An investigation of open government data utilization and impact: influential factors and stakeholders

Xiaoyi Yerden

University at Albany, State University of New York, amyzhao724@gmail.com

The University at Albany community has made this article openly available. Please share how this access benefits you.

Follow this and additional works at: https://scholarsarchive.library.albany.edu/legacy-etd

Part of the Library and Information Science Commons

Recommended Citation
https://scholarsarchive.library.albany.edu/legacy-etd/2839

This Dissertation is brought to you for free and open access by the The Graduate School at Scholars Archive. It has been accepted for inclusion in Legacy Theses & Dissertations (2009 - 2024) by an authorized administrator of Scholars Archive.
Please see Terms of Use. For more information, please contact scholarsarchive@albany.edu.
AN INVESTIGATION OF OPEN GOVERNMENT DATA UTILIZATION AND IMPACT:
INFLUENTIAL FACTORS AND STAKEHOLDERS

By
Xiaoyi Yerden

A Dissertation Submitted to the University at Albany,
State University of New York
In Partial Fulfillment of
the Requirements for the Degree of
Doctor of Philosophy

College of Emergency Preparedness, Homeland Security, and Cybersecurity
Fall 2021
ABSTRACT

Since 2009, increasing numbers of countries have started their open data development. Through these efforts, they aim to improve government transparency, advocate innovation in public services and business, boost economic development and improve quality of life. In order to achieve these values of Open Government Data (OGD), just making the data open is not enough; OGD has to be actually used by people. Among previous research, there is a limited discussion on the actual utilization of OGD and the generation of OGD impact. Focused on OGD utilization and impact generation, this dissertation is designed 1) to identify research status, gaps and directions for future studies through a comprehensive literature review; 2) to analyze the relationship between government related factors, OGD related factors, user related factors, community related factors and the generation of different types of OGD impact among different countries over time; and 3) to investigate the roles of stakeholders and identify their perceptions of the facilitators and obstacles in the process of OGD utilization through a case study focusing on the New York State Geographic Information System (NYS GIS) Web Services. This dissertation provides a better understanding about the influential factors and the interaction among stakeholders in the process of OGD utilization and impact generation. Practical implications are also discussed to better guide OGD practice in the future.
ACKNOWLEDGMENTS

I am thankful for all the help, support, and guidance that my committee, Dr. Luis F. Luna-Reyes (Chair), Dr. Teresa M. Harrison, and Dr. J. Ramon Gil-Garcia have provided during the dissertation process. I also appreciate all the interviewees from different fields that participated in my case study and the valuable feedback they have given to me.

I would also like to express my sincere gratitude to my father Quanqing Zhao and my mother Qiuhong Ma back in China for their unconditional love and generous support that make it possible for me to study abroad and advance in my academic learning and research. The weekly video calls with them have given me tremendous comfort and encouragement to overcome the difficulties I have encountered at different stages of my study and research. I would also like to thank my family and friends here in the United States, who have helped me conquer different challenges in the past nine years.

Last, I would like to thank my husband Austin L. Yerden, who is always there supporting me, loving me, and helping me improve myself day by day. This dissertation is dedicated to him and my daughter Yulan Yerden.
TABLE OF CONTENTS

CHAPTER ONE ........................................................................................................................................1
1. INTRODUCTION ..................................................................................................................................1
   1.1 Background and Related Work ..........................................................................................................2
   1.2 Problem Statement .............................................................................................................................3
   1.3 Structure of the Dissertation ..............................................................................................................5
CHAPTER TWO ..........................................................................................................................................6
2 LITERATURE REVIEW ...........................................................................................................................6
   2.1 Research Collection ..........................................................................................................................6
   2.2 Select Criteria ...................................................................................................................................7
   2.3 Descriptive Analysis ..........................................................................................................................8
      2.3.1 Time and area of publications ......................................................................................................8
      2.3.2 Research methods .....................................................................................................................10
   2.4 Thematic Analysis ............................................................................................................................12
      2.4.1 Theoretical frameworks ..........................................................................................................12
      2.4.2 OGD utilization .......................................................................................................................17
      2.4.3 Stakeholders .............................................................................................................................21
      2.4.4 Factors influencing OGD utilization .........................................................................................26
      2.4.5 OGD value and impact .............................................................................................................35
   2.5 Discussion .......................................................................................................................................39
CHAPTER THREE ....................................................................................................................................44
3 METHODOLOGY ..................................................................................................................................44
   3.1 Rationale .........................................................................................................................................44
   3.2 Research Methods ...........................................................................................................................47
CHAPTER FOUR .........................................................................................................................................50
4 QUANTITATIVE STUDY .......................................................................................................................50
   4.1 Hypotheses Development ..................................................................................................................50
      4.1.1 OGD political impact ..................................................................................................................51
      4.1.2 OGD economic impact ..............................................................................................................51
      4.1.3 OGD social impact .....................................................................................................................52
      4.1.4 Government policies ...............................................................................................................52
      4.1.5 Government actions .................................................................................................................53
Appendix A1. Example of Measurement of Independent Variables adopted from the ODB Expert Survey ................................................................. 158
Appendix A2. Examples of OGD Assessment ........................................................................... 164
Appendix A3. Examples of Measurement of Dependent Variables adopted from the ODB Expert Survey ................................................................. 165
Appendix B. Assumptions Check ............................................................................................ 175
Appendix C1. Interview Consent ............................................................................................. 178
Appendix C2. Interview Protocol ............................................................................................. 179
LIST OF FIGURES

Figure 1. The number of publications in 2009-2019 ................................................................. 9
Figure 2. The number of publications by authors’ countries ....................................................... 9
Figure 3. Overview of research methods ...................................................................................... 10
Figure 4. First generation activity theory model VS extended activity theory model .................. 14
Figure 5. Modified and extended DOI research model................................................................. 15
Figure 6. The original technology acceptance model ..................................................................... 16
Figure 7. Unified theory of acceptance and use of technology (UTAUT) model ......................... 17
Figure 8. Subject domains of OGD utilization ................................................................................. 18
Figure 9. Framework of current research on OGD utilization ......................................................... 40
Figure 10. Framework of research rationale ................................................................................. 46
Figure 11. Research model: OGD political impact ........................................................................ 58
Figure 12. Research model: OGD economic impact ...................................................................... 58
Figure 13. Research model: OGD social impact .......................................................................... 58
Figure 14. Framework of case analysis .......................................................................................... 79
Figure 15. Stakeholders interviewed related to the utilization and production of the NYS GIS web services ............................................................................................................. 84
Figure 16. Screenshot of Municipal Parking shown in the Public GIS viewer in the town of Huntington ...................................................................................................................... 95
Figure 17. Coordination among stakeholders of the NYS GIS web services. .............................. 96
Figure 18. Enriched framework based on the quantitative and qualitative Study ....................... 129
LIST OF TABLES
Table 1 Overview of the Article Selection ................................................................. 8
Table 2 Overview of Theoretical Frameworks/models Related to OGD Utilization ............ 12
Table 3 The Number of Articles about OGD Utilization at Different Stages .................... 19
Table 4 Stakeholders in OGD Utilization .................................................................... 22
Table 5 Influential Factors of OGD Utilization and Impact Generation ......................... 27
Table 6 An Example of Measurement of Indicator of the Independent Variable ............... 60
Table 7 The Number of Countries by Regions ............................................................ 61
Table 8 Cluster of OGD ............................................................................................. 63
Table 9 Constructs and Data Sources .......................................................................... 65
Table 10 Descriptive Analysis .................................................................................... 68
Table 11 Cronbach’s alpha ......................................................................................... 69
Table 12 The Results of Multiple Regression .................................................................. 70
Table 13 True Hypotheses ......................................................................................... 72
Table 14 The Type and Number of Stakeholder Interviewed .......................................... 81
Table 15 The Number of The NYS GIS Web Services by Types and by Agencies (by 4/21/2021) ............................................................................................................ 83
Table 16 Users’ Motivations of Using the NYS GIS Web Services ................................. 91
Table 17 Facilitators of Producing the NYS GIS Web Services .................................... 102
Table 18 Obstacles Encountered in Producing the NYS GIS Web Services ................. 104
Table 19 Facilitators of Using the NYS GIS Web Services .......................................... 105
Table 20 Obstacles Encountered in Using the NYS GIS Web Services ......................... 110
CHAPTER ONE

1. INTRODUCTION

Open data, specifically open government data (OGD), is released by governments in machine readable formats and with licensing that allows anyone to have free access and re-use for any purpose. Since 2009, the United States, United Kingdom, Canada and other developed countries have first announced their open data initiatives towards opening up public information and data (Open Knowledge International, 2010). The development of OGD has become popular all over the world. According to the World Bank (2019), more than 250 governments at national, subnational and city levels, almost 50 developed and developing countries, and entities such as the World Bank and United Nations have launched OGD initiatives, and more are launched every year. Membership in the Open Government Partnership (OGP) has gone from eight in 2011 to seventy-eight countries and fifty-six local members in 20191. More and more countries are motivated to follow the trend and start to design and implement their own OGD initiatives. By doing so, governments aim to improve transparency and enable innovation in public service and business products to facilitate the development of economy and the improvement of life quality (Janssen et al., 2012).

In order to reveal the value of OGD, just making the data available is not enough. OGD has to be actually used by people (Janssen et al., 2012; Attard et al., 2016). Different stakeholders, such as government agencies, developers, non-profit organizations, for-profit organizations, academic researchers, general citizens, etc., are involved in OGD development (Sangiambut & Sieber, 2017; Kassen, 2018; Lassinantti, et al., 2019). Various factors related to

1 https://www.opengovpartnership.org/our-members/
governments, users, communities and open data itself interact with each other and affect OGD utilization and impact generation (Zuiderwijk et al., 2014; Zuiderwijk & Janssen, 2014). However, limited research to date has investigated the process of OGD utilization and little is known about how different factors and stakeholders play a role in OGD utilization and impact generation (Safarov et al., 2017). This dissertation attempts to bridge the gaps in our knowledge through a comprehensive literature review, a quantitative study to assess the weight of different influential factors of OGD impact, and a qualitative study to explore the interaction among OGD stakeholders and their perceptions of the facilitators, obstacles, and the benefits in OGD utilization. By providing empirical evidence and practical implications, this dissertation aims to better understand issues related to OGD utilization and impact generation and help improve the practice of OGD utilization and OGD development in the future.

1.1 Background and Related Work

OGD, as an emerging concept, has drawn a great deal of attention of researchers from different fields, including information science, digital government, public administration, etc. It is possible to identify four types of research on OGD in the past decade (Saxena, 2018b). First, theoretical and conceptual OGD-focused research has been discussed a lot at the early stage of OGD development. These research studies mainly focused on explaining the concept of OGD and OGD initiatives (Janssen et al, 2012; Zuiderwijk et al, 2012; Wirtz & Birmeyer, 2015). Second, applied research on OGD has concentrated on the implementation of OGD initiatives in different contexts, including developed countries (Janssen & Zuiderwijk, 2014; Thorsby et al., 2017), and developing countries (Piotrowski et al., 2009; Afful-Dadzie & Afful-Dadzie 2017a; 2017b). A comparative study approach was also adopted to assess different contextual factors that affect the implementation of OGD initiatives (Rothenberg, 2012; Styrin et al., 2017). Third, user-focused
research on OGD, of which the concentration has shifted from OGD provision to OGD utilization. It recognized the diversity of OGD users and explored the influential factors on OGD adoption and utilization from users’ perspectives (Saxena & Janssen, 2017; Saxena & Muhammad, 2018; Ruijer et al., 2017a, 2017b; Wirtz et al., 2019). Fourth, with the development of OGD and increasing utilization of OGD, researchers also started to explore the generation of OGD value and impact (Ahmadi Zeleti et al., 2016; Alexopoulos et al., 2016; Attard et al., 2016).

1.2 Problem Statement

Governments expect to improve transparency and enable innovation in public service and business development through opening government data (Janssen et al., 2012). With the belief that making more data available in more formats will lead to more use and impact, increasing numbers of open data portals have been launched by local governments to improve access to OGD. As of early 2018, there are about 90 cities in the US possessing some types of OGD portals (Brown, 2018). However, even though the amount of open data has increased, the citizen participation and engagement in OGD utilization have not yet reached the anticipated levels (Helbig et al., 2012; Jetzek, et al., 2014; Susa et al., 2015b; Ruijer et al., 2018; Lassinantti, et al., 2019). Despite the potential significant impact of OGD, the lack of OGD utilization has made it far from generating the OGD impact that most people anticipated (Jetzek, et al., 2014; Zuiderwijk & Janssen, 2015; Smith et al., 2016).

