bodies, about the work process and activities of public bodies, financial reports, and it also specifies from BUMN, BUMD, political parties, and non-government organizations which data that have to release
Data format and standards	In an open format that can be retrieved, downloaded, indexed, and searched by commonly used web search applications. An open format is one that is platform independent, machine readable, and made available to the public without restrictions that would impede the re-use of that information. At NSF it is converted to CSV or XML	From public data principles: using open standards. From the Code of Practice: must be, as far as reasonably practicable; in electronic form in a re-usable format such as CSV format.	Comply to open standards and be computer readable. The data is offered 'as-is' by the organizations. It is preferable to have it in open standard format (e.g. if in XLS then transfer to CSV) but it's a tradeoff between the amount of work to be done. The open standards are XML and CSV instead of XLS and CSV if possible	File types used by the Kenya Open Data portal CSV, TXT, KML, RDF, RSS, SHP, XLS	At the moment only available in PDF

Framework element	USA	UK	Netherlands	Kenya	Indonesia
Data accessibility	Should be made accessible in online format, also through federal websites of each agency	Easily found through a single online access point www.data.gov.uk without application or registration from the users	Links to data available on portal, data sources are scattered	Through the portal	By request with the information commission
Data quality	Subject to information quality act and meets the information quality guidelines, 'utility,' 'objectivity,' and 'integrity' are the constituents	No tool provided to ensure the quality of data	According to each organization that opens the data, it is of sufficient quality to use.	All information accessed through www.opendata.go.ke should have undergone internal quality control by the submitting/source agency. As the authoritative source of the information, submitting Departments/Agencies retain version control of datasets accessed through "www.opendata.go.ke"	No mechanisms in place.
Metadata	Is not included in policy	Yes. Should be published about the datasets through the access point along with descriptions of format provenance and meaning of data.	Should be included in the description before publishing the data	Citation for each dataset it to be mentioned in the metadata. Was not heard of in practice.	No
Interoperability with other data	Because it is converted to machine readable format it is implied that it will be interoperable	Yes, having linked data forms and using the Government Principles for Open Standards	Making it machine readable should make it interoperable	Not focused on	No
Encouragement for data re-use	Yes	Yes by applying Sir Tim Berners-Lee Five Ranking system	Yes	Yes	No

Framework element	USA	UK	Netherlands	Kenya	Indonesia
Policy impact					
a. Re-use of published data	Not measured directly but there are some cottage industries and it is encouraged	Is evaluated through forums, blogs, and contact with data users	Measurement of re-use is difficult, but it is encouraged to have interaction with re-users and to keep track of the applications that use open data by inviting to a mailing list of re-users	Not enough data available for reuse	Not focused on with no data to reuse available yet
b. Possible predicted risks	n/a	Combination of anonymous data will lead to the revealing of personal information	Possible identification of persons in confidential data, misuse of data, misuse of rights, the disclosure of the data will distort competition	Will lead to challenges by the political parties because of the opening of political agendas	That information will be falsified instead
c. Benefit alignment with motivation	Yes	Yes	Not yet	Not yet	Not yet
d. Value of open data	n/a	n/a	n/a	n/a	n/a

Table 9 Comparison of open data policies

c. Comparison result analysis

The analysis of the comparison conducted is explained in the following section based on the sub-categories of the compared elements: policy context, policy content and policy impact. From Table 9, it can be seen that between the countries there lies some similarities and differences in terms of the compared elements. The results presented are considered the benchmarking of the countries in terms of open data that was compiled based on the research methodology. The analysis in this section answers a sub-question of the research question formulated for this thesis.

Comparing the policy context

In Table 9, it can be seen that all countries that are compared have similar motivations. The three important motivations that are stated from each country are to stimulate innovation, for improved accountability, and greater transparency. The difference lies in the focus of each country towards these motivations. Each country has a different focus that is influenced by the needs of the country. As an example, in Kenya, the main focus was to create new applications based on the data because the country needed to develop new job opportunities and open data was seen as a way to achieve that. A difference between all of the countries is that each country has different forces and counter-forces that influence the opening of data.

Other similarities exist between US and UK that are considered to be the frontrunners of open data in this study. An interesting finding in terms of policy context similarities is that the Netherlands also has many similarities with US and UK. The elements that these three countries have in similar include the readiness of the IT infrastructure, high technology penetration, government type, considerably higher GDPs, an acceptance towards the concept of openness, and similar cultural contexts. These similarities can be identified to have an influence on the development of open data in general however in terms of open data policies there are also other factors that play a role. These similarities help explain why these three countries seemingly have been more receptive towards the concepts of openness but it must be noted that the Netherlands still needs to define their open data policies. The difference that the US and the UK has compared to the other three countries is that both countries have enthusiastic support from the government and a huge movement to make the data available under existing legal frameworks. Both countries have policies that encourage the release of data in digital format.

From the bigger picture of the legal framework, countries US, UK, Netherlands, and Indonesia all have a FOI law that gives citizens the right to access government information. However in the US and Netherlands this law has existed for a considerably longer time which explains why having an open government and being transparent has not been the needed change in mindset. The change lies with the usage of IT infrastructures to make data readily available. Kenya on the other hand is the only country that currently has no laws concerning the release of data.

An important finding is that except for Indonesia, each country has a specific individual that either in the beginning or until now is still involved with the open data initiative in that country. These countries with a leader figure in open data have also been able to launch their open data portals something that has yet to be completed by Indonesia. The launch of a portal for the data is important because it serves as the connecting media between the citizens and the government that enable the benefits to be achieved.

Comparing the policy content

When comparing the policy content of the different countries it must be noted that none of the interviewees in the countries recognized a national open data policy existed in their country at the time of the interview. However for this research, the US and UK are considered to have national open data policies

under the definition of policy. Under this definition in the previous chapter (p. 20) policies for open data are guidelines that can help stimulate the continuous release and reuse of data through a well-thought plan and these characteristics are well founded in the current guidelines that exist under the US and UK legal frameworks in regard to open data such as the Open Data Policy in the US and the Code of Practice from the UK. Guidelines that are available on the portals that suggest the best practices for publishing data are also included in the comparison because it shows that there is a concern in regard to that area even if it is not formalized as a policy. For example, in the Netherlands there are specific guidelines that are suggested as the best practices but there are no policies about any of them. It should also be noted that even the US and UK are still constantly developing their policies through regular meetings with the involved organizations and agencies to discover the current barriers in the process. The interviewees from the two countries state that even though they the countries have been involved in open data longer there is still a need to continuously update and develop the related policies. These two countries also find that the process of opening data has led to better collaboration between the agencies within the government. They also state that the existence of a national policy also helps with the collaboration because it makes it possible to have these datasets interoperable. On the other hand from the Netherlands, it can be observed that all the guidelines for opening are detailed an in place. As previously analyzed, in the Netherlands there is a tendency for there to be very loose guidelines that exist at the moment. This is because there is a need for there to be more data release first before there can be policies implemented in fear that the policies will stifle the release of data. From the results of the interviews it can be seen that from the countries that are still developing their open data initiatives the focus at the moment is not around the detailed policies but more about stimulating the release of data.

When looking at the policy content there are some common aspects that are mentioned in the countries that already have an open data portal (US, UK, Netherlands, and Kenya) such as the need for the data to be machine readable or raw data, data formats, no fees for access, mechanisms to communicate with users, and accessibility through the portal. These elements seem to be common regulations that are included in guidelines that assist the open data process. In Indonesia however there are no guidelines available about such issues. The only mention within the FOI law at the moment is what types of data should be available to the public. This is a similarity amongst all the countries for the data to be public information. There is however a certain emphasis at the moment for the types of data that should be opened in Indonesia is more aimed towards increasing transparency because it should be about the work processes and activities of public bodies and also financial reports.

Issues such as data quality and metadata, even though are important, are only regulated by certain countries. The US mentions about data quality by having an Information Quality Act that is required of all data before publishing but there is no regulation regarding metadata. For metadata, interviewees from the UK and Netherlands state that this is an important aspect that is encouraged in the process of opening data. Kenya also states that the data should have a citation that describes the dataset. All of the interviewees do insist that metadata is an important aspect of the dataset that has to be included to each. One interviewee does bring a different point of view by saying that not only is metadata important but also a full description on how to interpret the data. From the Netherlands, Kenya and Indonesia, although the interviewees recognized the importance of metadata each state that at the moment there is a need for data release and details such as metadata are only mentioned and not focused on.

As seen from the comparison there is a tendency for more guidelines about the open data process when there is a portal available as the platform for the release of data. At the moment the US and the UK are seen

to be the leaders in their development of an open data platform. The platform is seen to have a large variety of datasets, enable the easy search of datasets, and also have options for data processing directly on the website. Kenya also has a well-designed platform at the moment which is partly due to the large IT developer community that is involved in the open data initiative. The Netherlands on the other hand has a very basic national portal which only refers to datasets instead of presenting the data for download. From the interviews it can be found that there is a different approach to open data in terms of policies. In Kenya and Indonesia it is more of a bottom-up approach where the need for data came from the citizens or the users which creates the need to develop a policy or law regarding data release first before the more detailed guidelines. In the Netherlands there are efforts to create a bottom-up approach in order to have more publishers involved, this is also by holding back on the formal policies and allowing the data release flourish first. But in the US and the UK it is more top-down by having the central government push for the release of data first. The common aspect between them is that attention is given firstly to the opening of data before more detailed policies are announced.

Comparing the policy impact

The policy impact between the countries is also provided in Table 9 in relation to how the data is reused, the predicted risks and also the correlation between the achieved and intended benefits. From the previous content of the comparison it can be seen that all the countries are still continuously improving the way they are publishing the data and ensuring that the benefits are felt by all segments. Each of the interviewees spoke highly of the benefits that they wish to see from the open data initiatives. However, whether or not there has been enough impact is not measureable in any country because there are no measures in place to identify them. As has been stated previously, there are different approaches being taken each of the countries. Some are more focused on changing the mindset of the civil servants to release the data hence there is little attention on achieving the benefits.

In the US and UK it is discovered that there is more attention on how the data released is being used. Although there are no direct measures being done the continuous development that is done in these two countries also incorporates the perspective of the data users and not only from the publishers. In the UK it is said that the data reuse is observed via forums, blogs, and contact with data users but it is not done so in the US. In the Netherlands it is encouraged, although not required, for data publishers to keep contact with the data users so there can be some notion as to what is being done with the available data. Interviewees from Kenya and Indonesia however state that there is not yet enough data published for such a measure to be done. This is an interesting finding because through several media outlets it is seen that there is a tendency for questions being raised about what the benefits are for opening data. On the other hand there are no measures being done to be able to observe it.

The risks that are predicted from the release of data varies amongst the countries are risks that are generally predicted with the release of data. However in Kenya and Indonesia there are more specific reasons that relate to the history of these countries in corruption and collusion. There are concerns that information will be falsified or support certain political agendas. During the interviews each of the interviewees felt that it was difficult to identify whether or not the current release of data has been beneficial in practice. Little evidence has been brought to light on how it has affected the average citizen on a national level. Because of this reason the interviewees found it hard to see whether or not the benefits at the moment were aligned with the motivation for opening. Especially in Kenya and Indonesia this was harder to identify because there has not been enough data reuse to observe. In the US and UK the key motivations that are hoped to be reach include transparency, government efficiency and innovation. To some degree, the interviewees of

these two countries, that the benefits were on the right track even if they were still minimum. There is hope that there will be more cases from which more thorough observations can be made. Both also agree that there has certainly been much progress in each of the three goals.

d. Lesson drawing from the comparison

As was mentioned in the methodology, the next step of the research will be the lesson drawing process that follows the ten steps that is presented by Rose (2002). From the ten steps, it can be observed that the first four steps have already been completed in the previous sections of this research and this section will be focused on elaborating steps 5 through 10. However, all the steps will be further detailed below with the steps completed cross reference to their respective pages in the report:

***Step 1**:* a diagnosis of the problem is given in the description of the problem in Chapter 1 (p. 2)

***Step 2**:* where to look for a lesson is identified in Chapter 1 (p. 6)

***Step 3**:* investigation on how a program works there is identified based on literature and interviews in Chapter 2 (p. 13) and Chapter 3 (p. 29)

***Step 4**:* abstracting a cause-and-effect model is completed by the comparison in the framework (p. 48)

Step 5: Designing a lesson

In Rose's (2002) steps there are three ways in lessons can be designed that is adaptation of the model, a hybrid, and a synthesis. Adaptation is a way where a program is very closely related to a single program from a one place. It does not need to be identical but it is very much based on the original. A hybrid is a combination of compatible elements of programs in two or more countries. Furthermore a synthesis is combining in a distinctive way elements familiar in similar programs in different countries (Rose, 2002, p. 13). Based on this and the previous analysis of the compared policies, the following lessons were designed for implementation in Indonesia which includes points which Indonesia could improve:

1. A more robust legal framework

   The legal framework should include a law that regulates the continuous release of data from the government officials in order to reach the critical mass of data that is available to the public. At the moment the law only gives citizens the right to request the data that the government owns. Additionally, the law would further stimulate the opening of a diverse number of datasets and ensure that there will continuously be data that is published from the various governmental entities. There will no longer be the need for the citizens to request the data but this law will ensure its release on appropriate platforms. Taken from the previous findings, it was suggested that each country finds the availability of legislation for data release to be helpful. In the comparison it is also founded that frontrunner countries have legal background for publishing data.

   The legal framework should also include regulation of the minimum number of datasets each government organization should release on a regular frequency. This further pushes the data publishers to prepare the necessary data that is expected to be released. Another aspect to be included in the legal framework is related to the privacy concerns that the data may contain. For example it should be stripped of personal information or information that may lead to the uncovering of personal information that is private. This will provide security to the data publishers that the data will not be used as a weapon in the future to attack the publishers for revealing privacy

sensitive information. It was suggested from the findings that at the moment there is concern from the government side that the information that is published will lead to the release of government secrets. By regulating this under the legal framework it should create more participation to publish data. This legal framework will be the grounds on which other policies are enacted and enforced on the more operational levels.

2. Operational policies

   As previously stated in the findings the policies of open data should be closely related to the actual process of open data which leads to operational policies. These operational policies include aspects of the policies that are enacted in most of the other countries but not in Indonesia. This is a list of many aspects because Indonesia has no policies about the open data process as of now. The operational policies are suggested to be important in order to provide guidelines to the data publishers and data users about the data that is published. For the data publishers these policies provide the guidance that is needed to prepare the data according to what is needed by the users. For the data users the availability of these guidelines convinces the users that the data is published in an understandable and reusable format. The policy should include the following aspects:

   a. The data that is published by the organizations should be presented in machine readable format. Based on the findings in other countries, this is the most effective presentation of datasets. By having data that is machine readable, it will be easier to achieve the intended benefits of opening data by reuse.
   b. Besides being in machine readable formats, the data that is published based should also be made accessible through a national open data portal. This is something that should also be stipulated in the policies in order for the published data to reach the data users in an easy way. Besides being published on the portal, the accessibility of the data should also be free of access and registration. Having the data available on an accessible platform will also provide ease to the data users to get the data that they need.
   c. Not only should the data be machine readable it should also be published in open formats according to standards that are available. Usually the data is presented in CSV, XML, or RDF file types which provide easier reuse for the users of the data. This process in open data also ensures that the data will be interoperable with other datasets and thus create the possibility of linked data.
   d. Datasets that are published should carry specific identifiers that explain and describe the data. The identifier should be in the form of an extra layer of interpretation on any file type which would increase the reusability of the data. The data should contain information about the publisher, the original intent of the data, and how the data was collected.