With the development of OGD, many related research studies have been conducted. However, most of them focused on the development of OGD initiatives and OGD provision, and very few rigorously discussed the process of OGD utilization (Zuiderwijk et al., 2015; Gill et al., 2017; Hivon, 2017; Safarov et al., 2017). We have limited knowledge about who is actually
using the OGD. How are they using it? What are they using it for? What are the influential factors that affect their use of OGD? In the meantime, while there has been much anticipation about the impact of OGD, there is a lack of empirical research that assessed the generation of OGD value and impact (Helbig et al., 2012; Ubaldi, 2013; Craveiro et al., 2016; Morelli et al., 2017; Safarov et al., 2017; Hossain, et al., 2018; ; Zuiderwijk et al., 2018; Jetzek et al., 2019; Meng et al., 2019), and little is known about what type of OGD impact has been generated? How has it been generated? Without a proper understanding of the reality of OGD utilization and impact generation, it is less efficient and effective to discuss the policies and technological aspects related to the sustainable development of OGD initiatives (Lassinantti et al., 2019).

Faced with this gap in academic research and the practical need to better understand and facilitate the OGD use, this dissertation is designed 1) to review and summarize the current research on OGD utilization; 2) to examine the relationship between the government related factors, OGD related factors, user related factors, community related factors and the generation of different types of OGD impact among different countries over time; and 3) to discuss the roles and interaction among stakeholders in OGD utilization with identifying their perceptions of the facilitators and obstacles during this process through a case study focusing on the New York State Geographic Information System (NYS GIS) web service. Practical implications are also provided to help improve OGD utilization and impact generation in future practices.

The overarching research question of this dissertation is: What are the key factors influencing stakeholders’ OGD utilization and impact generation?

To better address this research question, I have broken it down into several specific questions:
RQ1: What are the most influential factors in the generation of different types of OGD impact among different countries over time?

RQ2: How do stakeholders interact with each other in the process of OGD utilization?

RQ3: How do different stakeholders perceive the facilitators and obstacles in the process of OGD utilization?

This dissertation has adopted a multi-method approach, and the research questions above were addressed using two different research methods. For RQ1, using secondary data from different sources, a quantitative data analysis was conducted to assess what are the most influential factors that affect the generation of OGD impact among different countries over time. For RQ2 and RQ3, focusing on a specific case, the NYS GIS Web Services, a qualitative study was conducted to explore the process of OGD utilization with capturing the interaction among stakeholders and their perceptions of the facilitators and obstacles in the process of OGD utilization. Detailed descriptions about the research methods were illustrated at the beginning of Chapter Four and Five.

1.3 Structure of the Dissertation

This dissertation contains six chapters. Chapter One introduces the background and related work, research questions and the structure of the dissertation. Chapter Two presents a comprehensive literature review that summarizes the current research on OGD utilization. Chapter Three discusses the overall research rationale and the research methods. A quantitative study on the influential factors of OGD impact and a qualitative study on the interaction among stakeholders in the utilization of the NYS GIS Web Services were illustrated in detail in Chapter Four and Five. Conclusion, comparative discussion of quantitative study and qualitative study, implications, limitations and directions for future studies are presented in Chapter Six.
CHAPTER TWO

2 LITERATURE REVIEW

The popularity of OGD development all around the world has attracted the attention of many academic researchers. Several literature reviews have been conducted. Some summarized the current status of research on OGD in general with capturing the shift of research concentration (Saxena, 2018b), global progress (Lv & Ma, 2019); most studied themes (Corrales-Garay et al., 2019; Lv & Ma, 2019) and limitations of OGD research (Ali Hassan & Twinomurinzi, 2018). Some research summarized specific aspects of OGD development, such as assessing OGD initiatives with focusing on OGD provision (Attard et al., 2015), investigating OGD utilization in scientific research (Yan & Weber, 2018) and exploring the impact of OGD on democratic societies (Ruijer & Martinius, 2017). However, little attention has been given to the process of OGD utilization and impact generation. Safarov et al. (2017) summarized the types, stakeholders, influential factors and the effect of OGD utilization, however, this study only reviewed research articles from 2011 to 2015. More recent research about OGD utilization needs to be included in the discussion to better understand the status of research on OGD utilization and impact generation. In this research, the literature review was conducted based on the established procedures of systematic literature review (Okoli & Schabram, 2010; Petticrew & Roberts, 2006) to obtain an overview of the current research (2009-2019) about OGD utilization and generation of OGD value and impact. Following sections illustrate these procedures in detail.

2.1 Research Collection

The literature search used different databases, including Web of Science, Academic Search Complete, Business Source Complete, and ScienceDirect. Boolean operators were used to
combine multiple terms to find articles, which are ("Open government data" OR "open data") <subject> AND ("use" OR "utilization" OR "value" OR “impact”) <all fields> and ("Open government data" OR "open data") <title> AND ("use" OR "utilization" OR "value" OR “impact”) <all fields>. I used the same search terms but put them into different search fields (subject or title) to avoid missing related articles. The review only focused on English academic articles, book chapters and conference proceedings about OGD utilization and OGD value and impact that were published between 2009 and 2019.

2.2 Select Criteria

When the search was completed, a manual study selection was conducted based on the exclusion and inclusion criteria. My review included both theoretical and empirical research, but only studies about OGD utilization, value and impact were considered relevant. The article assessment was processed in two rounds. In the first round, duplications within the two-time search results of each database were removed, and articles were assessed based on their titles and abstracts. Articles related to linked open data, and open data provided by non-government agencies were excluded. In the second round, I combined all the articles from different databases together, removed the duplications and assessed the full text of the articles. Articles that related to OGD utilization, OGD stakeholders, OGD value and impact were included. This resulted in 146 publications as references for this literature review. Table 1 summarizes the process of the article selection.
Table 1 *Overview of the Article Selection*

<table>
<thead>
<tr>
<th>Keyword</th>
<th>Database</th>
<th>No. of Article Found$^2$</th>
<th>No. of Articles After 1st Round Review</th>
<th>No. of Articles After 2nd Round Review</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. (&quot;Open government data&quot; OR &quot;open data&quot;) &lt;subject&gt; AND (&quot;use&quot; OR &quot;utilization&quot; OR &quot;value&quot; OR &quot;impact&quot;) &lt;all fields&gt;</td>
<td>Web of Science</td>
<td>$N_1 = 2051$</td>
<td>$N_1 = 114$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Academic Search Complete</td>
<td>$N_2 = 611$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Business Source Complete</td>
<td>$N_1 = 165$</td>
<td>$N_2 = 325$</td>
<td>$N = 34$</td>
</tr>
<tr>
<td>2. (&quot;Open government data&quot; OR &quot;open data&quot;) &lt;title&gt; AND (&quot;use&quot; OR &quot;utilization&quot; OR &quot;value&quot; OR &quot;impact&quot;) &lt;all fields&gt;</td>
<td>Science Direct</td>
<td>$N_1 = 91$</td>
<td>$N_2 = 121$</td>
<td>$N = 38$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>$N = 3860$</td>
<td>$N = 222$</td>
<td>$N = 146$</td>
</tr>
</tbody>
</table>

2.3 *Descriptive Analysis*

2.3.1 Time and area of publications

This section provides a descriptive analysis of the background information of the selected articles, including the number of the publications over time, the number of the publications by authors’ countries and the research methods adopted by these articles.

Figure 1 presents the number of academic publications on OGD utilization over time.

Before 2010, there were no articles about OGD use, but as the development of OGD proceeds, more and more researchers start to investigate the utilization of OGD, and there are more and more publications related to this topic year by year. This increase also indicates that there is a need for a more comprehensive literature review on research related to this topic.

$^2$ Note: $N_1$ refers to the number of articles obtained by search using the first set of keywords. $N_2$ refers to the number of articles obtained by search using the second set of keywords.
By looking at authors’ countries (see Figure 2), I noticed that most of the research was conducted in developed countries, like The Netherlands, US, UK, etc., where there is a lot of OGD development in practice. In the meantime, I also noticed that developing countries started to pay attention to this topic and conduct more research about it, such as China, India, etc.
2.3.2 Research methods

OGD utilization has been studied by researchers from different perspectives using different research methods. Excluding articles that only mentioned OGD as a data source for their research, I examined 97 articles that focused on issues related to OGD utilization. Among these articles, 74 of them are empirical research studies, including 36 qualitative research studies, 24 quantitative research studies and 14 research studies with mixed methods, 13 of them are theoretical research studies and the rest 10 are literature reviews (see Figure 3).

![Figure 3. Overview of research methods](image)

For the qualitative research studies, most of them took the case study approach combined with document analysis, which helped them obtain more background information about the OGD development. Interviews were often adopted to collect more in-depth feedback from interviewees about their opinions on the use and impact of OGD. Lassinantti, et al. (2019) coded and analyzed the content of the 2011-2016 topic reports published on the knowledge portal hosted by EU commission to identify different OGD user groups outside the public sector and the motivations behind their utilization. Following an exploratory qualitative case study approach, Ruijer et al. (2017a) illustrated the utilization and impact of OGD in different types of democratic processes through workshop and follow up interviews with experts in academic research and public administration fields. Focused on the OGD utilization in the private sector in Pakistan, Saxena &
Muhammad (2018) reviewed the OGD portal and policy in Pakistan and interviewed 49 senior management representatives from private sector and NGOs to understand their perspectives on OGD utilization for professional purposes. McBride et al. (2019) investigated the use of OGD in food safety inspection forecasting model in the City of Chicago and concluded that motivated stakeholders, innovative leaders, proper communication, an existing OGD portal, external funding, and agile development are the six factors that played the key roles in allowing OGD-driven public service co-creation to take place.

For the quantitative research studies, based on previous research about related key concepts and theoretical frameworks and models, many studies raised hypotheses, adopted appropriate indicators, utilized survey questions to collect data, analyzed the relationship among variables to test hypotheses and answer their research questions. Talukder et al. (2018) conducted a quantitative cross-sectional survey to collect feedback about what affected the acceptance and use of OGD in Dhaka city. All measurements contained in the survey were adopted from concepts in previous research studies and were adjusted to fit the research model that combined the unified theory of acceptance and use of technology (UTAUT) and IS success models. Similarly, based on the technology acceptance model (TAM), Wirtz et al. (2019) extended the model and investigated the determinants of OGD use by citizens in Germany through hypothesis testing using survey data. Fitriani et al. (2019) explored the determinants of users’ continued intention to use open data websites in Indonesia. Besides survey methods, some research also conducted experiments to examine users’ interaction with OGD. Zuiderwijk et al. (2016) conducted a quasi-experiment to allow participants to operate on the OGD infrastructure prototype that they have built, and feedback from the participants was collected afterwards to evaluate the prototype.
2.4 Thematic Analysis

In this section, I will present the results of thematic analysis, including the theoretical frameworks used in OGD utilization related research studies, and aspects related to OGD utilization, such as types of utilization, stakeholders, influential factors and OGD impact.

2.4.1 Theoretical frameworks

As the development of open data initiatives around the world, more researchers started to focus on investigating the utilization of OGD. Different theoretical frameworks and models have been used to raise hypotheses and address research questions related to OGD utilization (see Table 2).

Table 2 Overview of Theoretical Frameworks/models Related to OGD Utilization

<table>
<thead>
<tr>
<th>Theoretical Frameworks/Models</th>
<th>No. of articles that used this framework/model³</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unified Theory of Acceptance and Use of Technology (UTAUT)</td>
<td>5</td>
</tr>
<tr>
<td>Technology Acceptance Model (TAM)</td>
<td>5</td>
</tr>
<tr>
<td>Activity Theory (AT)</td>
<td>2</td>
</tr>
<tr>
<td>Diffusion of Innovation Model</td>
<td>1</td>
</tr>
<tr>
<td>Technology Organization Environment (TOE) Framework</td>
<td>1</td>
</tr>
<tr>
<td>IS Success Model</td>
<td>1</td>
</tr>
<tr>
<td>Expectation Confirmation Theory</td>
<td>1</td>
</tr>
<tr>
<td>Theory of Social Construction of Technology</td>
<td>1</td>
</tr>
<tr>
<td>Institutional Theory</td>
<td>1</td>
</tr>
<tr>
<td>Total: 9 theoretical frameworks/models</td>
<td>Total 15 articles</td>
</tr>
</tbody>
</table>

OGD programs are viewed by researchers as “innovations that result from new government services that are offered via information technology (IT) platforms” (Wang & Lo, 2016). Among the selected articles, researchers have taken advantage of theories related to IT

³ Note: Some articles used more than one framework/model
development and adoption on researching the influence of factors on users’ intention of using OGD from the perspectives of different stakeholders.