3. Contact with data users

   At the moment, in Indonesia, contact with data users is mainly done through the request of datasets through emails. This does not create the needed relationship between the data publishers and data users. This relationship is needed to stimulate the provision of data and also the involvement of the data users in the process. In order to create that situation another lesson that should be drawn for Indonesia is to include the creation of public forums, mailing lists, or online portals that enable data users to post comments about the data and also feedback to improve the quality. This will also make it possible for the data publishers to be more efficient in identifying the datasets that are of high

interest and pay more attention on the release of these datasets. This can further attend to the needs of the users to continuously develop the data that is published. There does not need to be a registration process to access data but there should be mediums that can be used to voice concerns in an open forum. This will form the ecosystem between the data publishers and users. This can also be used as a mechanism to maintain the quality of the data that is published. By having users involved in the process, they can spot the mistakes that are in the datasets and have a medium to contact the publishers to have it re-evaluated. Besides the following lesson on data quality, having this form of contact with the data users is also an option to ensure the quality.

4. Data quality

   As mentioned previously, contact between the data publishers and data users can be used as a form of maintaining the quality of data that is published. However, the quality of data should be checked before publishing through mechanisms such as the Information Quality Act. This Information Quality Act is an act that provides the legal grounds that require each data publisher to issue guidelines that ensure and maximize the quality, objectivity, utility, and integrity of information that is published by the organization (OMB, 2002). This counts for each organization that publishes the data, in order for the organizations to have the data updated regularly, and updated according to received feedback from data users. This type of act is beneficial for Indonesia at this moment in order to ensure the quality of the data and further stimulate the release of newer and more updated datasets. This will also provide more motivation for the data users to access the data because the data will be relevant to current situations and applicable for innovative applications.

5. Creating a stronger support system for the open data process

   In order to enable the open data processes and enforce the policies there needs to be other more organizational level foundations laid out for implementation. Based on the findings in the previous chapters, it is suggested that designated agencies are useful to be in place either at a central level or within each of the data publishers' organizations. It would ease and create more reinforcement of the process if it was clearly appointed to a certain organizational entity. It was founded that at the moment that there were confusions of under what agency would the operations of open data and open government run. This also creates a clear responsibility and a more systematic approach to the process of publishing data. This will also address the concerns that data publishers have on where there data will go. Besides that, there will also be clear personnel in charge of the release of data from the organizations. Hopefully by having this organizational change there will be more people that are able to grasp the importance of open data and thus become the individuals that become leaders in the field of open data. As stated before in the findings, there is a certain importance of the individuals that determine the progress of open data. By having an organizational system that supports open data it will be easier for these individuals to work and permeate at all levels in terms of open data. There needs to be certain leaders in the agencies that are designated the responsibility of the open data process. These officials will then be in charge of ensuring the readiness of the systems that are involved in the opening of data.

   This leads to the next part of the needed support system. Besides needing an organizational foundation to build on, there also needs to be some infrastructural support systems in place for open data to be fully implemented. At the moment Indonesia is said to lack in sufficient IT infrastructure to support the entire open data process so this is a point that should be remediated.

However, by having a designated agency to be responsible for open data it can be easier for the infrastructural needs to be organized as well. These specific changes can be included at the policy or legal level as well to enforce the responsibilities that will be given to the designated agency which will further lead to continuous release of data.

6. Creating initiatives at the lower level that triggers more demand for data
   A way to disseminate the data that is published to all the citizens is to present it in an understandable format. Based on the previous findings, this lesson is designed to address the need for data publishers to see what the benefits of open data are and also for the data users to have a better understanding of what they are able to accomplish with the data. A way to do this is to encourage initiatives at the district and community level of the country that uses the data that is provided by the central government. Examples of initiatives in this case is the usage of open data to solve community issues such as that is being done in the Netherlands. It is observed in Indonesia that there are a lot of community development projects starting at the district government levels. A way to approach this is to provide the support these communities need by providing the necessary data and support. Another approach is to start problem solving meetings and try to solve the problems that are brought to the table by the citizens using open data. These efforts will address the need for the data publishers to see what the benefits are for opening data and also this will create citizen's initiative to search for available data. This gives a push to the government to release the data and thus create a demand supply chain for open data.

Step 6: Deciding whether to import

When deciding whether or not to import a lesson it should be both desirable and practical. In this step it is analyzed which of the lessons would be desirable from the point of view of the current government. From the interview that was conducted with the source in Indonesia and also based on findings during the desk research it can be seen that at the moment the focus of the Indonesian government is towards the publishing of data and giving citizens the right to access that data. In consideration of these aspects the designed lessons that are presented above need to be tailored accordingly.

In terms of giving citizens the right to access public data, this is already previously addressed with the 2008 FOI law but it can be further enforced using the points that are mentioned under the lesson about a more robust legal framework. It is observed that at the moment, the less desirable lessons are about the data standards or inclusion of metadata because it is less focused on the interoperability with other datasets. Thus, the operational policies should be mainly focused on making the data accessible and machine readable. Any further regulations about metadata and specific standards may create the unwillingness of parties to participate because they involve very technical work that some might not be prepared for yet. Along the way these points should also be included but it is suggested that at the moment this will only create more barriers. Through the lesson of preparing a stronger support system, it is intended that the involved organizations will be given the training and guidance necessary that will lead to the implementation of the mentioned aspects. However, at the moment they do not fall under the main concerns that need to be addressed. In conclusion of these requirements, it can be said that most of the lessons are suitable for import except for the specific operational technicalities of the datasets.

Step 7: Dealing with resource requirements and constraints

As stated in Rose (2002) 'there is no point in a poor developing country to apply a program that requires the money of a rich OECD country' (p. 14). From the data that is presented in Table 2, it can be seen that

Indonesia has a lower GDP in comparison with the exception of Kenya. However this fact does not influence the policy content related lessons that are already viewed desirable. The policy related lessons are merely further extensions of the current legal frameworks that exist and also additional stipulations to incorporate the more specific open data requirements.

However, the financial constraints do have a bigger impact on the lessons regarding the building of open data support systems and for the initiatives at the district levels that require a source of funding to carry out. The designed lesson about creating a support system is considered to be the one that will incur the most costs. This lesson involves improving the existing IT systems of the government to be ready to support the opening and hosting of published data. It also involves appointing certain individuals to be held responsible of the open data processes. At the moment there already exists a taskforce that is involved with OGI which is the UKP4. Placing the entire program of open data under this taskforce will create a centralized effort with an existing agency. Thus the costs will also be minimized because it does not require building a new organizational entity. Besides that, the improvement of the IT systems is something that needs to be conducted if the government is committed to releasing the data. Even though it will incur many costs, this IT development is also that is important for the development of Indonesia in general so should be done regardless of open data. In the beginning, as a quick solution, the release of data can be done through the available websites and platforms.

Furthermore, creating the community based initiatives may also incur costs to the nation but it is something that is already carried out in certain areas within Indonesia. The support for these initiatives can be by publicizing the initiatives through media thus creating a public awareness of what can be done using open data. Another form of support is by placing them also under the taskforce involved with OGI to make it official and thus creating the supply demand chain for data more possible.

Step 8: Handling the problem of context

Findings suggest that there is a need for the proposed lessons in order to further stimulate the release of data in Indonesia and to make it available to the public. It can also be observed that at the moment the important aspect that needs to be regulated is the supply and demand chain of data and not on the more technical details yet. In relation to the issue of the environment, the lessons that are suggested for Indonesia are also issues that are applicable to the general practice of open data that will be accepted and not only specific for a certain country.

A relevant contextual issue, found based on the cultural dimensions that were previously discussed, is that there is more attention on groups and how relations are built in collectivist countries such as Indonesia. This is why the creation of the supply and demand chain is an important lesson that can handle the contextual issue in Indonesia and lead to the better acceptance of the other lessons as well. Building this supply chain levels the playing field and puts the publishers and users of data within the same eco-system which groups them together. By grouping of this sort there will also be a stronger willingness of the data publishers to release their data because the data users are seen to be in the same grouping. The creation of this eco-system not only benefits the process of data release but also helps the feedback and data quality checking of the data by building this close relation. Once these aspects are well implemented and integrated, the policy related lessons will also be easier to implement in practice.

Step 9: Bounding speculation through prospective evaluation

Prospective evaluation is a form of evaluation that uses evidences from the national past or another country's present to identify what might happen in the future of the lessons that are suggested. The analysis

for this step is provided on a high level because there is currently little evidence on the effects of open data. As also mentioned before, open data is a new concept around the globe with little impact evidence presented. From the lessons that are presented however there shows promise in the prospective evaluation that the lessons will be well accepted by the nation. Based on the findings in the US and UK it can be seen that a robust legal framework and also the regulation of aspects such as suggested (presentation of data, feedback from the users, accessibility through the portal, maintenance of quality, and a support system) that the two countries have succeeded in placing their open data initiatives as the frontrunners. Both countries have been able to consistently publish various datasets from agencies across the government.

On the other hand small scale initiatives have been previously implemented in Indonesia with success. For example, a one-stop e-government service that is provided by Sragen (a small city in Central Java) has shown successful indicators and implementation of similar initiatives are predicted to bring similar successful results. From a prospective evaluation the lessons suggested have shown success in both other countries and also within the nation itself.

Step 10: Using foreign countries as positive or negative symbols

The last step of the process relates to how to market these lessons which is closely related to more of the political aspect of the analysis. Within literature it is suggested that the proposal is associated with a foreign country that can make it more or less attractive. In the case of the research that is presented here, it is clear that the lessons use foreign countries as positive symbols as throughout the research it is based on comparison with countries that are more advanced in this field.

Based on the lessons that are presented in the previous sections, it can be found that there are both policy level lessons and also more operational lessons that can be implemented in Indonesia. Both of these areas are covered by the lessons and not only specific to a certain area. It might also be observed that there are more operational level lessons that can be implemented instead of the policy level. The reason for this is, at the moment at the policy level, Indonesia is suggested to focus more on the release of the data from the organizations and less on the more specific open data details. Indonesia is observed to be in the very early stages of development of open data and there is a stronger need to stimulate the release of data as it is. By implementing the operational level lessons it will create more participation of both data publishers and data users. After there is more awareness of open data and a greater eagerness for the various stakeholders to participate then the more specific guidelines should be implemented.

Chapter 5
Conclusions and Discussions

a. Conclusions

Countries around the world have massive amounts of data that has been collected over the years. Not only is there central government data but also data from the local governments, agencies, various organizations, and also the private sector. There are many things that can be achieved with this collection of data that have yet to be fully discovered. In some countries freedom of information has been implemented within legal frameworks to create a culture of openness. With the recent shift toward a more digital era, there are several possibilities for data to reach its full potentials. Data is made available in formats that are easier to access and reuse. By making these changes and embracing the concept of open data it is possible for nations to improve their government efficiency, increase transparency, and create new possibilities through innovative inventions.

Research has increased in the field of open data to identify what changes can open data really bring to countries. With increasing research into the benefits, many barriers have also been identified. These issues come to play when more organizations are being asked to make their data openly available. The importance of identifying these benefits and barriers is because they lead to the development of relevant policies that highlight the benefits and help mitigate the barriers. Some of the benefits that are identified include economic gains, improved accountability and transparency, increase of public participation, development of new innovative applications, improvement in policy making processes, sustainability of data, and reduction of data duplication. On the other hand the barriers that need to be mitigated include the closed government culture, privacy concerns, lack of attention on data quality, lack of usability, accessibility of the data, and also difficulty from the data users to understand the data. These are issues that are mentioned based on previous research and are relevant to the formulation of open data policies.

As previously mentioned it is important that there are policies that help address the issues that arise when organizations are interested in opening their data. However around the world it can be observed that there is different progress being made between the countries in terms of the policies. Some countries are seen to have more regulation in the field and others are still developing the relevant legal frameworks. The difference in progress can be seen as an opportunity to have each of the countries to learn from each other on how to further improve and develop their open data policies. A way to conduct this learning process is by comparing the open data policies of different countries as a benchmark. The goal of the research is to develop recommendations for Indonesia's open data policy based on a benchmark of open data policies in the United States, United Kingdom, Netherlands, Kenya and Indonesia. Therefore, the main research question of this thesis has been formulated as:

What can Indonesia learn from the open data policies of the United Kingdom, the United States of America, the Netherlands, and Kenya in developing its own open data policy?

The main research question has been divided into sub questions which are answered in the following sections. The first sub question was defined as:

1. W*hich elements of open data policies should be compared to develop recommendations according to available literature?*

In this research, a framework was constructed based on desk research of existing open data policies and relevant comparison frameworks. The frameworks that were included as references varied from frameworks about data access domains, e-commerce diffusion, open data policies in different government levels in the Netherlands, and implementation of open data in various countries. Moreover, some elements were also founded in policy documents from the countries that are compared. The policy documents gave indications to what aspects are important to be included in an open data comparison. Other aspects that are taken into consideration when formulating the framework are the benefits and barriers of open data that relate to the open data policies.

The framework is sub-categorized into the elements that shape the policy context and environment, policy content, and policy impact. From the findings based on literature, there are eight elements that make up the policy context sub-category of the framework which includes the level of the government organizations, key motivations, open data platform launch, resource allocation and economic context (includes technology penetration, ICT infrastructure support, source of funding), legislation, social and political contexts (includes political influence, cultural dimensions, cultural beliefs), forces for opening data, and counter forces that hinder the opening. There are two further sub-categories under the policy content which includes technical aspects of the open data process (licensing, fees for access, data presentation, restricted data, contact with data users, amount of published data, processing of data before publishing, costs for opening) and technical aspects of the data (types of data, data format and standards, data quality, provision of metadata, interoperability with other data, accessibility of data, encouragement for data reuse). The last sub-category for policy impact is identified by three elements including evidence of reuse of published data, possible predicted risks, benefit alignment with motivation, and public values.

The second sub question was formulated as follows:

2. *What are the forces and counter-forces that drive and hinder the open data policies in the examined countries?*

Further research was conducted to discover additional elements or other findings that are relevant for the comparison. This research involved email communication and Skype interviews conducted with sources that were involved in open data in each of the compared countries. A total of eight interviews were conducted with sources that varied from researchers of open data, agencies directly involved with the process of opening data, project leaders, and also from standard bodies. The interviews were aimed at gaining further insight and also verify the initial literature findings in comparison to reality. Further research was also conducted into cultural aspects that influence the opening process in each of the countries. This part of research led to further refinement of the elements in the policy context of the framework. The cultural dimensions were further refined to Hofstede dimension of individualism versus collectivism, Minkov dimension of exclusionism versus collectivism, and Trompenaars dimension particularism versus universalism.

Furthermore specific forces and counter forces are seen to influence the process of opening data in each of the countries based on the interviews that were conducted. These findings combined with the cultural findings based on the dimensions suggest that there is a stronger relation of the development of open data with these specific forces and counter forces compared to cultural aspects. Some of the findings that can be concluded as forces and counter forces include the need for more evidence of the benefits of opening data in order to stimulate further publishing of data, opening data as a big change in mindset in all of the observed countries, individuals are important in determining the development and progress of the open

data movement, there needs to be a large diversity of data types that are published, data must be understandable for all segments even if it is already in an application, a balance is needed between being completely open and regulations, and lastly a strong foundation for open data must be laid before implementation. The reason why these findings suggest a strong impact on open data and specifically on the open data policies is because even though Netherlands is seen to be culturally close to the US and UK, they have not been able to develop an open data policy as well. It suggests that culture has an influence on the openness of a government because the US, UK and Netherlands have in common that they are considered to have an open government culture. However when factoring in the fact that data needs to be published online, the Netherlands still lacks in the area.

The third sub question is formulated as follows:

3. *What are the differences and similarities between the countries when compared using the comparison framework that is based on the elements identified for the examined countries?*

After the comparison of the five countries was conducted using the framework, findings can be concluded based on each of the sub-categories of the framework. In comparing the policy context it can be found that the UK, US and Netherlands, countries which have a stronger tendency to be open, have in common a supporting IT infrastructure, high technology penetration, government type, higher GDPs, and cultural dimensions. However, only the UK and the US are founded to have the needed government and legal support. Another finding is that Kenya is seen to be the only country to not have a legislation regarding freedom of information, which is found to have an impact on the acceptance of open data. It is suggested that context plays a role in deciding the acceptance towards openness in a country but does not influence the existence of an open data policy. This sub-category of the framework also includes findings from the forces and counter forces that are previously explained. Furthermore the results of the policy content comparison suggest that, except for Indonesia, all the countries have already placed certain guidelines for the open data process even though not all of them are formalized as policies. There is a tendency for these guidelines to be available in countries that have an open data portal. The guidelines include aspects such as the need for the data to be machine readable or raw data, data formats that follow open formats, no fees for access, mechanisms to communicate with users, and accessibility through the portal. The inclusion of metadata is only focused on in the UK and Netherlands whereas data quality regulations are only available in the US. This is an interesting finding as all three countries also state reusability of the data as a motivation for opening which is made thoroughly possible if metadata and data quality are well regulated. In terms of the policy impact it can be concluded that all the observed countries are more focused on the process of releasing the data than the reuse of the data. Even though it is part of the goal to stimulate innovation from data release at the moment no measures have been taken to evaluate the reuse of data that has been done. This is where the issue of remaining completely open but being able to evaluate the progress being done plays a role.

From the results of the comparison certain lessons were developed for Indonesia's open data policy. The suggested lessons for Indonesia were developed based on the findings from the different countries and adapted to fit the situation in Indonesia. They were synthesized based on the elements that were found in each of the countries and also based on the analysis that was done on the situation in each country. Based on the steps of lesson drawing, the lessons that are designed for Indonesia are the following:

1. A more robust legal framework that ensures the continuous release of data from the data publishers. The legal framework should include the minimum number of datasets that should be

published on a regular basis and regulation of the privacy issues involved with opening data on a publicly accessibly platform.

2. Operational policies that cover more aspects of the open data process such as machine readability of the data and the accessibility of the datasets on an open data platform without registration.
3. Creating an ecosystem between data publishers and data users by maintaining contact.
4. Ensuring quality of the published data through specific information quality acts.
5. Creating an open data support system through organizational and infrastructural changes. Organizational changes include the designation of open data to a specific agency centrally or specific organizational entity within each data publisher's organization that is responsible for the published data. Infrastructural changes in this case are related to the development of a more supportive IT infrastructure and also building of an open data platform.
6. Creating initiatives at the district levels of the government to trigger more demand for data.

These lessons are suggested to be included in the future development of Indonesia's open data initiative. At this point in Indonesia's development it is suggested that the policies remained focused on the release of data and achieving a certain amount of data release before the focus is on the more technical issues that are involved. Another reason is that the main motivation for Indonesia to open data is for government transparency and accountability. It is seen that issues such as data formats, metadata, and also data standards are more related to the reuse of data that is not the primary focus for Indonesia at the moment. Indonesia is still in the beginning phases of entering the domain of open data and much groundwork needs to be laid down before there can be further implementation.

The other lessons that are suggested also relate to the need for Indonesia to stimulate the data release. The development of a stronger support system is related to the need for a leader to take charge of open data in Indonesia and also for further development of the current IT infrastructure. In terms of the IT infrastructure, related to the data accessibility lesson previously mentioned, Indonesia needs to launch their open data platform to support the release of data that is encouraged. On the other hand, the need for a leader in this case is seen as the important factor because findings in the research suggest that the main reason for most of the countries to have open data initiatives is because of individuals that strongly encourage it and advocate strongly for open data. An individual needs to step up in Indonesia and fight for open data in Indonesia. The other structures of course need to be in place but in order to start the movement there needs to be a leader figure that initiates it.

From the comparison findings it is suggested that the US and UK have a more top-down approach for the opening of data, this means that the process started from centrally enacting certain laws and policies to be followed at the lower levels in the government. On the other hand, in Indonesia, it is suggested that a bottom-up approach was more effective because e-government services started from the district governments and this is the approach that is suggested to be followed. Also in Kenya and the Netherlands this bottom-up approach is seen to be the approach taken by creating initiatives at the lower levels in the government to create a larger demand for the data. This is the importance of the lesson regarding the development of community level or local government level initiatives that use open data.

Based on the findings in this research, it can be concluded that at the moment countries that do not have an open data policy such as the Netherlands, Kenya and Indonesia need to be careful in developing their policies. On one hand there needs to be enough regulation to ensure that there is enough guidelines for the release of data. Each of these countries needs to be more actively involved in releasing the data first. Especially in Indonesia, where there has not yet been launched an open data platform, measures need to be

completed to accommodate the release of data before it can be implemented. The structures need to be put in place for the release of data before attention can be paid to the technical aspects of the data. It is suggested that by regulating the technical details for the datasets to be published extra burden is put on the publishers of data which might create a new barrier and reluctance to open. This must be avoided in the beginning stages.

b. Recommendations for Indonesia

Based on the conclusions that are presented in the previous section it can be suggested that there are a few recommendations for implementation in Indonesia. The conclusions include lessons that are suitable to be implemented in Indonesia and also some other findings that are relevant to the implementation process itself. It can be observed that cultural findings are seen to have a minimum impact on the open data policies and a greater relation is founded from the forces and counter forces. As previously mentioned, the cultural aspects are observed to have a greater influence on the acceptance towards openness. For the situation in Indonesia, this implies that the lessons that were drawn will be affected by the overall response towards openness. It is recommended that Indonesia attempt to implement that various initiatives that use open data before implementing the lessons targeted at the policy level. This will address the cultural aspects that influence the acceptance of openness because it creates a community involved with open data thus bringing down the barriers of collectivism. The issue that was identified is that Indonesia as a collectivistic country is reluctant to share information –by open data- because the government is seen as a different grouping than the common citizen which has different rights. For the policy making process itself it is observed that in Indonesia there is little attention on rules and more towards relationships. In this case it is also important to create the needed relationships by strengthening the ecosystem between data publishers and data users while persisting with the policy making process. It is also suggested that there is a need for a legal framework as a foundation to opening data, despite there being a small influence, the cultural aspects should still be taken into consideration for the implementation.

At the moment, it can be observed that the actors involved with open data in Indonesia, based on the analysis of this thesis, include mainly the central government, district government, and the Indonesian citizens. In implementing the lessons, it is suggested that the first steps involve the further development of initiatives at the community level. This is to build more awareness towards open data at all levels of the country. The government or the data publishers will also be more receptive towards the concept because they are able to see the benefits. On the other hand data users or citizens will be able to see the possibilities that are presented to them by having government data available to use. The district government will be the connection between the central government and the citizens in realizing the initiatives. In the meantime the necessary organizational changes should be implemented as well in order to create stronger reinforcement. This will create a clear line of responsibility for open data and also help in the policy making process. This organization will be an addition to the current actors that will help bring all the parties together and continue to mediate problems that arise.

c. Contributions to research

At the moment open data is considered to be an emerging topic in the research field. More scientific research is being done to identify how newer information technology can be used in creating a more transparent and efficient government. Existing research in the field of open data is often focused on the usage of open data or other activities that can benefit from the usage of open data. This research is focused on guidelines that enable the usage of open data which are open data policies. Because the scope of the

research compares open data policies in different countries, this creates an international setting to the analysis. In the author's opinion, this aspect makes this research unique.

Comparisons of open data programs and open data policies have been completed in previous research (Zuiderwijk and Janssen, (to be published), Huijboom and Broek, 2011, Rothenberg, 2012) but on a different scale and complexity. There is still considerably limited research conducted in the field of open data, specifically open data policies, and the frameworks that were founded in the literature were also limited. The adjustments to the existing framework that is conducted in this thesis are related to the international setting under which the comparison is conducted. With these additions it is possible for other countries to be compared using the same framework and methodology that was conducted in this research.

The thesis also presents a thorough analysis of an engineering concept (open data as an IT-based concept) in a socio-technological setting. It addresses the causes and effects of implementation of open data within societies and also the necessary policies that can regulate the process. The analysis that was conducted balanced the technical aspects and the policy level aspects that are involved with the topic. Interviews were also conducted to broaden the perspective and incorporate stakeholder views on the problem at hand. These elements are highly related to the study that is followed by the author in the Engineering Policy and Analysis (EPA) master program at the TU Delft.

d. Implications for practice

The findings in this thesis also have practical implications for the countries that are compared (US, UK, Netherlands, Kenya, and Indonesia) in terms of continuing the development of their open data programs. For each of the countries, this thesis can be used to draw lessons and also to observe the strengths and weaknesses of their current situation. This thesis provides an extensive comparison which enables each country to learn from each other. With the continuation of development that is possible using the lessons that are presented through the comparison, this thesis also provides the means to help countries in reaping the benefits from publishing data. The research that is presented in this thesis adds to the existing research that is available in guiding organizations, agencies, and countries in reaching their objectives of opening data to the public.

Another implication that is founded in the research for practice is the findings that were observed through the interviews and desk research. It is suggested that there are specific forces and counter forces that play a hand in the development of open data which should be taken into consideration when planning further development. These forces and counter forces have quite an influence on a countries' acceptance towards openness and also towards the development of the open data initiatives themselves.

e. Reflection

In this section, reflections on the research process are presented from the author's personal experience in conducting the research. This includes steps that would have been changed or conducted differently if the author had more time and also a discussion of the complexity of the problem that was encountered.

During the early stages of this research, an exploration of the open data field was conducted to be able to identify a problem that would contribute to the practical and scientific world if addressed. This was in itself a complex procedure because of the many un-addressed issues that follow open data at the moment. The choice of open data policies were chosen because of the high correlation with the EPA master program that is followed by the author. The next step that brought a high complexity to the research was to identify countries that would be suitable to compare. It had to be kept in mind that the chosen countries would have

to be countries that were contactable for interviews and/or meetings for the data collection. Besides that, it should also be countries with sufficient documentation to conduct the necessary desk research. The choice of countries, in the author's opinion, was quite ambitious and proved to bring some difficulties at the data collection stage. However, this also brings the uniqueness and attraction to the topic because the countries that are compared are quite "high profile" countries in the field of open government.

The choice to create an extension on the framework that is presented by Zuiderwijk and Janssen (to be published) in their research was because of the high relevance that the research presented to the topic of choice. Also, there was limited research founded on the topic of open data policies and the comparison and this extension would further contribute to the relevant scientific literature. The framework by Zuiderwijk and Janssen (to be published) provided the necessary categories and aspects that should be compared in an open data policy and it had to be adapted to be able to compare different countries. This adaptation was also a difficult process because there was limited literature on the subject to rely on. The choices that the author made in this regard were also done based on knowledge of the subject in general and not entirely on available literature.

In conducting the comparison itself the author read through many various sources that were available for each of the five countries to be able to completely fill the framework that was designed. Some aspects of the comparison that were not founded in the written sources were then asked, through interviews, to the chosen respondents. This resulted in an extensive comparison involving five different countries which, in combination with the conducted interviews, took the most time to complete for the research. For the comparison the author also had to be critical in determining which aspects were the comparable aspects of the policy to maintain the balance between the technical aspects and the policy level of the comparison.

The author also observed during the research that there is a distinct characteristic to open data policies that differentiate these policies from others. From the observations that were conducted it can be said that data is still a sensitive topic for all countries. Previous behavior in regard to data has always been that the government holds important data that is of high security without the realization that this is not always true. However, this causes open data policies to be a complex policy to formulate. Releasing data, not only from the government to the public, but between any two parties can be a difficult concept to grasp. Open data policies are different from freedom of information laws because they need to regulate the continuous release of data on certain platforms and not only when requested. There are still many people that are even reluctant to fill forms with information unless completely certain what the data will be used for. This is a simple way of looking at the problems with open data. The policies should be able to erase this reluctance by providing guidance on what the data can/will be used for. This is why the author views open data policies as unique policies because there is a need for these policies to be able to change a certain mindset and ease the minds of the data publishers.

Another aspect that is reflected on by the author is the differentiation between 'good' and 'bad' open data policy that was done throughout the research. It is difficult to define exactly what is considered to be good and what is considered to be bad. The author took the understanding that the UK and US are considered to be the countries that have the better policies because there are more aspects that are regulated and has been proven to be successful in their efforts to publish public data. The differentiation also related to the current progress of the open data initiatives in the countries. For instance in the Netherlands although there are no policies in writing, it is still considered to be a country in between the good and bad because of the available guidelines regardless.

As part of the analysis, the cultural aspects of the countries were also observed. In the beginning the author had an opinion that cultural dimensions would certainly affect open data policies. However, as the research continued, there was no literature founded about the influence of culture on open data. After further investigation, the author reached the conclusion that there is very little influence from culture in a specific country on open data. It was suggested through the findings that it is a general change in mindset that happens globally. The determining factors of open data lie in the country specific events, historical backgrounds, and the people that are involved.

Looking back at the research that was conducted during the period of February 2013 to August 2013, this limited time period has implications on the depth of the research. In the event that there was more time to conduct the research the author would have tried to compare two more countries and conduct more interviews for each of the countries. The author would also try to send surveys to people involved with open data in the governments to observe what issues are currently being faced at the operational levels with publishing the data. This will help analyze the open data processes that should be addressed by the necessary policies. Moreover, by adding more countries to the comparison the findings would be more robust and give further depth to the issue.

f. Limitations

There are however some limitations to the research that was conducted and should be taken into consideration when adopting the conclusions presented in this thesis. The limitations are discussed in the following section.

The first aspect of the research limitation is that the study relies heavily on a combination of information that is available on various open data portals of the countries that are studied, various policy documents, and also responses from a number of interviewees that are involved with open data initiatives in each country that were willing to contribute to the research. Much of the information can be considered a matter of opinion. No direct observations of the involved actors were part of the methodology which also limited the analysis of the concepts that were founded in the documentation.

Second, each of the countries that were observed had differently structured open data policies. There was a fragmentation of the information across available policy documents. As a result information that was included in the comparison relied on the definition of open data policy that is presented in this research. This is considered to be a limitation, because this limits the scope of the research to guidelines that fall under the definition. The official policy itself may have been less extensive than is described in the findings.

Third, the policies that were observed for this research were limited to the policies and guidelines that are available at a national level. This creates a limitation because in some countries the policies that are developed at a lower level are more detailed compared to those at the national level. Moreover, not all open data platforms that were available in the countries were observed as the research was focused on the national platforms and the regulations that apply to those. Platforms available at the ministry level or the district governments were out of the scope.

Fourth, due to the short time frame of the study it was difficult to verify the findings of the benchmark in detail. However it is believed that the benchmark gives an aggregate view of the state in the several countries. The interviews that were conducted managed to address some of the issues and give a clearer understanding of the information that was initially founded in the documentation.

Fifth, because the conclusions in the research hope to improve the situation in Indonesia, this creates a limitation of the lessons that were drawn from the findings. The comparison itself can be used to develop lessons targeted at any of the other countries that are included in the comparison as well.

Sixth, as mentioned also in the theoretical overview, this research is limited in its findings because it does not compare the impact of the policies on the countries themselves. Because open data is considered to be a fairly new development, there is still little evidence about the concrete impacts that it has brought. If possible to include in the comparison, this will also affect the findings that are presented in this thesis.

Finally, other limitations include lack of access to people with sufficient knowledge to open data processes in different countries, very little previous research to draw from, and a lack of time to complete the extensive research.

g. Discussion for further research

The current issue at hand that is true for all the countries that were observed is that there lies concern from the publisher side of what is being done with the data that is published. From the US and UK, questions are being asked about what is being done with the large amounts of data that has been published already. On the other hand in the Netherlands this question actually creates the barrier to the release. Evidence of the benefits is needed before they are willing to release the data. In Kenya and Indonesia the same concerns are also voiced from the government but there is a tendency to be concerned about security and privacy. These reasons show the importance of discovering evidence on the impact of open data at the moment. This is an important field of research that needs to be completed based on case studies to show how open data has affected nations across the world.