Based on the theory of social construction of technology (Williams & Edge, 1996), researchers stated that as a multidirectional process, the OGD development can evolve different outcomes under the influence of related human actors (Pinch & Bijker, 1987). In this case, through document analysis from a social technical perspective, five related social groups were identified with different motivations to use OGD, including exploration, creating business value, enabling local citizen value, addressing global societal challenges, and advocating the open data agenda. It revealed the diversity of OGD use and engagement among different types of users (Lassinantti, et al., 2019).

Similarly, rooted in constructivism, The Activity Theory (AT) proposes that aiming at an object, human behavior is mediated by artefacts within the surrounding environment (Vygotsky, 1978). Based on the first generation of the AT model, Khayyat & Bannister (2017) built an extended AT model with ten components (see Figure 4) to investigate how people engage with OGD portals/tools and co-create value as the outcome, and how other factors, including barriers, motivations, division of labor, level of awareness, environment, and rules affect the users’ activities. In the context of democracy, using the AT, Ruijer et al. (2017a) suggested that different OGD uses were driven by different types of democracy (monitorial, deliberative and participatory democracy) (Meijer, 2012), and mediated by different levels of access of OGD.
Figure 4. First generation activity theory model VS extended activity theory model

Rather than focusing on the impact of individual human actors, researchers also studied the impact of institutional factors on the utilization of OGD. Altayar (2018) discussed the motivational factors for adopting OGD among government institutions in Saudi Arabia from an institutional theory perspective. Institutional theory views institutions as “multifaceted, durable social structures, made up of symbolic elements, social activities, and material resources” (Scott, 2013). Three elements of the institutional theory were adopted as theoretical foundations, including institutional pillars, institutional pressure, and rationalized myths, to guide researchers to explore their impact on OGD adoption among government institutions.

Opposite to the impact of human actors on technology development, researchers also looked in the other direction: influential factors affecting individuals’ intentions to adopt innovative technology. Derived from the Diffusion of Innovation Theory (DOI), Weerakkody et al. (2017a) modified and extended the original model by adding the risk factors together with function value, open data usability and users’ perceptions to examine their influence on the intentions of OGD use (see Figure 5).
Figure 5. Modified and extended DOI research model

The Technology Acceptance Model (TAM) is one of the most frequently used models in information technology research. The basic idea is to investigate two basic factors: perceived ease of use and perceived usefulness and their impact on the individual’s acceptance and intentions to use certain technology (Davis, 1986) (see Figure 6). Many researchers have modified and extended the original model and applied it in their research. Through literature review, Wirtz et al. (2018) identified the ease of use, usefulness, intrinsic motivation, extrinsic motivation and Internet competence, and examined their impact on citizens’ intention to use OGD. The motivation and internet competence factors make the original model more explicit.

Besides perceived usefulness and perceived ease of use, the model by Weerakkody et al. (2017b) also included the impact of (perceived) social influence factors on the users’ intention of OGD use. Wirtz et al. (2019) extended the base model by including transparency expectation and participation expectation to investigate the impact of users’ expectations. They also added word of mouth concerning OGD as a further outcome after the intention to use OGD to see what the intention to use OGD will lead to.
Another popular theoretical framework used in these selected articles is the Unified Theory of Acceptance and Use of Technology (UTAUT) (see Figure 7). Its main idea is that several factors have influences on users’ intention to accept and use a system or technology, and this influence is mediated by different individual characteristics, including gender, age, experience and voluntariness of use. The behavioral intentions will further affect the actual use behavior (Sykes et al., 2009). Compared to other models, UTAUT explains much more variance (about 70%) in the behavior intention to use a system or technology (Venkatesh et al., 2003).

Developed from the TAM, UTAUT contains some similar factors with different names, such as performance expectancy, which is perceived usefulness in TAM and effort expectancy, which is perceived ease of use in TAM. Besides that, UTAUT also includes (perceived) social influence, (perceived) facilitating conditions, and other mediated factors of behavior intention. An adapted version of the model has been utilized in research about OGD acceptance and usage (Zuiderwijk et al., 2015; Saxena & Janssen, 2017). Based on UTAUT, Bin (2013) re-structured the model to include: perceived performance (perceived benefit, perceived effort), (perceived) social influence, and personal characteristics (personal innovativeness, experience), and it was used to discuss the influential factors of user innovation using OGD (Susha et al., 2015a). Researchers also combined the UTAUT with other models and theories to explore the utilization of OGD. Zuiderwijkz & Cligge (2016) conducted an integrated model using the UTAUT and the two-stage (pre and post stage) Expectation Confirmation Theory of Information Systems Continuance
(ECT) to investigate the influence on acceptance and use of OGD infrastructure and continuance of use in the future. Talukder et al. (2018) combined the UTAUT and the Information Systems (IS) Success Model to capture the indicators of IS success: system quality and information quality, and their impact on performance and effort expectancy and further impact on users’ behavioral intention to use OGD.

**Figure 7.** Unified theory of acceptance and use of technology (UTAUT) model

Unlike most of the researchers, Wang & Lo (2016) recognized the limitations of the TAM and the UTAUT, which are that they mainly provide explanations for individual adoptions behavior rather than adoption at an organizational level. They took advantage of the Technology-Organization-Environment (TOE) Framework by Tornatzky and Fleischer (1990) and examined the impact of technological context (perceived benefits and barriers), organizational context/readiness and external environmental/pressure on OGD adoption.

2.4.2 OGD utilization

Among the selected articles that are related to the practice of OGD utilization, some just mentioned OGD as a data source for their research in different subject domains, such as environment, public health, transportation etc. (see Figure 8). The others are more focused on
discussing issues related to OGD utilization, such as influential factors, stakeholders, value and impact, etc.

Figure 8. Subject domains of OGD utilization

Focused on the utilization of OGD on data.gov.uk website, Davies (2010) viewed the utilization of OGD as a process which can be categorized into five stages: 1) data to fact, in which, users directly identify a specific fact through searching, browsing and extracting from OGD; 2) data to data, in which, users combine and manipulate datasets to form and share a new dataset; 3) data to information; in which, users create a presentation and interpretation of datasets in texts or graphic; 4) data to interface, in which, users create a method to interactively access and explore the datasets; and 5) data to service, in which OGD is used to provide online and offline services. These stages are not mutually exclusive in OGD practice, and the utilization of OGD often goes through multiple processes. However, for the convenience of categorization, I grouped these examples of OGD utilization mentioned in the articles into different stages by considering the end products of the utilization (see Table 3).

Table 3: Subject Domains of OGD Utilziation

<table>
<thead>
<tr>
<th>Subject Domains</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Environment</td>
<td>7</td>
</tr>
<tr>
<td>Public health</td>
<td>5</td>
</tr>
<tr>
<td>Transportation</td>
<td>5</td>
</tr>
<tr>
<td>Urban Planning</td>
<td>5</td>
</tr>
<tr>
<td>Public administration</td>
<td>4</td>
</tr>
<tr>
<td>Emergency management</td>
<td>3</td>
</tr>
<tr>
<td>Public Safety</td>
<td>3</td>
</tr>
</tbody>
</table>

4 Here is the list of subject domains that appear in more than 2 articles.
Table 3 *The Number of Articles about OGD Utilization at Different Stages*

<table>
<thead>
<tr>
<th>Stages of OGD Utilization</th>
<th>Number of Articles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data to Fact</td>
<td>0</td>
</tr>
<tr>
<td>Data to Data</td>
<td>2</td>
</tr>
<tr>
<td>Data to Information</td>
<td>31</td>
</tr>
<tr>
<td>Data to Interface</td>
<td>9</td>
</tr>
<tr>
<td>Data to Service</td>
<td>7</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>49</strong></td>
</tr>
</tbody>
</table>

Among the examples of data to information, open data analytics strategies have been greatly used in different research fields, including mining crime data to explore crime behaviors and predict future crimes (Saltos & Cocea, 2017), analyzing educational data to construct indicators that capture the evolution of research careers (Cañibano et al., 2016), providing an alternative method to estimate the value of public procurement using open economic data (Marin de la Iglesia, 2014; Surasvadi et al., 2017), tracking vulnerability of climate change (Tiepolo & Bacci, 2017), exploring methods for emergency management (Tkachenko et al., 2016; Noymane et al., 2017; Eem et al., 2018; Wu et al., 2018; Jato-Espino et al., 2019), and assessing the impact of neighborhood environment on residents’ health (Yen & Li, 2017; Hasthanasombat & Mascolo, 2019). The availability of OGD also made it possible to test and validate theoretical models, such as quantifying the potential impact of taxi ridesharing and shared bicycles on public transport travel times (Jeppesen et al., 2013; Barann et al., 2017), and predicting spatial crime distribution (Belesiotis et al., 2018). OGD visualization is another important method that helps general users better understand the complicated datasets and facilitate the communication among stakeholders with different backgrounds (Graves & Hendler, 2014; Estrada et al., 2015; de Mendonça et al., 2015; Brugger et al., 2016; Hasegawa et al., 2019). For example, visualizing the electric power consumption data to help citizens better understand the situation and engage in discussion and policy making (Kukimoto, 2014), creating maps of the infestation of Aedes
aegypti, also called yellow fever mosquito, which can spread dengue fever, chikungunya, Zika fever and other disease in urban areas, to better inform the community (Mendonça, et al., 2015), analyzing textual data using visual analytic approaches to potentially improve the effectiveness and efficiency in public policy making (Hagen et al., 2019).

With the development of urbanization, emerging issues within the community call for collaborative efforts in decision making, and data driven decision making is becoming more and more popular across the development of OGD. However, utilization of OGD often encounters challenges in reality due to the volume and diversity of OGD. A great number of OGD datasets in different formats and in different subjects using different coding schemes, makes it very difficult for general users to locate and understand the datasets. Researchers have devoted great efforts on developing interfaces, models and applications to help users better access and understand OGD. For example, many city dashboards were built to allow planners, policy makers and the communities to easily access information about city performance in different aspects and collaborate on understanding and managing complex community issues (Pettit et al., 2017). Incorporated with 3D simulation techniques, environment and geographic data have been used to represent areas of flood risk, which served as a great reference for emergency management (Yang, et al., 2018).

With better access and understanding of OGD, more and more researchers and developers have started to explore OGD based innovation and service design in public and private sectors. Through system building using OGD about nutritional information, game-based training has been provided to engage citizens and improve their health literacy and make a healthier life choice (Dunwell et al., 2016). Systems combined data about the city’s daily environment collected by citizens and through open sensor data to provide suggestions for
citizens’ daily commutes, which in the long term, helped improve their quality of life (Nallur et al., 2015). OGD innovation plays a significant role in empowering citizens and advancing the development of society. Hackathons and innovation contests have been hosted to connect different stakeholders and increase civic engagement in OGD based innovation (Juell-Skielse et al., 2014). On the other side, internal and external challenges against adopting OGD for service innovation especially for commercial use have been discussed (Magalhães & Roseira, 2016, 2017; Zuiderwijk et al., 2016).

By summarizing the examples of OGD utilization, it is clear to see that most cases of OGD utilization fall into the later three stages: data to information, data to interface and data to service. This indicates that people have recognized the value of OGD and are working on using it to obtain useful information, design interfaces and applications, and provide better services for society. However, the utilization mainly stays at the data to information stage. It seems that the more advanced utilization requires a higher level of data literacy and analytic skills that most of the general public are not equipped with.

2.4.3 Stakeholders

Among the selected articles, researchers also explored users’ experience of OGD utilization by discussing their roles in the process of utilization and the influential factors related to OGD utilization from their perspectives. Most researchers categorized these stakeholders based on the types of organizations that they come from or their roles in OGD utilization (see Table 4). Different types of stakeholders may play a similar role in the process of OGD utilization. Following is a detailed description of different types of OGD stakeholders.
Table 4 *Stakeholders in OGD Utilization*

<table>
<thead>
<tr>
<th>The Types of Stakeholders</th>
<th>The Roles of Stakeholders</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Governments (Federal, State, City/Local)</td>
<td>• Communities</td>
</tr>
<tr>
<td>• Non-government Organizations (NGOs)</td>
<td>• Developers</td>
</tr>
<tr>
<td>• Private Organizations (Companies, Entrepreneurs)</td>
<td>• Academics</td>
</tr>
<tr>
<td></td>
<td>• Media/Journalists</td>
</tr>
<tr>
<td></td>
<td>• Citizens</td>
</tr>
<tr>
<td></td>
<td>• Data providers</td>
</tr>
<tr>
<td></td>
<td>• Infomediaries</td>
</tr>
<tr>
<td></td>
<td>• End users</td>
</tr>
</tbody>
</table>

Governments and their departments act as the dominant control points in OGD development. They oversee publishing, modification, update, access and management of OGD (Sangiambut & Sieber, 2017). Through OGD, governments aim to improve transparency as well as achieve social and economic value (Janssen et al., 2012). In order to achieve these, governments are facing the challenges of transformation.