Another point of discussion that can start from this research is the development of lessons for the other countries that are compared in the framework. Each country is still continuously developing their initiatives because open data is still a relatively new field which means there is still room for further improvement. This can be done by using the framework that is presented in this research. Further research could also be done by using the framework to compare other countries with open data initiatives to draw lessons for those countries as well. Countries such as Denmark, Australia, Canada, India, and Philippines are examples of countries that are growing fast in the open data domain that can be reliable references.

Besides comparing different countries, comparisons can also be conducted on different aspects of open data. A comparison could be conducted on the functionalities of existing open data platforms, open data processes within certain areas of government, or on the open data action plans. These comparisons can also be used to gain further knowledge to improve the existing practices.

Based on another finding in this research another suggested path for research is the identification of important individuals in the domain of open data. Findings from the interviews and desk research showed that there is a strong relation between the existence of an individual with the progress and development of an open data initiative in a country. Not only in a nation but individuals are also shown to have an influence on the lower scales within governments as well. Research can go into identifying why certain individuals have a bigger influence than say the existence of only policies. This could also only be a coincidence but also there could be scientific reasons.

Research in the field of open data is currently progressing fast on the global scale with many organizations and governments seeking to find answers about open data. From this research it can be seen that there are

still areas that are not yet discovered. Research can be conducted to be more focused on identifying the awareness of open data in the average public, creating mechanisms to raise awareness towards open data, comparing services that utilize open data, or identifying the efficiency levels of the government because of open data. The many changes and fast development of this domain brings many possibilities for further research.

Bibliography

ANDERSON, J. E. 2010. *Public policymaking: An introduction*, Wadsworth Publishing Company.

ARZBERGER, P., SCHROEDER, P., BEAULIEU, A., BOWKER, G., CASEY, K., LAAKSONEN, L., MOORMAN, D., UHLIR, P. & WOUTERS, P. 2004. Promoting Access to Publich Research Data for Scientific, Economic, and Social Development. *Data Science Journal,* 3**,** 135 - 153.

AUSTRALIAN_GOVERNMENT 2012. Declaration of Open Government. Australia: Australian Government.

BERGMANN, G. 1957. Philosophy of science.

BLATECKY, A. 2 May 2013. *RE: Open Data Policy Research in US.* Type to NUGROHO, R.

CAMERON, D. 2011. *Letter to Cabinet Ministers on transparency and open data* [Online]. Gov.UK. Available: www.number10.gov.uk/news/letter-to-cabinet-ministers-on-transparency-and-open-data [Accessed March 2013].

CENTRAL_INTELLIGENCE_AGENCY. 2013. *The World Factbook 2013-14* [Online]. Washington, DC. Available: https://www.cia.gov/library/publications/the-world-factbook/index.html [Accessed 7 June 2013].

DOLOWITZ, D. P. & MARSH, D. 2000. Learning from abroad: The role of policy transfer in contemporary policy-making. *Governance,* 13**,** 5-23.

ESKELINEN, H., FRANK, L. & HIRVONEN, T. 2008. Does strategy matter? A comparison of broadband rollout policies in Finland and Sweden. *Telecommunications Policy,* 32**,** 412-421.

EUROPEAN_COMMISSION 2003. Directive 2003/98/EC of the European Parliament and of the council of 17 November 2003 on the re-use of public sector information. *In:* EUROPEAN_PARLIAMENT (ed.). Brussels: European_Commission.

GEIGER, C. P. & LUCKE, P. D. J. V. Open Government Data Free accessible data of the public sector. *In:* PARYCEK, P., KRIPP, M. J. & EDELMANN, N., eds. Conference for E-Democracy and Open Government, 5-6 May 2011 Danube-University Krems, Austria. Edition Donau-Universitat Krems.

GIBBS, J., KRAEMER, K. L. & DEDRICK, J. 2003. Environment and policy factors shaping global e-commerce diffusion: A cross-country comparison. *Information Society,* 19**,** 5-18.

HM_GOVERNMENT 2012. Open Data White Paper Unleashing the Potential. *In:* MINISTER_OF_STATE_FOR_THE_CABINET_OFFICE_AND_PAYMASTER_GENERAL (ed.). United Kingdom: The Stationary Office.

HOFSTEDE, G., HOFSTEDE, G. J. & MINKOV, M. 2010. *Cultures and Organizations: Software of the Mind, revised and expanded 3rd ed*, McGraw-Hill, New York, NY.

HUIJBOOM, N. & BROEK, T. V. D. 2011. Open data: an international comparison of strategies. *European Journal of ePractice,* 12**,** 1 - 13.

HUNJA, R. & ESTEFAN, F. 2013. *Open Government sees Promise after Kenya Elections* [Online]. People, Spaces, Deliberation. Available: blogs.worldbank.org/publicsphere/node/6262 [Accessed February 2013].

JAEGER, P. T. & BERTOT, J. C. 2010. Transparency and technological change: Ensuring equal and sustained public access to government information. *Government Information Quarterly,* 27**,** 371-376.

JANSSEN, K. 2011. The influence of the PSI directive on open government data: An overview of recent developments. *Government Information Quarterly,* 28**,** 446-456.

JANSSEN, M., CHARALABIDIS, Y. & ZUIDERWIJK, A. 2012. Benefits, Adoption Barriers, and Myths of Open Data and Open Government. *Information Systems Management,* 29**,** 258-268.

KENEI, S. 2012. Open data: Learnings from the Kenya Open Data Initiative for CSOs. Development Initiatives.

LA PORTE, T. M., DEMCHAK, C. C. & DE JONG, M. 2002. Democracy and Bureaucracy in the Age of the Web Empirical Findings and Theoretical Speculations. *Administration & Society,* 34**,** 411-446.

LUNDQVIST, B. 2012. 'Digital Agenda: Turning Government Data into Gold': The Regulation of Public Sector Information–Some Comments on the Compass-Case. *Available at SSRN 2148949*.

MAJEED, R. 2012. Disseminating the Power of Information: Keny Open Data Initiative. 2-11.

OBAMA, B. 2009. *Transparency and Open Government* [Online]. Washington. Available: www.whitehouse.gov/the_press_office/Transparency_and_Open_Government [Accessed 9 April 2013].

OMB. 2002. *Data Quality Act* [Online]. Washington: The_White_House. Available: http://www.whitehouse.gov/sites/default/files/omb/fedreg/reproducible2.pdf [Accessed March 2013].

OPEN_GOVERNMENT_PARTNERSHIP. 2011a. *The Netherlands Country Commitment* [Online]. Available: http://www.opengovpartnership.org/countries/netherlands [Accessed March 2013].

OPEN_GOVERNMENT_PARTNERSHIP. 2011b. *OGP Official Brochure* [Online]. Available: www.opengovpartnership.org/files/page_files/OGP_Oficial_brochure_1.pdf [Accessed 12 May 2013].

OPEN_GOVERNMENT_PARTNERSHIP. 2011c. *United States Country Commitment* [Online]. Available: http://www.opengovpartnership.org/countries/united-states [Accessed 27 May 2013].

OPEN_GOVERNMENT_PARTNERSHIP. 2012. *Kenya Country Action Plan* [Online].

PELED, A. 2011. When transparency and collaboration collide: The USA open data program. *Journal of the American society for information science and technology,* 62**,** 2085-2094.

PUBLIC_ACCOUNTS_COMMITTEE 2012. Implementing the transparency agenda. *In:* HOUSE_OF_COMMONS (ed.). London.

ROSE, R. 2002. Ten steps in learning lessons from abroad. *Future Governance: Lessons from Comparative Public Policy.* British ESRC programme on Future Governance.

ROTHENBERG, J. 2012. Case Study International Benchmark: Towards a Better Supply and Distribution Process for Open Data. Logica_Business_Consulting.

SCHWEGMANN, C. 2012. Open Data in Developing Countries. *In:* PLATFORM, E. (ed.) *European Public Sector Information Platform Topic Report.* ePSI Platform.

SEKRETARIAT_OGI. 2012a. *Inisiatif OGI (OGI Initiative)* [Online]. Jakarta. Available: http://opengovindonesia.org/?page_id=548 [Accessed February 2013].

SEKRETARIAT_OGI. 2012b. *Makna Keterbukaan (The Definition of Openness)* [Online]. Jakarta, Indonesia. Available: opengovindonesia.org/uncategorized/makna-keterbukaan [Accessed 20 February 2013].

SINAI, N. & DYCK, H. V. 2013. *Recap: A Big Day fo Open Data* [Online]. Available: www.whitehouse.gov/blog/2013/05/13/recap-big-day-open-data [Accessed 27 May 2013].

STEWART JR, J., HEDGE, D. M. & LESTER, J. P. 2007. *Public policy: An evolutionary approach*, Wadsworth Publishing Company.

SUNLIGHT_FOUNDATION. 2010. *Ten Principles for Opening Up Government Information* [Online]. Washington. Available: www.sunlightfoundation.com [Accessed March 2013].

TAUBERER, J. 2009. Open Data is Civic Capital: Best Practices for Open Government Data. *URL: http://razor. occams. info/pubdocs/opendataciviccapital. html*.

THE_WHITE_HOUSE 2009. Memorandum for the Heads of Executive Departments and Agencies: Open Government Directive. *In:* OMB (ed.). Washington.

THE_WHITE_HOUSE 2013. Memorandum for the Heads of Executive Departments and Agencies: Open Data Policy - Managing Information as an Asset. *In:* OMB (ed.). Washington, D.C.

TROMPENAARS, F. & HAMPDEN-TURNER, C. 1998. *Riding the waves of culture*, McGraw-Hill New York.

UNITED_NATIONS 2012. E-Government Survey 2012: E-Government for the People. New York: United Nations.

YIN, R. K. 2008. *Case study research: Design and methods*, SAGE Publications, Incorporated.

ZHANG, J., DAWES, S. S. & SARKIS, J. 2005. Exploring stakeholders' expectations of the benefits and barriers of e-government knowledge sharing *The Journal of Enterprise Information Management,* 18**,** 548-567.

ZUIDERWIJK, A. & JANSSEN, M. 2012a. A comparison of open data policies and their implementation in two Dutch ministries. *dg.o '12.* College Park, MD, USA: ACM.

ZUIDERWIJK, A. & JANSSEN, M. 2012b. A framework for comparing open data policies, their implementation and impact at various levels of government. 1 - 33.

ZUIDERWIJK, A. & JANSSEN, M. (to be published). Open Data Policies, their Implementation and Impact: A framework for comparison. *Government Information Quarterly*.

ZUIDERWIJK, A., JANSSEN, M., CHOENNI, S., MEIJER, R. & SHEIKH_ALIBAKS, R. 2013. Socio-technical Impediments of Open Data. *Electronic Journal of eGovernment*.

ZUIDERWIJK, A., JANSSEN, M., MEIJER, R., CHOENNI, S., CHARALABIDIS, Y. & JEFFREY, K. 2012. Issues and Guiding Principles for Opening Governmental Judicial Research Data. *In:* AL., H. J. S. E. (ed.) *EGOV 2012.* Kristiansand, Norway: LNCS.

ZURADA, J. & KARWOWSKI, W. 2011. Knowledge discovery through experiential learning from business and other contemporary data sources: A review and reappraisal. *Information Systems Management,* 28**,** 258-274.

Appendices

The results of the 45 minute interview can be found in the following appendices. The only different interview was the interview conducted with Anneke Zuiderwijk-van Eijk. This interview was done in the early stages of the research and thus used a different approach. The purpose of this interview was explorative into the open data situation from the researcher's point of view. Another interview that was conducted in a different approach was with Rushda Majeed, a source of Kenya open data, which was completed in typed question and answer with follow up email correspondence. This was by request of Mrs. Majeed and the questions were the same as the questions for the other interviews. The following table provides a brief summary of the people that were contacted.

Country	Name	Field of work	Date of interview
United Kingdom	Dr. Juan Bicarregui	STFC Rutherford Appleton - Head e-science scientific applications	9 May 2013
United Kingdom	Phil Archer	W3C on E-Government and Open Data Projects	7 June 2013
United States	Dr. Jose Munoz	National Science Foundation – Chief Technology Officer	7 May 2013
Netherlands	Anneke Zuiderwijk-van Eijk	Delft University of Technology – PHD student	18 December 2012
Netherlands	Paul Suijkerbuijk	Data.Overheid.NL – Project leader	17 May 2013
Kenya	Rushda Majeed	Princeton University – Senior research specialist	22 June 2013
Kenya	Jay Bhalla	Open Institute – Executive Director	4 July 2013
Indonesia	Tina George	World Bank Jakarta – Public sector specialist	15 May 2013

Appendix A: Interview with Anneke Zuiderwijk-van Eijk

This explorative interview was conducted with a PHD student at the Technology, Policy and Management (TPM) faculty at the Delft University of Technology that is conducting research on a topic in the field of open data. Anneke Zuiderwijk, the interviewee, has been conducting research in the field of open data under the supervision of Dr. Marijn Janssen. The idea of the interview was to find a starting point for further research which resulted in an interview that asked more broad or general questions about the topic to gain an overview instead of more specific questions on one subject. The interview was structured according to certain themes of questions starting from her personal opinion and views on open data, the current research that she is conducting, her opinion on current open data process and policies, and the future of open data.

The interview was opened with general questions about her and her background. Afterwards more specific questions about the open data topic were asked. The following list contains the questions that were asked during the interview structured according to the themes of questions:

A. Views on open data
    a. What drove you towards the field of open data?
    b. What is your personal interest in open data?
B. Current research: ENGAGE project
    a. Can you explain what the ENGAGE project is?
    b. What are the difficulties so far with the project?
    c. How do these difficulties relate to difficulties of open data itself?
C. Existing policies and processes

a. What policies currently exist for open data?
b. Are the policies seen as a driver or barrier for the open data process? How?
c. Is there an example of a policy that has helped?

D. The future of open data
a. Is open data attractive for all countries or just specific to some countries?
b. Should all countries follow the open data process?
c. If most organizations open their data, what is the next step to be done?

The end of the interview was concluded with a brief summary of the topics discussed. For the first part of the interview Mrs. Zuiderwijk explained that because of her background in criminology, it led her to work with the WODC which conducts research on crime. However it was then discovered that the WODC held a lot of data that was a result of the research conducted there and that they were searching for a method to open this data for the public to reuse. The WODC worked together with Dr. Janssen to develop a way to open their data and actually do something with the data that had been gathered. She also said that currently in Europe there is a sudden boom in organizations willing to take part in the open data process because of the initiative started by the European Commission. However not many people are aware of the initiative and that open data even exists. She also explained that her personal interest in open data was the fact that it is a complicated topic and she hopes to make it easier for the public and for researchers to use open data. Another reason is that a benefit from open data is for transparency of the organizations involved which she sees to be something important. She thinks that the government especially should be more transparent and more accountable for their actions. Most of the public probably do not completely trust their government, especially in the current situation, which makes the open data process a way for the public to gain more confidence and trust towards them by being able to see what they do.