The first challenge is to change the traditional mindset about data, which used to be mainly for internal use, but now is to open to the public (Gill et al., 2017; Ma & Lam, 2019). The launch of OGD initiatives and OGD portals is a significant first step to make data available not only to different government departments, but more importantly to the general publics (Johnson & Greene, 2017; Saxena & Muhammad, 2018).

The second challenge is that standards and technical training are needed to guide governments to provide data in ways that are compatible with other governments at different levels (Gill et al., 2017). Low data quality and missing metadata are some of the common issues that impede the use of OGD. It is difficult for users to understand the datasets, generate useful information, and create successful applications using low quality OGD and data with incomplete or missing metadata. (Zuiderwijk et al., 2016; Chorley, 2017; Beno, et al., 2017; Benitez-Paez et al., 2018; Talukder et al., 2018).
The third challenge is concern about privacy and information security. As the manager of public data and information, governments often struggle between the need of publishing data and making sure the information security and privacy are protected (Susha et al. 2015c; Beno et al., 2017; Ingrams, 2017).

The lack of a citizen-centric approach in OGD provision is another challenge. Most data are made open by governments’ own choices. Due to the lack of feedback from the users, governments have little knowledge about what is the type of OGD users, how are they using it and what are their user experiences (Susha et al., 2015b). In the end, the data published by the governments may not suit the needs of users, which leads to limited use of OGD by the public and potential waste of governments’ efforts in OGD development (Johnson & Greene, 2017).

Last but not least, due to the lack of resources and funding, especially in small size municipalities, some local governments do not have the capacities to make open data available (Beno et al., 2017; Gill et al., 2017). Governments have engaged with external parties to leverage their expertise to help deal with these challenges; outsourcing is now a common way for governments to solve problems. For example, a technical company has helped governments design the interface of the OGD portals (Sangiambut & Sieber, 2017).

OGD infomediaries, also called OGD intermediaries, is another frequently mentioned type of stakeholder in the articles. It refers to a specific group of OGD users who bridge the gap between OGD providers and citizens in that they extract, aggregate, process and transform the data into a format that is more useful and beneficial for the end users (Magalhães et al., 2013; Chan et al., 2016; Johnson & Greene, 2017). In the process of OGD utilization, different types of OGD stakeholders are playing the role of infomediaries with different motivations (Brugger et al., 2016).
Developers are the traditional infomediaries. With a high level of OGD and technology literacy, they are very innovative in creating business and social value through OGD utilization. All these characteristics make them the most active actors in OGD utilization. Usually, they are either outsourced by governments or work for IT firms, and they work on designing platforms, applications, and creating innovative services for public and private sectors (Sangiambut & Sieber, 2017). They also often participate in hackathons and innovation contests, not only aiming to create new products, but also promoting OGD initiatives and expanding the OGD community (Lassinantti et al., 2019). The developers in general are goal-driven, so their major concerns are the scope and quality of the data. The lack of access to open data and limited data formats will hinder their products and services design (Smith & Sandberg, 2018).

Private organizations are one of the few groups of actors that have recognized the value of OGD and tried to take advantage of it at the early stage. With the development of OGD, many OGD-based startup companies are emerging. In the meantime, traditional companies started to bring in new techniques to leverage the potential of OGD to stay ahead of other competitors. The availability of OGD provides them with the opportunities to create innovative services and obtain business value (Smith & Sandberg, 2018; Lassinantti et al., 2019). With the support from financial and technology resources and the demands for OGD, companies are actively building public-private partnerships with the governments who are managing OGD and aim to combine advantages of both sides to better utilize OGD and achieve mutual benefits (Lassinantti et al., 2019). However, in reality, the lack of data, the lack of support, the complexity of the project, leadership, business strategies, differences in organizational culture all hinder the success of this partnership (Maccani, 2016; Magalhães & Roseira, 2016; Susha et al., 2015a).
NGOs usually refer to non-profit organizations that are independent from the governments. They usually serve as advocates who promote the idea of free access to government information and data and encourage civic engagement (Kassen, 2018). Even though they themselves lack the resources and techniques, they have worked on providing suggestions for opening government data and facilitating community engagement in OGD utilization (Johnson & Greene, 2017).

Similar to the NGOs, different types of media have been used to promote OGD through online and offline channels. Social media networks constitute one of them that have a great positive impact on improving the awareness of the general publics about OGD. They help form civic online communities where people share their experience of OGD use (Johnson & Greene, 2017). Media itself is also influenced by the development of OGD. Data journalists obtain the new skills to engage with readers and audiences through providing information and visualization generated by OGD (Lassinantti et al., 2019). However, some researchers claimed that the time cost of understanding the dataset negatively affects the wide use of OGD (Burgger et al., 2016).

Academic researchers are another important stakeholder group that benefits from using OGD in their research. In the meantime, they are also contributing to OGD development through investigating OGD related issues (Yan & Weber, 2018). OGD, as open and free data sources, make it feasible for researchers to conduct certain research and enhance the research results. OGD helps provide useful information for knowledge improvement and decision making. On the other hand, through academic papers and reports on issues of OGD development, researchers investigate the status of OGD utilization, identify the facilitators and obstacles, provide suggestions for its future utilization and impact generation (Johnson & Greene, 2017; Kassen, 2018).
Citizens are usually viewed as the OGD end users. Due to the dynamics in their background, OGD and technology literacy, and interests, their awareness of OGD benefits and intention to use OGD are quite different. Some take advantage of OGD for their own benefit, whereas some explore solutions to address societal issues. Some promote OGD use, whereas some know very little about OGD and think it is irrelevant to them. The perceived usefulness and the ease of use are the most common factors mentioned in the literature that affect citizens’ intention of OGD use (Weerakkody et al., 2017a; Wirtz et al., 2018).

From the summary of research on OGD stakeholder, we can see that government policies and actions that support OGD development are the prerequisite for OGD utilization. The availability and easy access to high quality data make it possible for users to interact with OGD and use it for their own purposes. With technological knowledge and interest in OGD, the OGD infomediaries is the most active actor group that engages with OGD utilization and helps to bridge the gaps between OGD and general users who may have limited knowledge about OGD and lack the abilities to work on it.

2.4.4 Factors influencing OGD utilization

Among the selected studies, I have identified 45 of them that specifically focused on the influential factors of OGD utilization in different stages, including publishing, accessing, analyzing and re-production. Research studies were conducted focusing on OGD utilization in different organizational environments (public sector, private sector, academic community and among general public), and in different countries (developed and developing countries). The findings were extracted and summarized from previous research, including literature review studies, document analysis, surveys and interviews with OGD stakeholders. Influential factors were categorized as government related, OGD related, user related, and community related (see
Table 5). Following is a detailed description of factors in each category that affect the OGD utilization and impact generation.

Table 5 *Influential Factors of OGD Utilization and Impact Generation*\(^5\)

<table>
<thead>
<tr>
<th>Influential Factors</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Government related</strong></td>
<td></td>
</tr>
<tr>
<td>Resources (budget, funding, IT infrastructure, staff, etc.)</td>
<td>Attard et al., 2016; Maccani, 2016; Wang &amp; Lo, 2016; Ingrams, 2017; Kučera, 2017; Ruijer et al., 2017a; 2017b; Jetzek et al., 2019; Jiang et al., 2019; McBride et al., 2019; Ma &amp; Lam, 2019; Zuiderwijk &amp; Janssen, 2013; Susha et al., 2015c; Maccani, 2016; Ingrams, 2017; Khayyat &amp; Bannister, 2017; Kučera, 2017; Hellgren &amp; Nilsson, 2018; Jiang et al., 2019; Ma &amp; Lam, 2019; Zuiderwijk &amp; Spiers, 2019</td>
</tr>
<tr>
<td>OGD legislation and policies</td>
<td>Susha et al., 2015c; Attard et al., 2016; Magalhães &amp; Roseira, 2016; Beno et al., 2017; Khayyat &amp; Bannister, 2017; Ruijer et al., 2017a; 2017b; Hellgren &amp; Nilsson, 2018; Jetzek et al., 2019; Ma &amp; Lam, 2019; Zuiderwijk &amp; Janssen, 2013; Kaasenbrood, et al., 2015; Susha et al., 2015b, 2015c; Maccani, 2016; Magalhães &amp; Roseira, 2016; Khayyat &amp; Bannister, 2017; Smith &amp; Sandberg, 2018; Ma &amp; Lam, 2019</td>
</tr>
<tr>
<td>Privacy and security</td>
<td>Kučera, 2017; Ruijer et al., 2017a; 2017b; Hellgren &amp; Nilsson, 2018; Ma &amp; Lam, 2019</td>
</tr>
<tr>
<td>Interaction with OGD users</td>
<td>Zuiderwijk &amp; Janssen, 2013; Lee et al., 2015; Attard et al., 2016; Maccani, 2016; Kučera, 2017; Ruijer et al., 2017b; Ma &amp; Lam, 2019</td>
</tr>
<tr>
<td>OGD management</td>
<td>Susha et al., 2015b, 2015c; Ruijer et al., 2017a; Smith &amp; Sandberg, 2018</td>
</tr>
<tr>
<td>Concern about unpredicted/negative impact</td>
<td>Kučera, 2017; Ruijer et al., 2017a; 2017b; Hellgren &amp; Nilsson, 2018; Ma &amp; Lam, 2019</td>
</tr>
<tr>
<td>Organizational structure</td>
<td>Attard et al., 2016; Maccani, 2016; Ma &amp; Lam, 2019</td>
</tr>
<tr>
<td>Government IT leadership</td>
<td>Wang &amp; Lo, 2016; Jetzek et al., 2019; Jiang et al., 2019</td>
</tr>
<tr>
<td>Knowledge of OGD users</td>
<td>Susha et al., 2015b; Ruijer et al., 2017b; Benitez-Paez et al., 2018; Hellgren &amp; Nilsson, 2018</td>
</tr>
<tr>
<td>Provision of support to OGD users</td>
<td>Susha et al., 2015b, 2015c; Ruijer et al., 2017a; Smith &amp; Sandberg, 2018</td>
</tr>
<tr>
<td>Organizational culture</td>
<td>Hellgren &amp; Nilsson, 2018; Ma &amp; Lam, 2019</td>
</tr>
<tr>
<td>Develop and promote best practices</td>
<td>Susha et al., 2015c; Khayyat &amp; Bannister, 2017; Ruijer et al., 2017a</td>
</tr>
<tr>
<td>Government’s awareness of OGD</td>
<td>Wang &amp; Lo, 2016; Chorley, 2017; Kučera, 2017; Altayar, 2018</td>
</tr>
</tbody>
</table>

\(^5\) The factors are listed based on the frequency mentioned by previous research (from the most mentioned to the least mentioned)
<table>
<thead>
<tr>
<th><strong>Influential Factors</strong></th>
<th><strong>References</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Provision of incentives for OGD use</td>
<td>Susha et al., 2015c; Hellgren &amp; Nilsson, 2018; Ma &amp; Lam, 2019</td>
</tr>
<tr>
<td>OGD education</td>
<td>Susha et al., 2015c; Maccani, 2016</td>
</tr>
<tr>
<td>City characteristics</td>
<td>Maccani, 2016</td>
</tr>
</tbody>
</table>

### OGD related

<p>| <strong>Metadata</strong> | Charalabidis, et al, 2014; Kaasenbrood et al., 2015; Alexopoulos et al., 2016; Attard et al., 2016; Brugger et al., 2016; Maccani, 2016; Zuiderwijk et al., 2016; Beno et al., 2017; Khayyat &amp; Bannister, 2017; Kučera, 2017; Ruijer et al., 2017b; Benitez-Paez et al., 2018; Saxena, 2018a; Saxena &amp; Muhammad, 2018; Smith &amp; Sandberg, 2018; Talukder et al., 2018; Jiang et al., 2019; Ma &amp; Lam, 2019 |
| <strong>Data accessibility</strong> | Benitez-Paez et al., 2017; Beno et al., 2017; Kučera, 2017; Ruijer et al., 2017a; 2017b; Hellgren &amp; Nilsson, 2018; Smith &amp; Sandberg, 2018; Talukder et al., 2018; Jetzek et al., 2018; Jiang et al., 2019; Ma &amp; Lam, 2019 |
| <strong>Data quality</strong> | Hellgren &amp; Nilsson, 2018; Saxena, 2018a; Smith &amp; Sandberg, 2018; Talukder et al., 2018; Ma &amp; Lam, 2019 McBride et al., 2019; Zuiderwijk &amp; Spiers, 2019 Charalabidis, et al, 2014; Kaasenbrood et al., 2015; Alexopoulos et al., 2016; Attard et al., 2016; Beno et al., 2017; Khayyat &amp; Bannister, 2017; Ruijer et al., 2017b |
| <strong>Data format</strong> | Hellgren &amp; Nilsson, 2018; Saxena &amp; Muhammad, 2018; Smith &amp; Sandberg, 2018; Talukder et al., 2018; Jetzek et al., 2019 |
| <strong>Data availability</strong> | Kučera, 2017; Ruijer et al., 2017a; Smith &amp; Sandberg, 2018; Jiang et al., 2019; McBride et al., 2019; Meng et al., 2019 Charalabidis, et al, 2014; Kaasenbrood et al., 2015; Alexopoulos et al., 2016; Maccani, 2016; Ojo et al., 2016 |
| <strong>Data usability</strong> | Benitez-Paez et al., 2017; Ruijer et al., 2017b; Weerakkody et al., 2017a; Jetzek et al., 2019; Jiang et al., 2019; Zuiderwijk &amp; Spiers, 2019 |
| <strong>Data currency</strong> | Charalabidis, et al, 2014; Kaasenbrood et al., 2015; Alexopoulos et al., 2016; Ruijer et al., 2017b; McBride et |</p>
<table>
<thead>
<tr>
<th>Influential Factors</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open license</td>
<td>Attard et al., 2016; Maccani, 2016; Magalhães &amp; Roseira, 2016; Beno et al., 2017; Benitez-Paez et al., 2018; Talukder et al., 2018; Jiang et al., 2019; Zuiderwijk &amp; Janssen, 2013; Kaasenbrood et al., 2015; Attard et al., 2016; Magalhães &amp; Roseira, 2016; Benitez-Paez et al., 2017; Hellgren &amp; Nilsson, 2018; Saxena, 2018a; Jetzek et al., 2019; Zuiderwijk &amp; Janssen, 2013; Charalabidis, et al, 2014; Alexopoulos et al., 2016; Ojo et al., 2016; Khayyat &amp; Bannister, 2017; Beno et al., 2017; Benitez-Paez et al., 2018</td>
</tr>
<tr>
<td>Data discoverability</td>
<td>Kaasenbrood et al., 2015; Alexopoulos et al., 2016; Maccani, 2016; Saxena &amp; Muhammad, 2018; Jiang et al., 2019; McBride et al., 2019; Charalabidis, et al, 2014; Alexopoulos et al., 2016; Talukder et al., 2018; Ma &amp; Lam, 2019; McBride et al., 2019</td>
</tr>
<tr>
<td>Channel for user feedback</td>
<td>Zuiderwijk &amp; Janssen, 2015; Ojo et al., 2016; Zuiderwijk et al., 2016; Beno, et al., 2017; Benitez-Paez et al., 2018; Zuiderwijk &amp; Janssen, 2015; Ojo et al., 2016; Zuiderwijk et al., 2016; Benitez-Paez et al., 2018</td>
</tr>
<tr>
<td>Data accuracy</td>
<td>Maccani, 2016; Magalhães &amp; Roseira, 2016</td>
</tr>
<tr>
<td>OGD portal capabilities</td>
<td>Talukder et al., 2018; Ma &amp; Lam, 2019; McBride et al., 2019</td>
</tr>
<tr>
<td>Data quality indicators</td>
<td>Zuiderwijk &amp; Janssen, 2015; Ojo et al., 2016; Zuiderwijk et al., 2016; Beno, et al., 2017; Benitez-Paez et al., 2018; Zuiderwijk &amp; Janssen, 2015; Ojo et al., 2016; Zuiderwijk et al., 2016; Benitez-Paez et al., 2018</td>
</tr>
<tr>
<td>Participation mechanisms</td>
<td>Mazzetti et al., 2015; Attard et al., 2016; Ojo et al., 2016</td>
</tr>
<tr>
<td>Data interoperability</td>
<td>Maccani, 2016; Magalhães &amp; Roseira, 2016</td>
</tr>
<tr>
<td>Data standardization</td>
<td>Maccani, 2016; Jetzek et al., 2019</td>
</tr>
<tr>
<td>Data affordability</td>
<td>Kaasenbrood et al., 2015; Meng et al., 2019</td>
</tr>
<tr>
<td>Data credibility</td>
<td>Khayyat &amp; Bannister, 2017</td>
</tr>
<tr>
<td>User related</td>
<td>Charalabidis, et al, 2014; Susha et al., 2015a; Zuiderwijk et al., 2015; Alexopoulos et al., 2016; Maccani, 2016; Wang &amp; Lo, 2016; Zuiderwijk &amp; Cligge, 2016; Kučera, 2017; Weerakkody et al., 2017a; 2017b; Wirtz et al., 2017; 2018; Talukder et al., 2018; Zainal et al, 2018; Ma &amp; Lam, 2019; Zuiderwijk &amp; Spiers, 2019</td>
</tr>
<tr>
<td>Perceived usefulness/Perceived</td>
<td>Charalabidis, et al, 2014; Susha et al., 2015a; Zuiderwijk et al., 2015; Alexopoulos et al., 2016; Maccani, 2016; Brugger et al., 2016; Maccani, 2016; Zuiderwijk &amp; Cligge, 2016; Saxena &amp; Janssen, 2017; Weerakkody et al., 2017b; Wirtz et al., 2017; 2018; Talukder et al., 2018; Zainal et al, 2018; Zuiderwijk &amp; Spiers, 2019; Ma &amp; Lam, 2019</td>
</tr>
<tr>
<td>benefit/Performance expectancy</td>
<td>Maccani, 2016; Zuiderwijk &amp; Cligge, 2016; Saxena &amp; Janssen, 2017; Weerakkody et al., 2017b; Wirtz et al., 2017; 2018; Talukder et al., 2018; Zainal et al, 2018; Zuiderwijk &amp; Spiers, 2019</td>
</tr>
<tr>
<td>Influential Factors</td>
<td>References</td>
</tr>
<tr>
<td>---------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>General interest/awareness and perception about OGD/Attitude towards OGD</td>
<td>Attard et al., 2016; Brugger et al., 2016; Maccani, 2016; Khayyat &amp; Bannister, 2017; Morelli et al., 2017; Weerakkody et al., 2017a; 2017b; Gascó-Hernández et al., 2018; Fitriani et al., 2019; Zuiderwijk &amp; Spiers, 2019</td>
</tr>
<tr>
<td>(Perceived) Social influence/Subject norms</td>
<td>Susha et al., 2015a; Zuiderwijk et al., 2015; Zuiderwijk &amp; Cligge, 2016; Saxena &amp; Janssen, 2017; Weerakkody et al., 2017b; Zainal et al, 2018; Talukder et al., 2018; Zuiderwijk &amp; Spiers, 2019; Fitriani et al., 2019</td>
</tr>
<tr>
<td>Knowledge, skill, capabilities</td>
<td>Brugger et al., 2016; Maccani, 2016; Zuiderwijk et al., 2016; Kučera, 2017; Ruijer et al., 2017a; 2017b; Weerakkody et al., 2017a; 2017b; Gascó-Hernández et al., 2018; Hellgren &amp; Nilsson, 2018; Smith &amp; Sandberg, 2018; Wirtz et al., 2018; Jetzek et al., 2019; Ma &amp; Lam, 2019</td>
</tr>
<tr>
<td>(Perceived) Facilitating conditions/Perceived behavior control</td>
<td>Susha et al., 2015a; Zuiderwijk et al., 2015; Saxena &amp; Janssen, 2017; Talukder et al., 2018; Zainal et al, 2018; Zuiderwijk &amp; Spiers, 2019; Fitriani et al., 2019</td>
</tr>
<tr>
<td>Motivation to use OGD</td>
<td>Davies, 2010; Smith &amp; Sandberg, 2018; Wirtz et al., 2018; McBride et al., 2019; Zuiderwijk &amp; Spiers, 2019</td>
</tr>
<tr>
<td>Trust</td>
<td>Weerakkody et al., 2017b; Zainal et al, 2018; Fitriani et al., 2019; Zuiderwijk &amp; Spiers, 2019</td>
</tr>
<tr>
<td>Voluntariness of use</td>
<td>Zuiderwijk et al., 2015; Maccani, 2016; Saxena &amp; Janssen, 2017; Zuiderwijk &amp; Spiers, 2019</td>
</tr>
<tr>
<td>Experience with OGD</td>
<td>Susha et al., 2015a; Zuiderwijk et al., 2016; Zuiderwijk &amp; Spiers, 2019</td>
</tr>
<tr>
<td>User’s innovativeness</td>
<td>Susha et al., 2015a; Maccani, 2016</td>
</tr>
<tr>
<td>User participation</td>
<td>Jiang et al., 2019; Ma &amp; Lam, 2019</td>
</tr>
<tr>
<td><strong>Community related</strong></td>
<td>Zuiderwijk &amp; Janssen, 2013; Maccani, 2016; Khayyat &amp; Bannister, 2017; Kučera, 2017; Ingrams, 2017; Ruijer et al., 2017a; Smith &amp; Sandberg, 2018; Ham et al., 2019; McBride et al., 2019; Meng et al., 2019; Ruijer &amp; Meijer, 2019; Maccani, 2016; McBride et al., 2019; Meng et al., 2019; Susha et al., 2015a; Gascó-Hernández et al., 2018; Maccani, 2016; Morelli et al., 2017; Maccani, 2016; Meng et al., 2019; Maccani, 2016</td>
</tr>
<tr>
<td>Coordination and communication among OGD stakeholders</td>
<td>Zuiderwijk &amp; Janssen, 2013; Maccani, 2016; Khayyat &amp; Bannister, 2017; Kučera, 2017; Ingrams, 2017; Ruijer et al., 2017a; Smith &amp; Sandberg, 2018; Ham et al., 2019; McBride et al., 2019; Meng et al., 2019; Ruijer &amp; Meijer, 2019; Maccani, 2016; McBride et al., 2019; Meng et al., 2019; Susha et al., 2015a; Gascó-Hernández et al., 2018; Maccani, 2016; Morelli et al., 2017; Maccani, 2016; Meng et al., 2019; Maccani, 2016</td>
</tr>
<tr>
<td>OGD use champions</td>
<td>Maccani, 2016; McBride et al., 2019; Meng et al., 2019</td>
</tr>
<tr>
<td>Training</td>
<td>Susha et al., 2015a; Gascó-Hernández et al., 2018</td>
</tr>
<tr>
<td>OGD contests (e.g., Hackathons)</td>
<td>Maccani, 2016; Morelli et al., 2017</td>
</tr>
<tr>
<td>Intermediary organization</td>
<td>Maccani, 2016; Meng et al., 2019</td>
</tr>
<tr>
<td>Community-led initiatives</td>
<td>Maccani, 2016</td>
</tr>
<tr>
<td>Volunteer engagement</td>
<td>Maccani, 2016</td>
</tr>
</tbody>
</table>
2.4.4.1 Government related factors

Faced with the international trend of open data initiatives and the increasing expectations from citizens, more and more countries have put OGD development on their agendas and aim to improve transparency and provide better services for the public (Altayar, 2018; Wirtz et al., 2017). The internal and external pressures have forced government to change their traditional ways of managing data and information from internal use only to making it available to the public (Wang & Lo, 2016). Management and leadership are the keys for this transformation, and they decide what data to be opened and how it is to be opened (Chorley, 2017; Kučera, 2017; Lee et al., 2015). Some government leaders perceive the benefit of OGD and commit to improve public services through OGD (Altayar, 2018; Chorley, 2017; Kučera, 2017; Wang & Lo, 2016), while the others are concerned about the unpredictable impact of OGD, have fears of losing control and are unwilling to take the risk (Kučera, 2017; Ruijer et al., 2017a; 2017b). Besides that, budget, IT infrastructure, staff time cost are the most mentioned factors that affect the capacity and capability of governments in OGD development, especially for those small size municipalities (Beno, et al., 2017; Maccani, 2016; Ruijer et al., 2017a). In order to make the OGD development sustainable, related policy and standards are needed to guide the process and guarantee the data quantity and quality (Altayar, 2018; Kučera, 2017). In the meantime, privacy and data security are one of the biggest concerns that governments have in the process of making government data open (Ingrams, 2017).