The next series of questions were focused on the current research being done in the field of open data by Mrs. Zuiderwijk for her PHD research. She is currently working for the ENGAGE project which is a FP7, Framework Program 7, project of the European commission. The ENGAGE project is focused on creating an infrastructure for providing open and linked governmental data that is targeted for use by research communities and citizens in the field of social sciences and humanities. The ENGAGE infrastructure is considered to be a socio-technical infrastructure which is the focus of Mrs. Zuiderwijk's research. However there are certain difficulties that have occurred during the process of the project. One of the main issues is the communication problems that occur between partners of the project. The project now consists of nine partners that include businesses and universities which have different interests for the project. Businesses can be considered only interested in a model that can be used after the project is over and gain profit from the model. On the other hand universities are interested in the actual infrastructure because it is usable for them. For Mrs. Zuiderwijk it is difficult to accommodate all interests and ensure that all the partner's needs are incorporated into the infrastructure. Fortunately at the moment communication issues are being handled fairly well in the project. At the moment the project has resulted in a basic platform which enables downloading, uploading, searching, and minimum visualization possibilities of data. However another difficulty that has risen from this is that it is difficult to identify the user requirements for all the different target groups. Sometimes you are sure that all the requirements are fulfilled but the only way to be sure is to conduct many evaluations for the different groups so the platform can continuously expand the user base. This is what the project is currently doing, going to different target groups (journalists, researchers, citizens, and developers) to see what their interest in the platform is and how they are going to use it. A point that was mentioned is that in order to have a good working platform, there needs to be a large user base to collaborate on the platform. If not there will be no value created with no data being uploaded and reused.

To gain this user base the ENGAGE project must attempt to attract users of other existing platforms to use the ENGAGE platform which is difficult to do at the moment because of the limited features that the ENGAGE platform has to offer.

Mrs. Zuiderwijk already identifies this as a major challenge in the future of the ENGAGE project. As for the open data process itself, still related to the ENGAGE project, there is a problem for organizations in opening the data because they lack knowledge on what the exact benefits are. This is especially because the value cannot be seen immediately but only after a certain period of time. Another problem is that there are privacy sensitivity issues of the data being published and the legal liability when something goes wrong with the data. If for instance a governmental organization publishes a dataset that is in violation of a person's privacy then they can be held liable. This might also create a bad reputation for the governmental agencies which would sever any confidence or trust that is being built. The other problem is ensuring that the data is properly searchable through proper meta data terms for each dataset. There is little use of many datasets if the data is not searchable through proper search terms. This is currently an issue because there is a lot of data available yet it is fragmented over different websites and is not easily discovered because of the lack of searchable meta data.

The third theme of questions was directed for the existing policies that either support or hinder the open data process. As in the previous answers, there are currently many issues with the open data process which are being handled through various policies and regulations to guide the process better. Mrs. Zuiderwijk finds that most of the existing policies help with the process but they are not specific enough to narrow down on a specific issue. On the other hand it could also be said that there are enough policies but the guidelines for the open data process are not derived from them making them unusable. Some policies that are viable include those from the European commission that started in 2003. Currently the policies exist on spatial data or geographic data because the enforce the addition of mandatory meta data for the spatial data. This supports the opening of the data and the addition of meta data in terms of spatial data. This is an example of open data policies that actually help the open data process. In this case it can be said that each governmental agency has its own policies regarding to the opening of data. In order for the various agencies that exist to formulate better policies, comparing the open data policies can help the learning process. By comparing the policies, agencies can learn from each other and implement certain aspects of different policies to improve their own.

For the fourth theme of questions, the future of open data is discussed. With the current advances and discoveries of the benefits in the open data field more countries are seen to be interested to open their data. However Mrs. Zuiderwijk thinks that the benefits might not be true for all countries globally. Transparency, as one of the major benefits of open data, is seen to be more important for certain countries such as most European countries, the US, and Australia. This narrows it down to most developed countries that focus more on transparency of the government. There are some exceptions to this and she identifies Kenya as a developing country that also is interested in developing an open data program which has started with the development of a platform. Mrs. Zuiderwijk does think that each country should eventually start an open data program although it is not at the same time. Some countries such as US, England and Australia can be considered as frontrunners that can make improvements in a short amount of time. But this does not necessarily mean that all countries should follow these countries. It does mean that the less advanced countries in open data can learn from the frontrunners. The next step that should be done is to create a strong platform that can support the user requirements. This is what the ENGAGE project is aimed to

accomplish. All the difficulties that were previously mentioned should be handled in order to support the future of open data.

Appendix B: Interview with Dr. Jose Munoz (United States)

Dr. Jose Munoz is the Chief Technology Officer at the National Science Foundation (NSF) in the US. The NSF is an independent agency whose mission is to promote the progress of science. They are the funding source for approximately 20 percent of all America's colleges and universities. In many fields such as mathematics, computer science and the social sciences, NSF is the major source of federal backing. One of Dr. Munoz's collateral duties is to act as a Senior Account Officer for the Open Government Directive (OGD) activities at the NSF. This includes being responsible for putting together the NSF open government plan. NSF has been actively involved in open data for a while now because every report that the NSF produces is always available. However that data was primarily PDFs and regular reports and not the current requirement for the data to be machine readable form.

Dr. Munoz says that the start of open government in the US can be pinpointed to the Obama administration because the first act he signed as President was the OGD which gave a lot of importance to what his intention was. This was something that he really wanted his administration to follow through on. The main goal of opening data in the US is that the data that has been produced by the federal government over the years should be readily available to the general public. The focus of that was actually with the Freedom of Information Act (FoIA). So this is an extension to that by making the government data openly available and not just by request. This is done so that US citizens can access the information which is thought to be needed to conduct their lives. One goal is definitely to take data from one government agency and "mash it" with data from another government agency and create new information from that. Dr. Munoz also mentions about the creation of cottage industries from this data, an effort that is currently encouraged by the government. The FoIA has been around for a long time in the US which helped the opening of data when it was specifically asked in the OGD. But Dr. Munoz says that the FoIA can at times be cumbersome. It presupposes that the potential requestor knows what information the various agencies have as a result of what they do. This is also one of the things that are addressed by the OGD, to make openly available what types of data each agency owns that is available for access. From the NSF it has always been the nature to make this information readily available.

In connection to the Data.Gov portal, data from the agencies is required to be put on there. In the beginning the OGD stated that each agency upload at least four datasets of high interest. On the portal there are mechanisms that allow people that access the data to get back to each agency and register how easily they were able to download the data or if they had any problems with the data. From this the agencies can respond accordingly. So this mechanism enables feedback from the users and can also be used to maintain the quality of the data. Before anything from the NSF is published on Data.Gov, it goes through a data quality group first. This group makes sure that the data is accurate, readable, and in a format that users can easily use. This quality checking is in accordance with the Information Quality Act and Information Quality Guidelines that are available. Amongst these guidelines though there is no mention of including metadata in the publishing of the data. Dr. Munoz reflects that making the data usable is one of the end goals which means that metadata should certainly be a part of that usability. He finds it interesting that it is not explicitly mentioned and it is necessary to include. Interoperability with other datasets is also something that is encouraged but it is not required. It is required to put high interest datasets available on Data.Gov. However in terms of interoperability it is up to the users of the datasets. If there is a case where users have problems

they are encouraged to write to the publishing agency and state that they are unable to use to datasets for specific reasons. Then the agencies can attempt to correct it.

At the moment there are some agencies that can be considered to be more advanced with their open data plans. Some agencies have an advantage that they have a lot of data that is of general interest to the public. For instance at the National Institutes of Health (NIH), Environmental Protection Agency (EPA) and NASA which have such a rich collection of data. They can also easily identify interesting datasets to publish. On the other hand the NSF's information is more in the nature of demographic information about the science education in the US or what they make. So it is of less interest to be reused by the general public. The advancements in open data in the agencies then depend on their line of work and what their missions are. Even so, the NSF still receives some feedback from the data users about the data that is published.

Across the agencies there are still barriers and challenges that hinder the opening of data. One of the challenges is the amount of work that is involved with keeping the data up to date. Not only were the agencies required to identify four high value datasets when the directive first came out but they were also expected to update it periodically and maintain is as necessary. Some agencies have discovered that there is certainly a cost involved in doing that but there is no additional money being provided to make it happen. The costs involved in making the data available are covered by the agencies themselves. Dr. Munoz did not have information about how much exactly the costs were but it depended on the amount of data that they own.

According to the NSF open government plan, datasets that may hold high importance to the public should be converted to different formats. This is also the case for the other agencies. Part of the OGD says that the data should be available in machine readable form. This includes formats such as comma separated values (CSV), excel spreadsheets (XLS), and XML. The data that has been previously requested through the FoIA should also be made openly available in machine readable form. At the NSF this is included as one of the annual updates that are conducted. Every year the FoIA requests are also made available on Data.Gov. Dr. Munoz says that the NSF has gone back and completed this for data from the past 10 years.

The impact of opening data is checked through two mechanisms through the available portals. First is the raw statistics about the amount of downloads each dataset gets. The other one is the development of a cottage industry using the available data. So, a cottage industry that use and market the data is also measured. Besides these two mechanisms, Dr. Munoz has not seen firsthand in what ways the public has used the data that is available from other agencies as well. He does have knowledge about an upcoming Hackathon in the US which will use the information that is available on Data.Gov, from local governments, city governments, and state governments. The purpose of the Hackathon is to see how cleverly people can use the data provided on Data.Gov through application developments. It is expected that the Hackathon will bring new knowledge and new tools to the agencies in regards to open data. For instance a mash-up that would be of interest to the NSF would be between NSF data and patent data from the Patent and Trade Office. It would be interesting to see how many patents came out indirectly out of something that the NSF funded. There are some difficulties however in discovering how the data is reused. The users of the data are not obligated to register or report that they are using a particular dataset. Dr. Munoz also says that there is no intention to make a policy to ensure that there is a reporting system of this kind. It is the goal to have as few barriers as possible to be able to use the data and to encourage people to use it more. Having a registration system for the data creates a barrier instead.

Currently to improve the open data program there is a lot of activity being generated to looking at mechanisms to making the data that is a result of federal funding openly available. For example, although the NSF makes the data that it has available and openly available, none of the data from the research that the NSF funds is openly available. All of the agencies are currently kicking off activities to make explore that possibility. There is also new activity being generated towards opening publicly funded research data and publications. When the NSF was working on the open government plan for NSF, they opened a mechanism that enabled various users to come back with comments. The number one request through these comments was to make the data that was generated as a result of NSF funding openly available. But it has always been NSF policy that data and information is owned by the institution that conducts the work so it was not NSF's data to make it openly available. Because of this, it will be a long process to open this data. The data that is produced by the researchers is their intellectual property and they can publish papers and receive recognition from that data. The current struggle lies in when to make the data available and when to let the researchers have it to get many publications. As far as the program has been ongoing, there have been at least two or three cottage industries started up. However, it is still quite unclear whether or not the benefits have been optimally reached. Once a month all the federal agencies get together to discuss the OGD and what data is being put together. Besides that collaboration between the agencies are now also easier because of this openness. This is also one of the topics that is discussed during the meetings on how the agencies can work together to make the data more federally useful. It is also encouraged in each of the agency's OGD plan to identify how they will collaborate with other agencies to make their data more useful.

Appendix C: Interview with Dr. Juan Bicarregui (United Kingdom)

Dr. Juan Bicarregui is involved in the opening of scientific research data with the Science and Technology Facilities Council (STFC) which is one of the seven research councils which fund research in the UK. Dr. Bicarregui represents the STFC on a body which is across the seven councils. This body formulates the data policy for research data in the UK. He also runs a group of people in the national laboratory (Rutherford Appleton Laboratory) which is responsible for managing the data that comes from the large facilities in the UK. The laboratory manages the large facilities which are cross disciplinary. This includes the large telescopes, high energy physics with CERN, synchrotron neutron source and other large scale facilities. The data that is opened from these facilities still fall under the umbrella of the open government directive from the UK but is much more specific to a certain area. Dr. Bicarregui says that the main goal of opening data within the research council is to maximize the scientific value of the data and have a broader use. Making the research data available enables peers to validate the results of the publication. Data is collected and analyzed by a person who then publishes it on a paper. Then on the other side the peers check the analysis that was done during the research. A different goal is that the data is made available for other research. This is more difficult to achieve because each individual has to interpret the data in a broader concept.

The STFC defines open data as data that is still under regulation by terms and conditions and not completely available for use. This definition is highly dependent on the discipline that defines it because the data is viewed differently across the disciplines. Because of this there is not a single definition across those disciplines. The different research councils opt for slightly different definitions. However it is certain that there will be certain terms and conditions that regulate it. Certain factors come into play such as privacy and consent. Data that was gathered from human subjects may have been collected for a particular research and consent was given for that research only and does not necessarily mean it can be used for other research.

When asked about the start of open data within the UK or more specifically the Research Council United Kingdom (RCUK), Dr. Bicarregui says that the idea that data behind a publication should be made available is a concept that has been around for a long time but without any reinforcements. The current move for opening up the entire scientific process has been ongoing for the past five years at least. Some disciplines in the council have been longer involved with open data such as the human genome project. For the past 20 years the human genome project has had a very open policy about the data that they collect. It has been seen to help accelerate the research in that field. It has been one of the flagships and evidence of the benefits to opening scientific data. However the data that is published from the RCUK is not included in the Data.Gov.UK portal.

In terms of policies, each of the seven councils has their own data policies with some principles that are cross disciplinary. There is the general reasoning that data should be opened unless proven otherwise included in these principles. Moreover the principles cover data management plans or policies which say that the proposers of data should know from an early stage how they will manage the data. If the data that is collected is seen to have long time value, then this data should also be maintained for the long term. Collection and publication of enough metadata so that the data can be understood by others is also included in the principles. The data itself is more valuable when there are identifiers that explain the data which also makes it easier to reuse. Regarding limitations, the principles recognize that there are certain constraints that must be acknowledged such as data protection, confidentiality, and ethical constraints. These factors may limit the degree of openness you want to make for the data. Another principle addresses the recognition that is expected from the researchers when publishing their research. In most disciplines, the researchers gain credit for publication of the results (i.e. published paper) and not for publishing the data. So there is an incentive to keep the data to publish more papers for themselves instead of making it public and having other people publish using that data. One of the last issues that are addressed in the principles is that it is appropriate to use public funds for this application of making the data sharing as well as the research. Data format standards however are not included in the principles that exist across the disciplines. The formats are specified within each disciplinary because it is hard to require only one data format for the whole of science. In the policies within the disciplines it is usually only stated what format is the best practice. In relation with data formats, interoperability with other datasets is something that is hoped to be achieved. But this is also something that cannot be written in the high level policies because it is highly dependent on the individual disciplines.

In the RCUK, the cost of opening data varies enormously on the disciplines. For some disciplines not much data is available so it is relatively easy to publish it on the existing repositories. In other disciplines, the data is very large so you have to make value judgments on what data must be published and what must be ignored. The value judgment must be made because sometimes researchers believe that the data that is collected cannot be used for any other application so there is no point in putting the extra effort to make it understandable for others. For an example, the Natural Environment discipline has a policy that all the data should be deposited in one of the data centers. These data centers only exist for the big disciplines and are run nationally. All data that is collected from all the programs should be uploaded to the data centers at the end of the research. In this case there are no ongoing costs, the costs are only in running those data centers. About 2% is allocated to the maintenance and management of these data centers which still varies amongst the disciplines. These data centers are also in charge of the data quality. There is a need from the researcher side to make the data ready to upload and publish on the data centers but the data centers then maintain the data.

Dr. Bicarregui reflects that the current policies that exist for open data is a new area of policy and is changing fast. The efficiency of the policies varies across the disciplines because some are more advanced than others. For scientific data the challenge lies in making the data understandable beyond the original team. It is a real task that has to be built into the project from the beginning. This makes the opening of data not easy for anybody. As for the current impact of opening the research data, only specific disciplines conduct such an assessment. The assessment is done in the data centers that manage the whole database of data and they conduct a review of this data. Every three to five years it is evaluated whether or not the data that is uploaded is useful. In order to see how the public responds to the data there are certain mechanisms which allow feedback from users of the data. However at the moment this is limited to the academic community even though one of the anticipated benefits is to have the private sector be able to access and reuse the data. Besides that, data centers also have ways to communicate with the data users and communities. Individual data on the other hand would probably be uploaded somewhere and left to manage on its own. So only the data that comes from large facilities or data centers have contact with the users. Applications or mash-ups that are created from the published data is heard of but not generally well known or focused on.

The policies that regulate the publishing of scientific data within the council are continuously refined. Guidance that is given to researchers is being constantly developed. Regular meetings that discuss the challenges to opening data are held and taken as meetings to also improve the available policies. Across the funding bodies and research institutions meetings are conducted where they share the current best practices which is a push from the policy side. Policy does not say what the best practice is but says to follow them. As for future developments, it is only to continue the refinement of the policy and guidance in the next three to five years.

Appendix D: Interview with Phil Archer (United Kingdom)

Background and the W3C

Phil Archer spends most of his time working for the World Wide Web Consortium (W3C) on eGovernment and open data projects. His involvement with open data is because he works with the world standards body he is more concerned with open data than the other colleagues and comes into contact with open data practitioners. He works with common vocabularies which not only open the data but makes the data more interoperable. The vocabulary here is the vocabulary that is used when the data is generated and also includes meta-standards. Because the W3C is a sub-contractor to Price Waterhouse Coopers which has a contract with the European Commission (EC) a lot of work that he does is related to the EC. That tends to be the development of vocabulary and documents that support governments that are in the process of opening their data. A recent vocabulary that was on a vocabulary that describes public services when data is published about public services. Although his involvement is not specifically for the UK, because W3C is a global organization, because he resides in Britain and lives an hour away from London he comes in contact with the situation in UK more often than anywhere else in the world. His focus is mostly European than Britain. The W3C is a body for open standards on the web which was founded 19 years ago. The current director of the W3C is Sir Tim Berners-Lee the inventor of the World Wide Web. They are considered the industry association where technical standards for the web are developed. Any standards that are developed by Mr. Archer are global standards even if his focus is in Europe. This is actually why the W3C is involved with the EC because that allows European standards to have a global audience.

Start of open data in the UK

In general there are three pillars that are considered the goal of open data everywhere which are transparency, efficiency in the government and to support innovation. This is also the goal for open data in the UK. Each of these pillars has a particular role. In the UK the open data portal, Data.Gov.uk, was the second portal to launch. First the US with Data.Gov and a few months later the UK version also came online. An interesting observation that Mr. Archer pointed out in regards to open data in the UK was that when it started the government was a central left government. They promoted open data particularly for the innovation side of things and to create new jobs in the economy. However, when the central right government came in May 2010, open data was promoted more for transparency than for innovation. Another important observation is that open data survived the transition between governments even though it was less than a year old proving that it was not a temporary thing. Change in government is usually when certain projects like this are stopped. Even with the change in government open data has survived and is now seemingly embedded. This change in promotion however did not affect how the citizens used the data it was still encouraged for all three pillars but the language that was used by the politicians to promote it shifted.

The event that started open data in UK was particularly by a campaign by the Guardian newspaper in 2006 called “Free our Data”. When it started not many people understood what the campaign was about and did not get the value of it. The point of the campaign was that citizens have paid for the data so can we have it. This event raised awareness for the issue and presented some information and research in the field to present their arguments. However this campaign only reached a group of people because the Guardian is not a mass market paper. Even to this day the majority of people in the UK may not have an idea what open data is. This campaign ran for three years before Sir Tim Berners-Lee had a lunch with the then Prime Minister Gordon Brown after which Data.Gov.uk was launched. An occasion that can be said as the point where things changed would be that particular lunch.

Incorporating open data

From Mr. Archer’s point of view, open data is not only a big change within the government in UK it is globally a big change in mindset and is difficult for individuals to go through. Most people are not used to having their work done in public. On the other hand in the W3C all working groups are done in public and the conversations that led up to the development of HTML4 is even still available online. It is part of the culture at the W3C but for other people find it too difficult. Mr. Archer says that even he does most of his work in public through public mailing lists. But when outsiders are asked to contribute to public mailing lists instead of through personal emails they always mention that they are worried about spam and other issues. He says that it is a change in attitude to make things publicly available and it is something that needs getting used to. It is also considered a cultural thing because most people do not work this way and neither do companies or governments. People want to keep things to themselves and they are happy to send an email or spreadsheet or something to their boss. But the idea that it will become public complicates things. An interesting view on the excuses that people have for not opening their data is in open data excuses bingo. It can be used in offices or during open data conference and use it when one of the excuses is heard. From all the reasons that people give the ones that are actually valid are very small.

Lawyers want a custom License	It's too complicated	We might want to use it in a paper	People may misinterpret the data
I don't mind, but someone else might	There's no API	There's already a project to...	It's not very interesting

Data Protection	What if we want to sell it later	We will get too many enquiries	We'll get spam
Thieves will use it	Terrorists will use it	It's too big	Poor Quality

Figure 11 Open data excuse bingo[1]

A stumbling block at the moment for the UK is a file that everyone has been asking for since the Guardian campaign in 2006. Mr. Archer says that the Netherlands is very lucky to have access to this particular file. The file under question is a Postcode Address File (PAF) which you need to pay a certain amount to gain access to in the UK.

Open data culture

It is in Mr. Archer's opinion that the culture of opening data is not national or government culture and is simply human nature. He thinks that cultural differences express themselves in different ways. In some countries the culture might be that you must get something right or physical danger is the punishment. In other cultures it comes from the fear of seeming unintelligent which is a natural response as well. If what is made public is 15GB of spending data and it contains many mistakes then the publisher will feel bad about publishing it but it still must be done. Governments around the world are struggling with what the web means in different ways whether it is about opening data or about tracking what people are saying about the government. These two points are actually two sides of the same coin. Publishing data is the government speaking to the public and it is just as much communication as what happens on social media and blogs. Even though people are trying harder and harder to figure out what to do with the internet, the change of culture is the hard part. It is also in Mr. Archer's opinion that this dilemma is purely human nature and not affected by UK culture.

Regulating open data

To the knowledge of Mr. Archer, there is no national legislation for open data. However there he knows a government policy which is called "Digital by Default". The government policy comes from the Cabinet Office which is a part of the government that runs the government. On the other hand, the official document that states the participation of UK in open data is in the Coalition Agreement. In the document it states that local governments are required to make local authority spending about 500 pounds to be made public. Anytime any local authority spends 500 pounds or more must be reported regularly. Mr. Archer was unsure whether it was monthly or quarterly but it was most certainly annually. The response to this requirement widely varied amongst the local governments. An example of a city that was willing to do so was Lichfield. They embraced it enthusiastically and were amongst the first to accept the task. Other councils on the other hand were on the complete other side. They firmly said that they were never going to participate unless David Cameron (the Prime Minister) stands behind them and tells them to do it. The reason for this was all the

[1]Pasted from http://data.dev8d.org/devbingo/bingo.php?n=1&w=4&h=4&title=%22Open+Data+Excuse%22+Bingo&tag=%23openDataExcuses&statements=Terrorists+will+use+it%0D%0AData+Protection%0D%0ALawyers+want+a+custom+License%0D%0APoor+Quality%0D%0AThieves+will+use+it%0D%0AWe%27ll+get+spam%0D%0AIt%27s+not+very+interesting%0D%0AIt%27s+too+complicated%0D%0AThere%27s+no+API%0D%0AWhat+if+we+want+to+sell+it+later%0D%0AI+don%27t+mind%2C+but+someone+else+might%0D%0AIt%27s+too+big%0D%0AThere%27s+already+a+project+to...%0D%0APeople+may+misinterpret+the+data%0D%0AWe+might+want+to+use+it+in+a+paper%0D%0AWe+will+get+too+many+enquiries&rules=%3Cp%3EFor+open+data+teams%3B+print+out+a+copy+and+put+it+on+your+office+wall.+Cross+out+each+excuse+people+give+you.+There+are+no+prizes%2C+but+you+can+tweet+%22bingo!+%23openDataExcuses%22+if+you+think+it+might+make+you+feel+better*.%3C%2Fp%3E%0D%0A%0D%0A%3Cp+style%3D%27font-size%3A80%25%27%3E*+it+won%27t%3C%2Fp%3E

reasons that are also available on the open data bingo. Instead some of these councils would take pictures of screens of data, put that in PDF format and publish it. Lichfield was eager to participate because of the enthusiasm of one person that was the webmaster that is Stuart Harrison. He did a lot of work on the Lichfield council website and did a lot to make the data open. Recently he has moved to work for the Open Data Institute. So what becomes next of the Lichfield open data is unknown. Mr. Archer emphasizes that individuals make a difference when it comes to movements like open data. If there is an enthusiastic individual that really wants to do things then things will happen. In different countries there are different individuals that help the progression of open data.

In terms of licensing of the data, the data published uses the Open Government license. All authorities are encouraged to use that. It is required to use this license unless there is a very good reason not to. Besides that, the users are free to use data that is provided on Data.Gov.uk without registration or fees. However Mr. Archer sees that in the future this might change. Even though it might not be specifically for the data on Data.Gov.uk, but open data in general has to start working with data that is not open thus might charge a certain fee. At this point Mr. Archer provided a report about a workshop that was recently conducted in London. The workshop was conducted between the W3C, Open Data Institute, and the Open Knowledge Foundation. This report gives an outline about the issues that were brought up during the workshop. Some of these issues include the fact that open data needs to have a secure foundation with a solid business case for the publication and use of open data, tabular data formats and the inclusion of metadata as a simple data format, from organizations working with open data and how to provide more guidance, product linked data and also mixing open data. The last topic is one of the important issues; open data needs to be seen as part of a system. Mixing open data with closed data which is data that exists behind a firewall or pay-wall will make the data more attractive. The new PSI directive from the EC allows for charging marginal costs for data which might happen soon. There are other parties that have a legitimate interest in charging for cleaned data. Publishing data that is already available can be done for free but if time is being spent in curating, cleaning and making something from it that costs time and money that people should be paid for. Mr. Archer sees this as a legitimate business, to take open data and clean it and present it in various ways as well as conducting interpretation and visualization. He would not be surprised if some governments charged a little bit in the future.

At the moment there exists no mechanism that checks the quality of data that is published on Data.Gov.uk. This is also partly because there is no registration needed to access the data that is provided on the portal which results in no feedback about the quality of data from the users. This is also a topic that was discussed during the workshop that was previously mentioned. It is in the interest of the developer to let the data publisher know that they are using the data. Data publishers have a legitimate concern about what the point of publishing is and why they should bother publishing. From the workshop several new working groups are being formed that are looking at developing a vocabulary that allows the developer to state on their application that it uses what datasets to make that application. By accomplishing this several things that is to make the application more discoverable but also signals back to the publisher that the data is actually valuable and is used for certain things. Conversely the developer needs to know about the quality of the data; is there a commitment from the data publishers to update the data, how accurate the data is, is it mostly accurate with a few estimates in it, is it only estimates, how is the data recorded and how the data is put together. Although there are many ways of expressing these points, there is not a standardized way of doing it so that is why it needs to be formulated by the working groups. The target is to form an ecosystem between the publishers and the users of data. The publishers can get some idea on whether their data is being used or not and the users can get some way of gaining more trust in the data.

Current policy and impact

It is Mr. Archer's opinion that the current policies that exist for open data works for most. There are so many datasets that are widely used but some of them are completely ignored because they are uninteresting or of no relevance to users. That is part of the problem because from the publishers point of view there is no interest to open the data. The form an opinion that they have tried to publish the data and nobody downloaded it or it was not used at all. As for the policy side of the data, the insistence of publishing data in machine readable format is crucial. Publishing datasets in PDF format is not useful and also the reason why it only gets one star on Tim Berners-Lee Five-Stars of open data. PDF is only considered one-star because there has been something done but it is of little use. However getting the one-star is the hardest one to get because of all the issues that was previously discussed. There is also a new executive order in the US that insists on machine readability which also verifies that the policy to ensure that the data is machine readable is really important.

The government policy in the UK specifically cites Tim Berner-Lee's Five-stars of open data but there are many other reasons that need to be considered as was discussed during the workshop. Most public data at the moment is in CSV format and the W3C needs to make sure that it is acceptable. CSV only gets three-stars for open data because there is a limit to how much information can be extracted from the data. From CSV data there is no way to know what the column and row headings are, what the data types are, and it not necessarily possible to get the annotations. There is now formed a working group from W3C that will work on a metadata standard so that when working with CSV data there will be a pointer in the HTTP to a metadata file that explains how to interpret the data. If the metadata file is correctly added to the CSV you can claim the data to be published as five-star data because you included the data in a package that is similar to RDF. The goal of the working group is to make it easier to create five-star data without expecting government employees to go through more work to make it so. The vocabulary will provide the extra layer of interpretation for the CSV.

For the impact of open data in the UK, Mr. Archer refers to research that is being conducted at the Web Foundation. In the UK there are always efforts being conducted to look into how open data is impacting the public. Tim Davies is a key person that works into identifying these impacts. One project of the Web Foundation that works in this area is the ODDC. The group is focused on exploring the emerging impacts of open data in developing countries.

The interview with Mr. Archer ended with this topic and he showed support for the research on the open data policies. Questions that are relevant at the moment for open data is about what the point of this is and what needs to be done. He also mentions that the Netherlands is the country to conduct such research because the research done is really trying to show the world about this.

Appendix E: Interview with Paul Suijkerbuijk (Netherlands)

Paul Suijkerbuijk is project manager for open data in the Netherlands, for the Dutch ministry. There currently exists an open data portal, Data.Overheid.nl whose main goal is to have as much data of the Dutch government available on this website as possible, by providing links to their sources. At the moment there exists a Freedom of Information Act which obligates the government to respond to requests for data by citizens without regulating the pro-activeness from the government to open data. There are some guidelines available in the policy guidelines for open data and the government is currently working on a new Freedom of Information law concerning this, which is not an easy process and will take some time. Under Neelie Kroes' guidelines, that are available on the PSI Directive, there are obligations to have the data ready in

machine readable data and exceptions for organizations which ask money for the data they own, but it does not obligate the opening of data itself. Many of the initiatives within the government with regard to open data are conducted by enthusiasts, people believing it will work or people seeing benefits from opening data for their own goals. The start of open data itself in the Netherlands can be said to be brought by Ton Zijlstra from Hack de Overheid. This was one of the organizations that brought open data to the Netherlands. In fact the real beginning was when Obama opened up US data and it made more people pick it up.

Opening data from inside the government organizations

The last two years Paul Suijkerbuijk has been talking to different governmental organizations about opening data but it could be moving faster. Open data or data at all is not appointed to specific authorities within governments. It is very individually arranged how data is being managed within organizations. At times it is managed by the ICT department, other times it is under the communication department, or the policy makers. It is not easy to go to all the cities in the Netherlands and find a specific contact person to talk about open data with. It is very diverse between different organizations. This makes the process of opening data less easy because it is unclear who to talk with to get the organization to open their data. In the beginning of the process they got about 500-600 datasets opened on Data.Overheid.nl.

Then they started thinking about how to speed up the process of opening up data because this rate was not fast enough. Not enough to build a robust group of re-users of data, because there is not enough data to actually do something with. Attention was drawn to the INSPIRE program of the European Union which states that we have to share all kinds of geo data with each other. Here there was already a lot of standardization and licensing arranged within INSPIRE. On the other hand, within the government we could look at the public domain data and the creative commons data and open it very easily, because it was already available under the right circumstances. Integration was made between the national geo register (where this data, and all kinds of geo data was registered, open and closed data all together) and they took all the existing references to open data and brought these to Data.Overheid.nl. This made the number of datasets available for reuse increase to about 5000. We do have a lot of datasets but the change within the minds of the civil servants is not there yet. Civil servants should see that opening data could be used to serve a better purpose and ease the process to achieve their various end goals.