Besides these organizational and cultural challenges in making OGD available, the governments also find it is challenging to provide support to the OGD users who have different requirements and various levels of skills. On one side, governments expect users from different fields to use the OGD they provided, on the other side, they have invested little or no efforts to
better understand the user groups and their user experience (Susha et al., 2015b; Ruijer, et al., 2017a). In order to increase the citizen engagement of OGD and make more OGD to be used by more users, just releasing the data is not enough. Governments should establish mechanisms that help communicate with users and collect their feedback, which in return helps guide governments to provide better quality data that better satisfy users’ needs (Kaasenbrood et al., 2015; Susha et al., 2015c; Ma & Lam, 2019). Governments should improve users’ awareness of OGD through workshops or showcases that present what can be done using OGD (Maccani, 2016; Khayyat & Bannister, 2017). Governments should support training and OGD events that help improve users’ capabilities and engagement in OGD utilization (Susha et al., 2015b; 2015c).

2.4.4.2 OGD related factors

Among the previous literature, factors related to OGD have been mentioned the most that affect OGD utilization and impact generation. These factors have reflected two different dimensions about the properties of OGD: data openness and data liquidity (Jetzek, 2016; Jetzek et al., 2019). The first dimension is about the open access to data, such as data availability, affordability, free license, etc. Even though many datasets have been opened by government, there are still many complaints from the data users about the lack of OGD and difficulties in accessing datasets that are useful for them (Beno, et al., 2017; Magalhães & Roseira, 2016; Smith & Sandberg, 2018). This has to do with the way that governments open the data, which is often based on the government’s choice which may not necessarily fit the needs of the users (Ruijer et al., 2017b; Susha et al. 2015c). In the meantime, governments often encounter legal constraints that only allow data to be open within government departments rather than to the general public (Attard, et al., 2016; Ma & Lam, 2019).
The second dimension is about data liquidity. It refers to the technical characteristics of open data, including usability, discoverability, accessibility, etc. The usability of OGD greatly relies on OGD quality. Some previous research studies only mentioned OGD quality as one general influential factor, whereas some articles explicitly illustrated different indicators of OGD quality, including metadata, data format, data currency, data accuracy, data sustainability, etc (Zuiderwijk et al., 2016; Beno, et al., 2017; Benitez-Paez et al., 2018; Jetzek et al., 2019). With low quality OGD (non-machine-readable format, missing data and metadata, errors, not up to date, etc.), it is infeasible for users to extract useful information from it. All these will directly affect the usability of the data and the effectiveness of the OGD utilization (Beno, et al., 2017; Mazzetti et al., 2015). When open data is presented to the public through an OGD portal, the design of the OGD platform also affects the user experience of OGD. For example, how easy is it for users to locate a dataset through the navigation of the website? Are there any different presentations (visualization, map, etc) of OGD rather than datasets tables? Is there a channel for users to leave comments or submit inquiries? (Hivon & Titah, 2017; Kaasenbrood et al., 2015; Zuiderwijk & Janssen, 2015; Zuiderwijk et al. 2016).

2.4.4.3 User related factors

From the user perspective, many studies discussed the influential factors related to users themselves, including individual users and organizational users. Using the TAM, the UTAUT and other theoretical frameworks, many researchers have assessed users’ intentions to use OGD as an approximation to the actual OGD utilization. The perceived usefulness, perceived ease of use, perceived social influence and perceived facilitating conditions are revealed as the main factors affecting users’ intentions to use OGD and OGD utilization in general (Anneke & Martijn, 2016; Maccani, 2016; Saxena & Janssen, 2017; Susa et al., 2015a; Talukder et al.,
The lack of awareness of OGD among the users is another frequently mentioned factor of OGD utilization. Some people perceive that OGD is unrelated to their daily life, and they have no desire to use it (Anneke & Martijn, 2016; Dechman & Syms, 2014; Hivon & Titah, 2017). Some people have a basic understanding of the OGD, but they have doubts in the level of government transparency and the quality of the data they provided (Weerakkody et al., 2017b; Zainal et al., 2018; Fitriani et al., 2019). Both cases lead to limited users’ participation and engagement in OGD utilization and related activities. OGD utilization requires certain levels of OGD and technical literacy. The knowledge and skills of the users not only affect their perceived value of OGD but also their capabilities to innovate using OGD (Dechman & Syms, 2014; Smith & Sandberg, 2018; Wirtz et al., 2018).

2.4.4.4 Community related factors

The environment is important especially for the development of a new phenomenon. The popularity and professionalization of OGD utilization need the support from all stakeholders within the community. Communication and engagement among stakeholders affect the processes of co-production and co-creation using OGD (Hivon & Titah, 2017; Kučera, 2017; Lee et al., 2015; Maccani, 2016). Hackathons and other OGD contests provide the platform to facilitate the communication and to help explore the mechanisms of collaboration among different organizations and individuals (Juell-Skielse et al., 2014; Morelli et al., 2017). With considering the diversity in capacities and capabilities of the users, tutorials and training are mentioned as methods to help people improve their awareness and skills in OGD utilization (Susha et al., 2015a; Gascó-Hernández et al., 2018).
From this summary of influential factors of OGD utilization, we can see that from OGD provision to OGD utilization, there are mainly four categories of factors that play a role in the process, namely government related factors, OGD related factors, user related factors and community related factors. For the government factors, most of them can be categorized as government policies (e.g., OGD legislations, OGD policies, privacy and security policies, etc.), government actions (e.g., OGD management, resources, support to GOD, etc.). They are the prerequisites of OGD development. The OGD related factors include the various indicators of the properties of OGD (data openness and liquidity) and the capabilities of the OGD platforms. They are the essential aspects that affect the effectiveness of OGD utilization. For the user related factors, users’ awareness of OGD and skills are most mentioned factors that affect user engagement and participation in OGD utilization. The community related factors emphasize the interaction among stakeholders and the supportive environment that facilitates OGD utilization (e.g., training, OGD contest, etc.)

2.4.5 OGD value and impact

Moore (1995) believed that public resources should be used to increase value, not only in economic development, but more importantly in the aspects that are valued by citizens and communities. There are three main reasons for governments to make their data available to the public, namely transparency, participatory governance, and releasing social and commercial value (Attard, et al., 2016). OGD provides citizens and organizations with the access and opportunities to use government data, which makes it possible for them to monitor government performance, engage in policy and decision making and build partnerships with government agencies to collectively solve societal issues (Alexopoulos et al., 2016). It also provides the opportunities for citizens to re-use OGD for their own purposes, which enables innovation and
economic development (Gurin, 2014). Among the selected research studies, I identified 28 of them that focused on OGD value and impact generation. The political, economic and social impact of OGD are discussed in the following sections.

2.4.5.1 Political impact

As the main goal of most OGD initiatives, governments work on improving transparency and accountability through making government data and information available to the public. Through OGD, citizens obtain a better understanding of government work and start to play a bigger role in monitoring the government working process. It helps improve government transparency and accountability, especially those departments that are involved in the process of OGD provision (Attard et al., 2016; Mergel et al., 2018). The availability of OGD also enables collaboration among government agencies and other community organizations in policy making to better address community issues (Charalabidis et al., 2014; Worthy, 2015; Pereira et al., 2017). In the meantime, OGD, as a great resource, can also help governments facilitate their work across departments and improve its effectiveness and efficiency in delivering public services. (Linders, 2013; Mergel et al., 2018).

2.4.5.2 Economic impact

Economic impact of OGD is generated through the utilization of OGD, including aggregation, search, comparison, information extension, delivery, personalization, and translation (Henkel et al., 2017). The availability of OGD boosts OGD-based entrepreneurship and innovation (Mergel et al., 2018; Kukkamäki et al., 2019). New applications, products and services are designed to provide users with easy access to useful information and better services that are convenient and beneficial for their lives. In real cases, it is proved to help users obtain knowledge for decision making, problem solving, cost saving, efficiency improving, and other aspects related
product/service quality and job performance (Bílková et al., 2015; Henkel et al., 2017; Jetzek et al., 2014). However, the knowledge of OGD stakeholders about contextual factors (such culture-specific, organization-specific, etc.), the availability and quality of GOD resources, including data, technology and governance, and the cooperation among OGD stakeholders all affect the generation of economic value of OGD (Zuiderwijk et al., 2016). To better understand the mechanism of OGD impact and value generation. Researchers have adopted the concept of a “business model” to describe the generation of economic value of OGD (Bonina, 2013; Magalhães & Roseire). Ahmadi Zeleti et al. (2016) developed a 6-Value business value framework, including elements of value proposition, value adding process, and value in return, value capture, value management and value network to analyze. They used this framework to categorize the existing OGD business models used in practice into five different types: freemium (offers limited data free of charge and apply fees for additional request for complete and higher quality datasets), premium (data are offered in high quality and complete form at some cost), cost saving (cost of opening and releasing data are reduced by engaging participants and publishing data as Linked Data), indirect benefit (allows the business to develop its own data and data infrastructure by using the third-party infrastructures that are created because the data is open and available) and the parts of tools (offer first set of data at a discount, while offering complementary or dependent data at a considerable higher price). Each category provided a path that can be used to guide different kinds of OGD utilization and methods of value generation.

2.4.5.3 Social impact

Besides political and economic impact, OGD was often expected to have an impact on the life quality of individuals or the improvement of society as a whole (Bonina, 2013; Jetzek et al., 2014). The most mentioned social impact of OGD is increased citizen engagement (Attard et al.,
2016; Craveiro et al., 2016; Gagliardi et al., 2017; Jamieson et al., 2019). Some research studies show that the available OGD helps empower and encourage citizens to participate in civic discussion about various public issues (Jetzek et al., 2014; Bílková et al., 2015; Pereira et al., 2017; Hossain et al., 2018). As a result, the increased citizen engagement is expected to help achieve a better communication and mutual understanding between governments and citizens, which help design and deliver better services for local and international communities, such as education (Kool & Bekkers, 2016), energy use (Jetzek et al., 2014), immigrant services (Henkel et al., 2016), international aids (Linders, 2013) and other fields. However, some research studies also indicated that mismatch between the citizens’ demands, and the data provided, and the diversity in citizens’ needs all limit the realization of social value and confine it to the stage of awareness improvement (Craveiro et al., 2016; Kool & Bekkers, 2016).

From this summary of OGD value and impact, we can see that opening government data is expected to have an impact on governments, other community organizations and general citizens. For governments, OGD development can help motivate governments to improve their transparency and accountability. The OGD itself, is also a potential asset that can be used to help governments become more effective and efficient in their daily operation. Besides political impact, the economic impact of OGD has been discussed the most since the very beginning of OGD development. It is expected that OGD will bring potential opportunities for innovation and economic benefits for different actors in the community: increased OGD based start-ups, established companies using OGD to explore new products and services, etc. Great revenue from making government data open to the public and reuse by the public have motivated many different countries to implement their OGD initiatives. While chasing the economic impact of OGD, some research also discussed the social impact of OGD and emphasized its potential to
facilitate citizen engagement in discussion and policy making that helps community development and improve the quality of life. However, despite the potential impact and value of OGD, there is a lack of evidence that OGD has revealed its true potential and it is difficult to measure the actual impact and value that OGD has already generated.

2.5 Discussion

This literature review provides a comprehensive summary of current research related to OGD utilization (see Figure 9). The descriptive analysis shows that with the development of OGD, increasing numbers of academic research studies have been conducted on this topic. Most research studies were conducted by researchers from developed countries. However, more and more research studies are now conducted in the context of developing countries (González-Zapata & Heeks, 2015; Saxena, 2017; Saxena & Janssen, 2017; Saxena, S., & Muhammad, 2018; Talukder et al, 2018). Besides theoretical research, more and more empirical studies were conducted to discuss issues related to OGD utilization. Survey and interview methods were the most common methods to collect feedback from the stakeholders.

The thematic analysis shows that the UTAUT and the TAM are the most often mentioned theoretical frameworks that were used to explore the influential factors on OGD utilization from the stakeholders’ perspectives. People have utilized OGD for information extraction and analysis, designing interfaces for better access and for understanding datasets, and creation of new business and public services for economic development and life quality improvement.
Figure 9. Framework of current research on OGD utilization
During the process of OGD utilization, the roles of stakeholders varied based on the contexts and types of utilization. OGD infomediaries are often the most active ones among all OGD users to help advocate and facilitate the OGD utilization. Influential factors of OGD utilization were discussed in most of the previous research studies. All the identified influential factors can be categorized into 1) government related, such as law, regulations and policy, management and leadership, etc., 2) OGD related, such as data accessibility, quantity and quality, etc., 3) user related, such as awareness, perceived usefulness and ease of use, technological skills, etc., and 4) community related, such as training, communication among OGD stakeholders etc. Issues about OGD management, data quality, users’ perceived usefulness and effort of use, and the communication among stakeholders are the most mentioned factors that affect the OGD utilization. The political, social and economic value and impact of OGD have also been discussed by some researchers but mainly at the theoretical level. Following are some summarized limitations of previous research and directions for future study.

First, different stakeholders use OGD differently and play different roles in the process of utilization. It is not clear within the current research about the links between different stakeholders and different types of OGD utilization. Is there any type of OGD utilization more common among one type of stakeholder than the others? In the meantime, very limited study has focused on a specific case of OGD utilization and investigated the roles and interaction among different stakeholders in the process of OGD utilization. More empirical research about the process of OGD utilization is needed to better understand the dynamics of OGD utilization and the roles and interaction of OGD stakeholders in that process.

Second, within the research about influential factors of OGD utilization, theories and models were mostly adopted to explore the factors that affect the users’ intentions of OGD use
instead of users’ actual use of OGD. Even though the concept of intention to use has been widely applied and empirically confirmed as a useful variable for the use of technology, more research studies need to be done to better understand the link between intention to use and the actual use. In the meantime, most models and frameworks only focused on assessing the user related factors, such as user perceived usefulness, ease of use, etc. Little empirical study has been conducted on examining the relationship between other types of influential factors (such as government related factors, OGD related factors and community related factors) and OGD utilization and impact generation. We have little knowledge about which are the most influential factors on OGD utilization and impact generation. More research studies are needed to explore this relationship to help guide stakeholders to improve their OGD practice.

Third, the discussion of OGD impact remains more at the theoretical level rather than based on empirical analysis. Most of the studies indicate that through utilization, OGD are expected to generate different impacts, however, there is limited evidence that indicates OGD has generated an impact. The measurement of OGD impact remains difficult. About the process of OGD impact generation, we have a very limited knowledge about the relationship between different influential factors and the generation of different GOD impacts. More empirical studies are needed to better understand the OGD impact generation, the roles of different influential factors and the interaction among different stakeholders in this process. More insights on this can be used to better guide OGD utilization and help maximize the benefits of OGD.

Faced with these limitations of previous research, this dissertation was designed to focus on 1) assessing influential factors, not limited to user related ones, but also including government related, OGD related and community related factors, and examine their influence on different types of OGD impact across countries over time; 2) investigating a specific case of OGD
utilization and exploring the interaction among stakeholders and identifying their perceptions of the facilitators, obstacles, and benefits during the process of OGD utilization.
CHAPTER THREE

3 METHODOLOGY

This section explains the methodology of this dissertation by describing the overall rationale (section 3.1) and a summary of two different research methods: a quantitative data analysis using secondary data and a qualitative case study through interviews (section 3.2).

3.1 Rationale

Among the reviewed previous research studies, one third of them specifically focused on factors that affect OGD utilization and in different stages, including publishing, accessing, analyzing and re-production and OGD impact generation (Sangiambut & Sieber, 2017). Research studies were conducted focusing on the OGD adoption in different organizational environments (public sector, private sector, academic community, etc.), and in different countries (developed and developing countries) (Lassinantti et al., 2019). All the identified influential factors can be categorized into 1) government related, such as OGD legislation and policy, privacy and security, resources, management and leadership, etc., 2) OGD related, such as metadata, data accessibility, quantity and quality, etc., 3) user related, such as awareness, perceived usefulness and ease of use, technological skills, etc., and 4) community related, such as coordination and communication among stakeholder, training, etc. (Maccani, 2016 Ma & Lam, 2019; Jetzek et al., 2019). Most previous quantitative studies only focused on assessing the relationship between user related factors and the intention of OGD use at the individual level. There is a lack of quantitative research studies that analyze the relationship between other types of influential factors and the actual OGD utilization and impact generation (Safarov et al., 2017). Little is known about whether different factors play different roles in different contexts, and among these various influential factors, which factors play a more important role in generating a certain type
of OGD impact than the others. In this dissertation, a quantitative study was conducted to address this gap by investigating: RQ1: What are the most influential factors in the generation of different types of OGD impact among different countries over time?

The quantitative study was designed to focus on assessing the relationship between influential factors and OGD impact, however, there is a “black box” between the influential factors and the OGD impact that needs future explanation, which is the process of the OGD utilization (see Figure 10). Few empirical studies have been conducted to focus on a specific example of OGD utilization and to explore the roles and interaction among stakeholders, their methods of utilizing OGD, the facilitators and obstacles they have encountered and what impact has been yielded through their OGD utilization. In this dissertation, besides the quantitative study, a qualitative case study was conducted to explore the process of OGD utilization focusing on the following two research questions: RQ2: How do stakeholders interact with each other in the process of OGD utilization? RQ3: How do different stakeholders perceive the facilitators and obstacles in the process of OGD utilization?
Figure 10. Framework of research rationale
3.2 Research Methods

Multi-method approaches, namely the use of multiple methods (usually quantitative and qualitative) in conducting research, are recommended to use to address complex social phenomena (Creswell, 2003). OGD utilization and impact generation is a complex process where different types of factors and stakeholders are involved and play a role. In order to better understand this process and the influence of these factors and stakeholders, this dissertation has adopted a multi-method approach with two studies conducted simultaneously but focusing on different aspects related to OGD utilization and impact generation: a quantitative study was designed to answer the question about “what” by assessing what are the most influential factors that affect the generation of OGD impact (RQ1) at a country level, whereas a qualitative study was conducted to dive into a specific case of OGD utilization in a specific country. It was designed to answer the question about “how” by exploring how these factors affect the process of OGD utilization and OGD impact generation. This study mainly focused on the interaction among stakeholders (RQ2) and their perceptions of the facilitators, obstacles, and benefits in the process of OGD utilization (RQ3). This multi-method approach allows us to in some extent validate interpretations of triangulate results that are generated through using different research methods. It helps address the weakness of each individual method, enables us to ask a broader set of questions while considering different aspects of the research topics, and discovers new or paradoxical factors that could inspire future research (Gil-Garcia & Pardo, 2006).

In the quantitative study, multiple regression using secondary data is used to assess the relationship between influential factors and OGD impact among different countries over time. Multiple regression is an analytical tool that helps to examine the relationship between a single outcome (dependent variable) and several predictors (independent variables) (Cohen et al.,
2003). Further, it helps to test a theory about presumed causal influences on a dependent variable. In our case, four types of influential factors, namely government related factors, OGD related factors, user related factors, and community related factors were mentioned by previous research and were expected to affect OGD utilization and lead to the generation of different types of OGD impacts. Through multiple regression, in addition to identifying which factors are statistically significant (P-value is smaller than .05), I also explored which factor had a bigger influence on the generation of OGD impact (based on the size of the standardized coefficients).

Simultaneously, a qualitative study was conducted focusing on a specific case: the NYS GIS Web Service. Through semi-structured interviews, I explored the stakeholders’ interaction in the process of producing and using the NYS GIS Web Services and discussed their perceptions of the facilitators, obstacles and benefits during their interaction with the web services. A qualitative case study provides empirical evidence that helps to explain and refine theory (Ospina et al., 2018; Scapens, 1990). Through interviews, it also allows us to obtain detailed information from the stakeholders and study the research question in depth while leaving room for new findings that can form the basis for concrete hypotheses to be tested in future research (Yin, 2013; Marshall & Rossman, 2011). This is especially useful when there is little existing research on the topic (Yin, 2013), as is the case here.

In sum, this dissertation has adopted a multi-method approach with a triangulation design where a quantitative study and a qualitative study were conducted simultaneously. The findings of these two studies were discussed individually and comparatively to help us better understand the OGD utilization and impact generation. More specifically, in this case, a quantitative study

---

was designed to help better understand the influence of different factors on the generation of different types of OGD impact. Whereas a qualitative study added more evidence to explain how different factors affect the OGD utilization and ultimately affect the generation of the OGD impact from the perspectives of different OGD stakeholders. And by exploring the interaction among OGD stakeholders during the process of OGD utilization, the qualitative study also revealed the important role of OGD stakeholders in OGD utilization and impact generation. The findings of these two studies together help us better understand what factors and how these factors affect the OGD utilization and the generation of OGD impact.
CHAPTER FOUR

4 QUANTITATIVE STUDY

By reviewing the literature, I summarized the influential factors of OGD utilization and impact generation and classified them into four categories, namely government related factors, OGD related factors, users related factors, and community related factors (see Table 5). In this quantitative study, I used secondary data to assess some factors from each category and their influence on the generation of different types of OGD impact. Due to the limitation of data availability, this research is not able to test all the factors that are mentioned by previous research. However, from each category, at least one factor was included in the model to be analyzed, and these factors have been mentioned frequently in previous research. Following is the illustration of research design in detail, including hypothesis development (section 4.1), research model (section 4.2), data collection and measurement development (section 4.3), analysis results (section 4.4), and discussion and implications (section 4.5).

4.1 Hypotheses Development

In this quantitative data analysis, the dependent variables are three different types of OGD impact, namely, OGD political impact, OGD economic impact and OGD social impact. The independent variables are four types of influential factors, namely government related factors (government politics and government actions), OGD related factors (OGD properties), user related factors (user participation) and community related factors (OGD training). I also included three control variables that capture the diversity of different countries in terms of their political governance, economic development and social progress. Multiple hypotheses were raised to test the relationship between these different factors and different types of OGD impact. The conceptual definitions and operational measurements of dependent variables, independent
variables, and control variables, and hypotheses that need to be tested are illustrated in detail in the following sections.

4.1.1 OGD political impact

As one of the main goals of OGD initiatives, governments have been working on improving transparency and accountability by making public data and information available to the public. OGD provides citizens with easy access to government data, which helps them better understand the government work. In the meantime, it enables citizens to monitor government performance (Linders, 2013). To a certain extent, it motivates governments to increase their efficiency and effectiveness in delivering public services to better satisfy the community needs (Mergel et al., 2018). In this research, OGD political impact refers to 1) increased transparency and accountability, and 2) increased government efficiency and effectiveness resulting from OGD development.

4.1.2 OGD economic impact

Nowadays, economic development is increasingly data driven. OGD has been adopted as important business resources (Gurin, 2014). Impact on economic development is expected to be obtained through OGD utilization, including aggregation, search, comparison, information extension, delivery, personalization, and translation (Henkel et al., 2017). The availability of OGD boosts OGD-based entrepreneurship and innovation in different types of businesses (Mergel et al., 2018). According to Open Data 500 U.S. survey⁷ that investigated the OGD utilization among companies from fifteen different sectors in the United States, OGD has been proved to help users obtain knowledge for decision making, problem solving, cost saving, cost saving,

---

⁷ https://www.opendata500.com/us/
efficiency improvements, design new business model and other aspects related to economic development (Bílková et al., 2015; Henkel et al., 2017; Jetzek et al., 2014). In this research, OGD economic impact refers to 1) increased new business, and 2) economic development in general resulting from OGD development.

4.1.3 OGD social impact
Open data development values citizen engagement. It focuses on citizen empowerment and encourages citizens to better participate in policy discussion and decision making (Hossain et al., 2018). As a result, increased citizen participation is expected to lead to a better mutual understanding between government and citizens and potentially better decision-making to address local and international communities issues in different fields, such as education, energy use, urban planning, and in long term to help achieve the sustainable development of the society (Jetzek et al., 2019; Meng et al., 2019). In this research, OGD social impact refers to 1) improving environment sustainability and 2) increasing the inclusion of marginalized groups in policy making and accessing government services.

4.1.4 Government policies
As the biggest OGD provider, governments have experienced many organizational and cultural challenges in the process of opening up data. In order to make the OGD development sustainable, related policy and standards are needed to guide the process and guarantee the data quantity and quality (Altayar, 2018; Kučera, 2017). For example, concerns about privacy and information security are one of the most mentioned factors that negatively affect government’s willingness to open data (Gascó-Hernández, et al., 2018; Saxena & Muhammad, 2018). The well-defined OGD policies and regulations improve the legitimacy of OGD development and provide guidance for the implementation of OGD initiatives across government departments in a
right and effective way (Susha, et al, 2015c; Attard et al., 2016; Altayar, 2018). Therefore, we hypothesize that well-defined OGD related government policies and regulations have a positive influence on the generation of different OGD impact.

H1a: Well-defined OGD related government policies and regulations have a positive influence on the generation of OGD political impact.

H1b: Well-defined OGD related government policies and regulations have a positive influence on the generation of OGD economic impact.

H1c: Well-defined OGD related government policies and regulations have a positive influence on the generation of OGD social impact.