Various municipalities across the Netherlands are advised to publish references to their data on Data.Overheid.nl. In fact Data.Overheid.nl does not contain any data at all and are just focused on linking it to the portals that do contain data. It is done according to this model because the goal is to have the data as up-to-date as possible which is harder to do when the data is uploaded to a different site like data.overheid.nl. It is easier to forget to update data that is not stored within their own databases. The data should be located as close as possible to the owners of data. More municipalities are investigating the possibilities of opening up local data. Big cities are quite advanced with open data and have a lot of initiatives to stimulate the economy using the data. Utrecht just recently announced that they will be joining in, making their local data also available. Other cities, on the other hand, are just beginning to understand what open data actually is. Contact with the municipalities is very diverse but there is a process involved with meeting together to work on the various aspects of opening data. There are many efforts to develop the Freedom of Information law and to maintain contact with the municipalities.

Changing the mind-set of civil servants

The current mind-set is not being against the opening of data but more that they focus a lot on the problems with opening their data. It is needed for the process of opening data to be integrated in the everyday tasks of the civil servants. When this is achieved there will be less need for regulation. Civil servants are interested

in opening up their data but; quality of data, actuality of data (timeliness), security, privacy issues….etc. For instance the data currently being opened is a couple of years old because the process of opening data takes so long. This means that it might not be relevant anymore. The sentiment of most civil servants about opening data is positive, however, the problems they see are big, in their view. This is what keeping them from doing so. A lot of the problems that civil servant see are bigger in their minds than they are in reality.

All the reasons for the lack of participation to open data are mostly because of non-technological issues. Such as that it isn't their task, too much hassle, and too many risks in doing it etc. And this frame of mind is what needs to be changed. Even though open data might be outside the task of a civil servant it should be formulated as a way in which they can achieve their tasks, faster, better. With the uncertainty of what the benefits of open data are, the stimulus for opening data is even lower. The value of open data has not been and cannot be estimated. A critical mass is needed to ensure that the open data policy works. And how this critical mass is affected and achieved is what needs to be investigated.

Paul Suijkerbuijk changed his approach: getting out of the mind-set of thinking within the organization by going to talk to citizens, inhabitants of the Netherlands talk to them about societal issues. What is happening in their everyday situation and what kind of problems are they experiencing and seeing if these problems can be solved by using open data. Because there were a lot of issues he decided to narrow the focus to Groningen; the depopulation, and poverty districts, 'achterstandswijken'. He is arranging discussions in Groningen and the poverty districts and getting to the specific societal issues.

They take these issues and the data they deem fit to solve these issue and bring this combination to "solver meetings". At these solver meetings there are app developers, concept developers, artists hoping that they bring a creative way for solving these issues to the meetings. The first meeting will be on 20th June 2013. At the solver meetings the attendees might give notice to other types of information that they want to use to solve that problem. There will be a process of demand-supply of other data. Paul is focusing on having five real working solutions at the end of this year, which can give proof that opening your data is working for a government. Developing actual apps that solve actual problems that were made possible by the fact that these governments have opened their data can be used to show other governments the 'real world' benefits of opening data. The end goal is to get governments to open data and society will play their role in solving the issues that exist within these governments. Paul is convinced that this project will be the kick-start for other government to open their data. Setting an example of why they should open data. The ultimate end goal is government opening up all their data. But the reuse of this data also needs to be stimulated, which this approach can also do.

This is very closely related to the open government. This difference is that the open government initiative is more leaning towards a big system change within the government; government should think in a different way. This is also good, but the way Paul is looking at it right now is to start small, with the 'small' societal issues that are there and getting them resolved. And this will evolve in to solutions on a bigger scale which fits into the open government initiative. These two will come together in the longer run.

Standardization, metadata, and quality

Standardization of datasets would make it less easy for the government to open data. Raw data is what is being pushed as the way to motivate the publication of data. This encourages user feedback that leads to revision and improvement of the data. The focus is to publish, worrying about metadata, quality and other attributes later based on the user feedback. At the moment the standardization efforts are done at a low level, although it would make the data more reusable. Pushing for linkable data makes it less easy to get the

data opened. The focus is more towards the context of the data being stated because if the context of the data is not given misuse/abuse of the data is very easily done, maybe not on purpose. The meaning might be misinterpreted. There is the need for a good context description to enable people to reuse it in the best possible way. This also included the gathering methods because this gives a view of the quality of the data. First there needs to be this big change that is what open data is about at this point in time.

License

The website talks about licenses however these are no formal policies at the moment because these issues are still under investigation at the moment. The intention is that their open data is under the "creative commons zero", open standard format etc. However these decisions have to go through the policy process. The websites addresses the preference for the CC0 license. This again is an issue that depends on the diversity of the data and the involved areas of government. On certain areas of data it is already decided to adapt the CC0 license and on others it is not yet decided. Creative commons zero is what they see as being open data; however there is the "creative commons by…" which states that you have to mention the source/supplier of the data. This could be seen as a barrier to its openness.

Registration

There is (will be) the option for registration however no registration is obligated to download the data. Because of this there is no insight into the reuse which there is also no stimulus for the data supplier to open their data, as they do not see it being used. It is suggested for the owners of the data to contact the users of the data to monitor the re-use. However with the lack of the registration makes it more difficult. It is clear that this creates a barrier to those that open the data because there is no clear evidence of what is being done with their open data. Government organizations wonder about this, the only thing that is recorded is how many times a page in the system is seen. No further actions are being taken because they want to keep it that way to again emphasize that is completely open without barriers. However they are looking at a way in which the users of the data have the option to register and if they want publishers can contact them to provide updates or newer versions of the data that they have used.

Costs of opening data

Costs are dependent on the IT situation within an organization. Open data can be cheap if the system/infrastructure already supports it. Alkmaar just added a page in their content and published some CSV files which were just two days work. License plate data needs a heavier database which is more complex and more difficult.

Open data initiatives

They made a list of 150 initiatives being made based on open data. A few are successful but most of them are not. The reason for their demise is that they do not make profit or are currently non-existent and everything in between. This pushes the need to more actively create applications and solutions where you can see the benefits. Such as the previously mentioned community outreach programs. There are classical examples such as a publishing company that compiles information about working for the government, from the government, and they publish the book and sell it to the government. It has been doing so for many years now and is profitable with a good business model. They just make it easier instead of having to go through each of the government pages individually. For more of the application developments, 10000 schools etc. are good examples. Companies build open data apps without profit but use it to show customers what they can do and get more work by showing off their open data skills.

Appendix F: Interview with Rushda Majeed (Kenya)

The interview conducted with Rushda Majeed was done via email correspondence through a question sheet and also with further explanations through email. Rushda Majeed is a researcher at The Princeton University as a Senior Research Specialist at the Innovations for Successful Societies. Her relation with Kenya came through the publishing of several case studies that she conducted on the open data initiative in Kenya. Her published work titled "Disseminating the Power of Information Kenya Open Data Initiative, 2011-2012" (Majeed, 2012) was used as an important literature reference when conducting desk research about Kenya's open data. Table 10 contains the questions and answer sheet that was sent via email to Rushda. Through her answers she refers to Kaburo Kobia as a contact person from the Kenya ICT Board. Contact with the Kenya ICT Board was pursued but through a different source which is also presented in this report.

Question	Answer
In Kenya what is the main goal or strategy of open data?	The key people behind the Kenya Open Data Initiative wanted to put government data online for economic and social benefits to society. M-PESA, the mobile money transfer program, had taken off in 2007. Bitange Ndemo, permanent secretary at the Ministry of Information and Communication, made the case to President Kibaki that more government data online would fuel application development (along the lines of M-PESA) and create more jobs. In addition, data, especially on health and education, would enable his government to deliver better services.
How did the open government movement start in Kenya?	I would say that the push for providing more government information to citizens started with advocacy for the Freedom to Information Act in 2000, still a work in progress. But the clamor for open government came towards the end of the decade. The country's new 2010 constitution also provided a legal framework for making government information public. (Others based in Kenya may be able to offer specific years / dates.)
How is the open data publishing and re-use situation in Kenya as of now?	From what I understand, limited datasets went on the opendata.go.ke website in July 2010. The taskforce behind the launch added more by the end of the year. The most recent data that I put in the case study on the Kenya Open Data Initiative can be found at: http://www.princeton.edu/successfulsocieties/content/data/policy_note/PN_id206/Policy_Note_ID206.pdf. You may request the most recent data on number of datasets and usage from Kaburo Kobia at the Kenya ICT Board.

Question	Answer
	My sense is that the publishing and re-use has been disappointing. Interviewees in 2012 noted that (1) the government was not putting enough data online, and (2) Not enough developers were using the datasets on the site to create useful apps.
How do citizens, government officials, and politically involved people react to open data?	Each sector varies. Broadly, government officials and ministries seem to be skeptical although some, like health and education, have shared their data with the opendata.go. ke site. Civil society organizations have given a mixed response. Many would have liked to be included in the initial taskforce and feel disengaged.
How does the everyday culture in Kenya (in general) support the open data process?	Hard to say. It seems that M-PESA has set a good example. The vibrant Nairobi technology community has also pushed interesting and innovative open data projects (iHub, Ushahidi). Still, many citizens do not know about "open data" as it doesn't impact their lives directly.
How does the ICT infrastructure support this process?	The Kenya ICT Board built a brand new website for the Kenya Open Data Project. The fiber optic cable that came to the country in 2007 helped.
Why is it beneficial for the country at this point?	The open data advocates made an economic and financial case for opening up information. It would help grow the economy further, they said.
What legislation is formulated for the opening of data?	The Kenya Open Data movement relied on the 2010 constitution for making a case for putting government information online. Ndemo also forged an agreement with the Open Government Partnership to push for the portal. (http://www.opengovpartnership.org/). The Freedom of Information Act, which would have been the right legal framework, is stuck in parliament.
What are the current policies that help regulate the opening of data?	You may want to ask Kaburo Kobia at the ICT Board. The Board may have revised guidelines.
What areas do these policies regulate?	See above.

Question	Answer
How is the licensing or restrictions that are available on the process of opening data?	At the time of the launch, the terms of use stipulated that the public was free to use and reuse the data as long as it was not for commercial use. The opendata.go.ke website will have the most recent terms of use.
In which format does the data need to be published? Why these ones?	https://opendata.go.ke/browse.
What forms of pre-processing of the data are required?	You may want to check the most recent pre-processing decisions with Kaburo Kobia at the ICT Board.
What form of contribution can the citizens participate with?	The site has a "Suggest a Dataset" section to request specific information.
What are the estimated costs of opening data? What do these costs cover?	Kaburo Kobia at the ICT Board may be able to provide the most recent information. The Kenya Open Data case study has numbers on the launch of the site: http://www.princeton.edu/successfulsocieties/content/data/policy_note/PN_id206/Policy_Note_ID206.pdf.
How is the data quality maintained?	Kaburo Kobia at the ICT Board may be able to provide the most recent information.
Have you ever heard of metadata? If yes, is this something that is considered when talking of opening data?	I didn't hear it in Kenya during my research interviews.
How does the government support the open data process?	The Ministry of Communication and Information, the Kenya ICT Board, and the Kenya National Statistic Bureau are the three agencies that coordinate to open government data. The ICT Board has recently funded or helped coordinate hackathons and other similar events to encourage people to use the data on the site. Other initiatives like Code4Kenya build on the data on the portal.
How is the impact of opening data checked or measured?	Kaburo Kobia at the ICT Board may be able to provide the most recent information. At the time I conducted the research in 2012, no good measure existed.

Question	Answer
In your opinion are these policies enough? What is being done to improve or expand them?	In 2011, the usage of data on the site was low. Interviewees complained that the private sector had not taken advantage of the datasets to build applications. Kaburo Kobia at the ICT Board may be able to provide the most recent information.
What are the current challenges in opening data?	Kaburo Kobia at the ICT Board may be able to provide the most recent information. (I conducted my research in 2012.)
Are there any future improvements in store for this country in terms of open data or open government?	The Open Government Partnership is one avenue to push for putting more open data online in Kenya. Kaburo Kobia at the ICT Board may be able to provide the most recent information. (I conducted my research in 2012.)

Table 10 Interview question and answer with Rushda Majeed

Appendix G: Interview with Jay Bhalla (Kenya)

Jay Bhalla is an innovation strategist that helped pioneer Kenya's digital revolution. He helped develop the Kenyan government's ICT policy and also played a role in developing the Kenya Open Data Initiative. He is co-founder and current executive director for the Open Institute (OI) think tank. The interview was conducted in two parts through question and answer via email and also through a Skype call for follow up questions. The following Table 11 contains the questions and answers that were exchanged with the interview source. Below the table the summary to the follow up interview that was conducted is also presented.

Question	Answer
In Kenya what is the main goal or strategy of open data?	The main Goal of Open Data in Kenya was to empower an innovation and knowledge economy and to create jobs for the youth. We had the infrastructure in place and a budding software development community that was in place that wanted access to Data to build relevant applications. It was also to allow Government itself to run more efficiently.
How did the open government movement start in Kenya?	We had a Champion in Government in the former PS Bitange Ndemo who had a vision for Open Data in the Country. Together with stakeholders led by the World Bank and the Kenya Bureau of Statistics a task force was established to plan the roll-out and the launch.
How is the open data publishing and re-use situation in Kenya as of now?	Unfortunately there are no laws that are in place for Data release and therefore since the launch there hasn't been much new data released. In addition to this there are no structures in place that allow for dedicated staff to maintain the Open Data Platform and/or be the resource people for the Open Data Platform. Ideally there should be an Open Data Manager, Community Manager and Technical Manager. Some applications have been used and the traffic to the site has been constant through the year. Open Data has spawned other initiatives such as

Question	Answer
	Code4Kenya that embedded fellows or "Change Agents" in these Organizations to change their mindset as to the value of Data not just on the Kenya Open Data Platform but also their own Data.
How does the ICT infrastructure support this process?	Very Important. In Kenya we have the advantage of good fiber Bandwidth, a huge Mobile Community, Co-working spaces and accelerators as well as a growing pool of Software Developers.
What is the source of funding?	The Source of Funding for the Kenya Open Data Platform came from the World Bank through an ICT Infrastructure Grant.
Why is it beneficial for the country at this point?	If implemented well it is beneficial from a number of fronts. Allowing for Government Transparency, Allowing Government to use the data for better planning purposes, Empowering Citizens to make better decisions about their lives and creating an Innovation and Knowledge Economy.
What legislation is formulated for the opening of data?	At the moment there are only proposed legislation to open data and a Freedom of Information Law that is yet to be passed. Data re-use on the Platform is allowed for commercial and non-commercial use with attribution to the source. Going forward it is envisioned that other legislation including tying regular quality data release from Ministries to performance factor as one of the key metrics for Government Performance monitoring.
What are the current policies that help regulate the opening of data?	None in Place at the moment
What areas do these policies regulate?	This question is not applicable as yet however as stated above if put in place then this will allow for constant release of quality datasets from Ministries and better management of datasets released..
How effective have the policies been in ensuring the opening of data?	N/A
How is the licensing or restrictions that are available on the process of opening data?	The licensing is based on the UK open License and is straightforward as to the use and re-use of the Data.
In which format does the data need to be published? Why these ones?	It is generally advised for the data to be published in raw format with the general format being .csv, .xls, .shp and .json for mapping files. The reason for this is that it is easier for data to be used, manipulated, investigated and re-used as compared to being in closed formats such as .pdf where the data would have to be "scraped" first to create the dataset to be more useful.
What forms of pre-processing of the data are required?	Depending of the Dataset again it would be envisioned for the dataset to be in a compatible format first, where the dataset is converted and then it is cleaned. By cleaned we mean that for e.g. if the data is in a spreadsheet, the rows and columns are arranged accordingly with the correct data.
What form of contribution can the citizens participate with?	They can request for Datasets and they can engage with intermediaries and infomediaries to not only ask for data

Question	Answer
	release but ask for assistance in understanding the data that is already released. By doing this they are also providing vital feedback to the data that they require and problem statements that the data through technology through applications can fulfill.
What are the estimated costs of opening data? What do these costs cover?	Hard to give a finite figure on this but I would envision that the costs would cover putting in place dedicated staff to maintain the platform as mentioned above in addition to the ICT infrastructure for the release of data from the various line ministries and the Bureau of Statistics.
How is the data quality maintained?	Again this is dependent on the staff that is dedicated to maintaining the platform and data collection. There would be a need for a dedicated training program in place that would address the quality of data across the Ministries.
Have you ever heard of metadata? If yes, is this something that is considered when talking of opening data?	Yes Metadata is very important when talking about Open Data as it is useful to create order within datasets, create, Linked open data, improve finding and preservation of Linked Open Data, improve the visualization of linked open data, avoid duplication of datasets, among others.
How does the government support the open data process?	For the moment in the contest of the Kenya Open Data Platform the Government has prepared legislation that is awaiting parliamentary approval. A lot can still be done without this like hiring the correct staff to maintain the current infrastructure.
How is the impact of opening data checked or measured?	Unfortunately at the moment it is checked outside Government by external initiatives and entities like Civil Society and other Organizations in the Open Data Space and the mechanisms used range from actual interviews with citizens on the ground as to the awareness to seeing the metrics around the application use that have been built around this data.
In your opinion are these policies enough? What is being done to improve or expand them?	No they aren't. At the moment other than the bills that are awaiting Parliamentary approval not much. There needs to be policy in place that makes it law for Ministries to release data on a regular basis. There should also be consultation with other stakeholders and gate-keepers of data such as Media to work towards releasing their data.
What are the current challenges in opening data?	Lack of by-in from other Line Ministries to release data, lack of mechanisms in place to maintain the current platform, lack of coordination between Government and other data gatekeepers and the lack of incentives to use the data. In addition to this there is a lack of systems in place for educating the academic community in the value of training in handling data.
Are there any future improvements in store for this country in terms of open data or open government?	Again it is envisioned that once the legislation is passed there will be better release of data that will lead to better uses of the data and that will translate to advantages

Question	Answer
	such as greater transparency more applications being built and greater data use by citizens.

Table 11 Interview question and answer with Jay Bhalla

The following summary details the rest of the interview that was conducted after a review of the question and answers that were already given. In the beginning of the interview Jay referred to a new initiative that is called "What's the Plan" which aims to compare the action plans of the countries involved in the Open Government Partnership. This initiative is mentioned because it carries a similarity to the research in this report. The initiative is creating a peer review mechanism for governments and comparing action plans from the OGP. There is much talk about all the different action plans which are not in harmony.

He does confirm the literature finding that Kenya does not have an open data policy at the moment. There has also been no draft submitted. The policy is still to be designed and implement. The challenge is currently there are many changes within the government. There exist counties in exchange for the provinces that previously existed under the new constitution. Kenya never used to have a default system of government. There are changes that needs a certain degree of adaptation before there can be more focus on open data policies. Another change is the fact that Kenya no longer has ministries which are now cabinet secretaries. There is also a reduction in ministries because of the coalition from 42 to 18 ministries. Before there were various ICT bodies such as the ICT board (had the vision to make Kenya the ICT hub and open data fell under that), the directorate of e-government that rested in the office of the President, and another body. All these three bodies are merged into one. So there is no policy in place in terms of data release, where the initiative should rest (with an ICT agency or the bureau of statistics), and no dedicated staff yet that have the responsibilities relevant to the initiative because of these changes. Kenya has a taskforce but not really active because of the transitions in government. The Kenya action plan for the OGP is also not very comprehensive. It touches data release but there is no policy that is in place. There is a FOIA that is going through parliament but it has not yet been formalized.

With all the changes in the government the priority of open government and open data within the government is the next issue that Jay was asked. The first OGP Africa meeting was conducted about 3 weeks ago, and the cabinet secretary himself was there. It is fortunate that the cabinet secretary comes from a governance and transparency background. So his commitment is to see Kenya taking the full front in terms of how open government is concerned. However his thoughts were that the infrastructure and the right systems need to be in place first. I agree with it because it is not possible to start something and it ends being changed because its starting all over again. Open government and open data is high up on the agenda. But higher than that is the systems, infrastructure, and policies in place first before you can actually do some implementation. There's a lot of foundation to be laid but also while that foundation is being laid there needs to be more movement as far as legislation is concerned. Jay says that as much foundation as possible may be laid but without legislation then nothing will move. There needs to be legislation that requires, as a prerequisite, for ministries and government agencies to release data. That needs to be tied with performance contracts and things like that.

The former PS, which was one of the leaders in developing the Kenya Open Data portal, is no longer in the government which is also another worry about the future of open data. He did however raise the initial changes that were needed. After his efforts, the reaction to open data varied. The push for open data in Kenya was not based on governance or transparency or at least that was not the angle that was used within the government. The infrastructure was in place, a growing mobile culture, fiber optic cables, co-working

spaces, accelerators, and a youthful population that was a large pool of IT talent that provided the aspects that supported the open data movement. The IT talent that existed was calling for useful data that they could use to build applications because they did not yet have access to data. The PS' commitment to them was that I'll give you data so we can build this knowledge innovation economy. This was also his speech to the powers if you release this data you'll be able to do a number of things: 1) the youth that are looking for jobs will be able to use this data build applications, 2) government will be able to run more efficiently because they'll have a more structured data because they'll know where there spending is going know how to spend wisely, and 3) governance and transparency. He had huge support from civil society who was clamoring for release of data but they were more governance and transparency angle. Then there was huge support from the developer community. In government itself it was a mixed reaction. As further answer to this topic Jay suggested to look at the Kenya open data case study written by the World Bank that will give a lot of insight.

Another issue that was questioned was about the guidelines that are on the portal at the moment, whether they were to be included as the formulized policy or is that policy only for the release of data. In Jay's opinion it needs to be a combination of both. There needs to be a policy on release of data and a policy on the format of how the data should be released. This is a process that needs to be developed in tandem. A problem that arises here is that there has not been much consultation. To design policy there needs to be consultation from multiple stakeholders and multiple segments. These include from civil society, academia, technology community, private sector, and various government agencies. It has not been done. It was very important for the OGP action plan. It was an opportunity to get it right. Unfortunately the action plan was designed without using the opportunity to consult.

Jay also says that open data itself is still a new concept in Kenya until now and not just within the government. Even now the government is still wary. They think it means secrets and wiki-leaks. 48 hours before the launch of the platform the PS was called in by the President because there was a meeting with the head of public service and security to ask what kind of data that is actually being released. They thought it was data that contained government secrets. But this wariness and unawareness of what open data means is not only in government but in other segments as well. It is one of the reasons Code4Kenya was started. Open data was trying to be embedded not in government this time but in media and civil society. It was to change the mindset of an organization towards what the value of data is. It aimed to not only see the value of data on the government portal but what is actually owned by these organizations themselves. These organizations are also considered to be gatekeepers of data. Media for example sits on huge amounts of data. So the Code4Kenya initiative was aimed to changing the mindsets of organizations towards what they can gain from data, what they can do with data, and what data would mean to them. It enabled getting a greater understanding from these organizations and more participation in the entire process. At the moment it is something that we're doing but we hope that when the systems are in place government will be able to assist. Assistance from the government comes from having the right staff which includes an open data manager or technical manager and a community manager or outreach manager.

As of this moment there is not as much audience from citizens from Jay's opinion. It is very difficult for the average citizen to understand the data that is up on the portal. Secondly the average citizen does not care about that data. What he cares about are services, how my tax dollars are being used. In Jay's observation, the data needs to be disseminated to the citizens in a way they can understand. This means showing the data through applications, intermediaries, and community radio or media. On the portal Kenya has community development funded project. Every constituency will get money and they would go to various

projects such as building schools etc. There is a study about whether rural folk know about open data that is available for access on the Open Institute blog. The study is to approach people from across age groups and different walks of life in a rural area and we show them this data. Some of these projects were marked as completed but the project hasn't even started. Apparently the money had been misappropriated. The data for these projects was available and if they had a way to channel the data through media then they could respond. This is how open data can make meaningful impact in their life. There needs to be more structure and more dialogues with intermediaries so that the data can actually be taken and changed into a way the citizens can access and understand it easily. With Kenya, we had to run before we could walk is how Jay explained the start of open data in Kenya. The President had said to release this data and that open data is approved. To release this data we had a choice on working on structures first or just release it because people may change their minds tomorrow. Now the next steps are working on these structures.

Appendix H: Interview with Tina George (Indonesia)

Tina George is a public sector specialist with the World Bank, based in Jakarta, Indonesia. Her scope of work includes being counterpart to the President's Delivery Unit for Monitoring & Oversight (UKP4), supervising a specific investment lending project with the National Statistics Agency (BPS), and leading initiatives in technology-mediated transparency and social accountability. UKP4 is a working unit that is charged with the leadership of Indonesia's Open Government Program (OGP) together with several other ministries and governmental agencies. The BPS reforms project, called Statcap-Cerdas includes a fundamental transformation of all statistical processes, methods, IT systems and aligning these with e-government. This area is closely related with the opening of data. The 1997 Law on National Statistics along with a number of accompanying decrees equips BPS with far-reaching competencies and tasks for providing statistical data to the government and the public, to assist statistics divisions of government ministries and agencies in developing their own systems and to establish and promote standards on statistical techniques and methods. The legal framework regulating the access to government information is robust. The 2008 Law on Freedom of Information applying to all government institutions at the national, regional and municipal level, as well as political parties and non-governmental organizations regulates the handling of public information and prevents government ministries and agencies from withholding information that is in the public interest. As for the projects related to transparency and accountability, it is whether transparency should be an end to itself or if transparency for the sake of accountability should be the end.

The interview was conducted through VoIP calls using the Skype program which was recorded to be kept as reference. The length of the interview was approximately 40 minutes and followed an interview plan that consisted of three question categories. The categories included general opinions about open data in Indonesia from a national perspective, the policies and laws that are involved in the opening of data in Indonesia, and the open data process and impact. The goal of the interview was to gather more information for the research from the practical side of open data from people that are directly involved in the processes.

The first questions were more related to a high level view of open data in Indonesia in relation to the goals, strategies and drivers. Tina reflected that it was difficult to pin point the goal or strategy of open data in Indonesia because there are a lot of action plans that have been published in relation to it. But it could be said that the principles are transparency, innovation, and participation. However because openness is implemented at many different levels of the government in various contexts, it is not immediately visible whether these principles trickle through each of the levels. When questioned about the involvement of open data in the national effort for bureaucracy reform, it was Tina's opinion that the bureaucracy reform plan covers many things which also include transparency and openness. But a link between open data and

bureaucracy reform is yet to be made. In her experience the bureaucracy reform involves a lot of documentation and paperwork on processes. Based on this documentation, the officials from agencies that have completed their assessments will receive a bureaucracy reform allowance.

Tina identified that the start of open data can be distinguished by a few different events. One is the FoI law that is said to have led to openness on various levels. But can also be generally looked at from the post-Soeharto era where there was a lot of decentralization in the district levels. At these levels innovation started to arise. Some of the district governments could provide one-stop online services to the public which the central government has yet to do. For example, the municipality of Solok, West Sumatra provided a one-stop-shop for various services to the public and Sragen, Central Java established a citizen friendly e-government system of public services. So first is the FoI law which makes it easier to have people publishing the data and the other is the fact that Indonesia is on the path to reform, democratization, and participation. The voice of citizens can be considered as the drivers that will lead to more openness. On the other hand, even though there is the FoI law, the ICT infrastructure in Indonesia is still a barrier In the BPS project where Tina is involved, there is difficulty in drawing agreement on the design of an integrated system to process statistical data. The Ministry of Communication and Information Technology is also not at that place yet where it is able to facilitate a national open data portal such as Data.Gov (US national open data portal). It is not clear which agency will step up and lead the Indonesia Data.Gov portal, although there is potential for UKP4 to lead a taskforce of key agencies such as RISET, BPS, MoF and Menkominfo for this purposeAnother remark is on the source of funding for the opening of data in Indonesia. Tina was not certain whether or not there was a specific budget allocated to the open data movement. She knows that the UKP4 has a budget that can provide some resources for agencies that opt in openness. If they produce an action plan, they can approach the UKP4 and request for the resources to implement it. Tina also reflected that this was an interesting point on whether or not it should be included in the national budgeting.

The next group of questions was more related to the policies that are involved in the opening of data in Indonesia. As was previously mentioned, the FoI is considered to be the only policy that regulates open data in general according to Tina. A policy that encourages government agencies on national data sharing and accessibility along the lines of India's National Open Data Sharing & Accessibility policy is yet to be developed. The FoI is a law and not an open data policy such as those developed by UK, the US or India. There is work being done with the BPS to develop metadata for statistical datasets based on the Data Documentation Initiative (DDI) which relates to having clear metadata and listing of all the available data. A public listing of datasets would benefit from a detailed metadata standard. However at the moment there is no emphasis on a certain standard for publishing the data, pre-processing of the data before publication, or ensuring interoperability of the datasets. In terms of machine-readable (technical open data) there is still some ground to cover. For example, most data that is published through Satulayanan.Net is in PDF format. It is hoped that the next versions will be published in a machine readable format. As for participation from the citizens in this effort, Tina highlights Lapor as good example of citizen participation that was initiated by the UKP4. Lapor was one of the early systems that UKP4 released and has been covered in reports by McKinsey and a Princeton University. Lapor facilitates citizen reporting through an online application which can be about any public service. It was also emphasized by Tina that the UKP4 is one of the leading agencies in the open government movement in Indonesia.

Through her experience, Tina finds that there are still many internal bureaucratic barriers when it comes to getting support for openness. A lot still needs to be done to improve the understanding of what the benefits of an open government environment can be for government and governance. Also there needs to be more

evidence of how accountability can lead to transparency. In terms of outreach from the government, a lot still remains to be done. However Tina does not agree with the statement that the government does not feel the need to participate in the open government movement. Tina says that Indonesia is still in the development phase and has taken the opportunity in being a founder of the Open Government Partnership which brings strategic advantage on the world stage. Being a complex country system with wide geographic spread, it is difficult to completely open everything straight away and a lot more has to be developed from the central and local government, the ministerial level and also from the various agencies. There are many different reasons why the uptake has been difficult. The current government culture may also play a role in this difficulty. Government officials are not used to putting government data on websites and on other public domains. Tina also mentions that according to literature, a country with greater taxation which finances the government has more attention to the release of public data. Because the government is financed by taxes, it raises the question of to whom does the information belong to whether it is the taxpayers that financed the collection of data or to the data producers. So there are views that believe that in countries where taxation is lower it is harder to argue for transparency. Citizens have less of an interest to demand for that information and also the government feels less need to publish it. However, newer literature focusing on economies of windfall revenue (from oil and natural resources) is seeking to disprove this thesis, and show that citizens from such economies are also concerned about transparency and openness of government data.

As for the future outlook of open data in Indonesia, Tina sees that there is a greater demand side for participation. Quite a few projects are coming from the demand side of citizens being more engaged or agencies that are trying to engage the citizens more. Indonesia can be seen as a country that consists of several small communities. On occasion these communities work together in small community development programs. When these communities begin to raise questions to the government then open data will bring many benefits because the communities will make good use of it. However at the moment the direct impact of opening data this far has not been measured or checked yet. Especially since Indonesia is still in the very early stages, there is not much to be measure yet. In the coming year there is supposedly much activity to be done in the area of open government and open data in Indonesia. Indonesia will be taking lead chair role in the Open Government Partnership which causes a lot of outreach to be done. One of the projects financed by the World Bank involves civil society organizations for social accountability. This involves working with local government budget datasets that are already open and follow the money style applications to to inform citizens about gaps between budget allocation and expenditure improve public services. There is also some uncertainty on what is in store for Indonesia because of the upcoming election year in 2014. Institutional changes such as this will also take into account the political climate at the time. Tina is confident that Indonesia will develop towards the goal of being an open government over the long run.