4.1.5 Government actions

OGD development has forced governments to take actions to change their traditional ways of managing data and information from internal use only to make it available to the public (Wang & Lo, 2016). This transformation not only takes place at the national level, but also in the city or regional governments. Governments need to devote a great amount of budget, staff time, technology and other resources to support OGD implementation at different levels of government (Attard et al., 2016; Beno, et al., 2017, Ruijer et al., 2017a). During the process of OGD provision, various government agencies are involved in providing data and making it available to the public. Coordination among different agencies and departments, designing publication guidelines and data standard, etc., all these OGD management and publication related government actions help guide the workflow of OGD provision and guarantee the data quantity and quality (Kučera, 2017). Beyond data provision, substantial continuous commitment from governments to support and empower OGD based innovation also greatly increase the utilization of OGD. For example, governments provide funding and support to OGD boot camps,
OGD competitions, OGD hackathons, etc., through which it makes more people aware of the existence of OGD and encourages them to use OGD. These different actions help governments form strong operational capabilities that support OGD provision, utilization and sustainable development. Therefore, we hypothesize that well-resourced and supportive government actions, in the form of well-resourced OGD initiatives, consistent OGD management and publication practices, and support to OGD based innovation, have a positive influence on the generation of different OGD impact.

**H2a:** Well-resourced and supportive government actions have a positive influence on the generation of OGD political impact.

**H2b:** Well-resourced and supportive government actions have a positive influence on the generation of OGD economic impact.

**H2c:** Well-resourced and supportive government actions have a positive influence on the generation of OGD social impact.

4.1.6 OGD properties

The OGD properties refer to data openness and data liquidity, which contain different elements that affect OGD utilization and impact generation (Linders, 2013; Jetzek et al., 2014; Zuiderwijk et al., 2016; Mergel et al., 2018). Even though OGD development has started since 2009, there are still many complaints from the data users about the data openness and usability, such as the lack of OGD that suits users’ needs, difficulties in accessing datasets, non-machine-readable data formats, missing metadata, and inaccuracy of the data (Beno, et al., 2017; Magalhães & Roseira, 2016; Smith & Sandberg, 2018). These issues related to the OGD properties make it difficult for users to effectively use and conduct innovation based on OGD (Beno, et al., 2017; Mazzetti et al., 2015). Also, from previous research studies, we have noticed that different government
departments have opened up government data in different subject domains to the general public, such as crime data, transportation data, geographic data, etc. Different users have attempted to use them in different ways to achieve different goals (Susha et al., 2015b). However, not all OGD is equally valuable for users in achieving these goals (Gurin, 2014). Therefore, we hypothesize that the high-quality OGD in different subject domains will have a positive influence on the generation of different types of OGD impact.

H3a: High-quality OGD related to accountability has a positive influence on the generation of OGD political impact.

H3b: High-quality OGD related to innovation has a positive influence on the generation of OGD economic impact.

H3c: High-quality OGD related to social policy has a positive influence on the generation of OGD social impact.

4.1.7 User participation

As indicated in previous research, while efforts have sought to increase general public’s awareness about the potential of OGD, there are still many people who view OGD as unrelated to their daily life and have no desire to utilize it (Anneke & Martijn, 2016; Dechman & Syms, 2014; Hivon & Titah, 2017). In the meantime, concerns about data quality also hinder the utilization of OGD (Ruijer et al., 2017a; 2017b; Saxena, 2018a). Despite the potential significant impact of OGD, the lack of OGD utilization has made it far from living up to its full potential (Jetzek, et al., 2014; Zuiderwijk & Janssen, 2015; Smith et al., 2016). Therefore, we hypothesize that user participation in OGD utilization has a positive influence on the generation of OGD impact.
H4a: User participation in OGD utilization has a positive influence on the generation of OGD political impact.

H4b: User participation in OGD utilization has a positive influence on the generation of OGD economic impact.

H4c: User participation in OGD utilization has a positive influence on the generation of OGD social impact.

4.1.8 OGD training

Limited user participation and engagement in OGD utilization impede the generation of OGD impact (Jetzek et al., 2014; Smith et al., 2016). Technical skills, creative capabilities and data literacy are the necessities of OGD utilization (Morelli, et al., 2017; Susha et al, 2015a; 2015b). The lack of technical skills and OGD knowledge are mentioned by users as one of the main barriers for them to engage in OGD utilization (Brugger et al., 2016; Kučera, 2017). Tutorials and training are mentioned as methods, which not only can help improve users’ awareness, but also improve their skills and capabilities in OGD utilization (Susha et al., 2015a; Gascó-Hernández et al., 2018). Therefore, we hypothesize that the availability of OGD training has a positive influence on the generation of OGD impact.

H5a: The availability of OGD training has a positive influence on the generation of OGD political impact.

H5b: The availability of OGD training has a positive influence on the generation of OGD economic impact.

H5c: The availability of OGD training has a positive influence on the generation of OGD social impact.
4.1.9 Control variables: diversity of countries’ development

By reviewing the OGD development all around the world, it is clear that among different countries, the differences in governance lead to the diversity in OGD initiatives design and implementation (Styrin et al., 2017). The differences in economic development make the countries to be at different stages in OGD development (Susha et al., 2015a; Attard et al., 2016). The differences in the development of the society, including quality of life, citizen education, etc. lead to the differences in users’ awareness of OGD, engagement with OGD and the effectiveness of their OGD utilization (Dechman & Syms, 2014; Smith & Sandberg, 2018; Wirtz et al., 2018).

In general, the developed countries not only start the OGD development earlier, but also perform better compared to the developing countries (González-Zapata & Heeks, 2015; Saxena & Janssen, 2017). In this research, three control variables were included to capture the diversity of development in each country: 1) political governance, which consists of the traditions and institutions by which authority in a country is exercised 2) economic development, which is the creation of wealth from which community benefits are realized, and 3) social progress, which is a comprehensive measure of real quality of life that is independent of economic indicators.

4.2 Research Model

Figure 11, 12 and 13 present three research models that were used in this quantitative study.

They aim to investigate the influence of government actions, government policies, properties of OGD related to accountability, innovation and social policy, user participation and OGD training on the generation of OGD political, economic and social impact.
Figure 11. Research model: OGD political impact

Figure 12. Research model: OGD economic impact

Figure 13. Research model: OGD social impact

4.3 Research Method

4.3.1 Research design

In this quantitative study, three multiple regression analyses (see Equations 1, 2 and 3) were conducted using secondary data to assess the relationship between government policies ($X_1$), government actions ($X_2$), properties of OGD related to accountability ($X_{3.1}$), innovation ($X_{3.2}$), properties of OGD related to innovation ($X_{3.3}$), user participation ($X_4$), and OGD training ($X_5$) with OGD political impact ($Y_1$), OGD economic impact ($Y_2$), and OGD social impact ($Y_3$).
and social policy ($X_{3.3}$), user participation ($X_{4}$), OGD training ($X_{5}$) and different types of OGD impact, including OGD political impact ($Y_{1}$), OGD economic impact ($Y_{2}$) and OGD social impact ($Y_{3}$) while controlling for the diversity of each country in terms of their political governance ($X_{6}$), economic development ($X_{7}$), and social progress ($X_{8}$). Analysis was conducted using Stata.

$$Y_{1} = a_{1} + \beta_{1.1}X_{1} + \beta_{1.2}X_{2} + \beta_{1.3}X_{3.1} + \beta_{1.4}X_{4} + \beta_{1.5}X_{5} + \beta_{1.6}X_{6} + \beta_{1.7}X_{7} + \beta_{1.8}X_{8} + \epsilon_{1}$$

(Equation 1)

$$Y_{2} = a_{2} + \beta_{2.1}X_{1} + \beta_{2.2}X_{2} + \beta_{2.3}X_{3.2} + \beta_{2.4}X_{4} + \beta_{2.5}X_{5} + \beta_{2.6}X_{6} + \beta_{2.7}X_{7} + \beta_{2.8}X_{8} + \epsilon_{2}$$

(Equation 2)

$$Y_{3} = a_{3} + \beta_{3.1}X_{1} + \beta_{3.2}X_{2} + \beta_{3.3}X_{3.3} + \beta_{3.4}X_{4} + \beta_{3.5}X_{5} + \beta_{3.6}X_{6} + \beta_{3.7}X_{7} + \beta_{3.8}X_{8} + \epsilon_{3}$$

(Equation 3)

4.3.2 Data collection

This study has adopted secondary data from multiple resources to test the hypotheses mentioned in section 4.1.4 to 4.1.8. The secondary data includes the 2015-2017 Open Data Barometer (ODB) survey data, which were used as indicators of the independent variables and dependent variables, and the 2015-2017 Worldwide Governance Indicators\(^8\), Gross Domestic Product (GDP) per capita\(^9\) in 2015-2017 from the World Bank, and the Social Progress Index\(^10\) in 2015-2017 were used as indicators for the control variables.

The ODB was produced by the World Wide Web Foundation, and it is a global measure of the open data performance of governments in different countries in terms of their readiness for open data initiatives, implementation of open data programs, and the impact that open data is having on politics, business and civil society. The ODB data used in this research came from

---

\(^8\) [https://info.worldbank.org/governance/wgi/](https://info.worldbank.org/governance/wgi/)


\(^10\) [https://www.socialprogress.org/](https://www.socialprogress.org/)
peer reviewed expert surveys, where trained country specialists were asked to first, respond to questions about the open data development in a specific country following the research handbook indications\(^1\). Each question needed a response on a 0 - 10 scale, with detailed scoring guidance and thresholds provided. Researchers also provided detailed justifications and citations for all scores, for example, when assessing the freedom of information laws, the expert gave the score 5 with providing additional references as justification for the score (see Table 6\(^2\)). Second, researchers were asked to conduct detailed data assessment on 15 kinds of open government datasets in each country using 10 checklist questions with respect to the properties of the datasets. An aggregation logic and weightings were applied to the results of the checklist questions (see Table 9, IV3 for details) to generate a total score of each kind of dataset between 0 and 100. Experts also provided supportive evidence for the scores they gave to each dataset that they assessed (see Appendix A2 for examples). The responses were peer-reviewed, re-scored by researchers, and cross-checked by the research coordination and quality assurance team of ODB.

Table 6 An Example of Measurement of Indicator of the Independent Variable

<table>
<thead>
<tr>
<th>Indicators</th>
<th>United States (2017)</th>
</tr>
</thead>
</table>
| To what extent does the country have a functioning Right to Information (RTI) / Freedom of Information (FOI)? | Score: 5  
Explanation by experts: 
The Freedom of Information Act (1) states that any person “has the right to request access to federal agency records or information except to the extent the records are protected from disclosure by any of nine exemptions contained in the law or by one of three special law enforcement record exclusions” (2). Its various exemptions are listed on the US Department of State website (2). More information on submitting a request, and FOIA appeals can be found on the Department of Justice website (3). Users can also look at detailed data on each agency’s FOIA activities on FOIA.gov (4). In 2015, the Obama administration has been noted to withhold a large number of FOI requests (5). In addition, the Obama administration sets |


\(^2\) More examples can be found in Appendix A1.
Information (FoI) law? record for unfulfilled FOIA requests, as the “U.S. government in 2015 came up empty-handed when asked for public records more than 100,000 times, The Associated Press disclosed.” (6). In 2015, the White House announced it was exempting itself from records requests (7). Since 2009, the US government has confronted whistleblowers and official leakers. There have been eight prosecutions under the 1917 Espionage Act, more than double those under all previous US presidents combined (8).

## Sources
(2): http://foia.state.gov/Learn/FOIA.aspx
(3): http://www.justice.gov/eoir/foia-facts

The sample represents 112 countries in different regions of the world (see Table 7). The ODB survey data used in this research captured subject experts’ evaluation of OGD development in these countries from 2015-2017. It includes data from 92 countries in 2015, 115 countries in 2016 and 30 countries in 2017. I combined these data from 2015 to 2017 and obtained an effective sample size of 230 for the multiple regression analysis after excluding the missing data.

Table 7 The Number of Countries by Regions

<table>
<thead>
<tr>
<th>Region</th>
<th>Number of Countries (N=112)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Europe</td>
<td>35</td>
</tr>
<tr>
<td>Asia</td>
<td>27</td>
</tr>
</tbody>
</table>
4.3.3 Measurement development

The independent variable *government policies* is operationalized as a construct consisting of three equally weighted indicators: the availability of a functioning right-to-information law, the availability of a well-defined OGD policy/strategy and the availability of privacy and security policies. These three indicators come from the ODB expert survey and are measured using a scale from 0 to 10.

The independent variable *government actions* is operationalized as a construct consisting of four equally weighted individual indicators: supportive OGD resources, OGD implementation, OGD management, and provision of support to OGD users. The supportive resource is measured by the availability of well-resourced OGD initiatives in the country. OGD implementation is measured by the availability of OGD initiatives at different levels of governments (e.g., regional, county, city, town, etc.). The availability of a consistent OGD management and publication approach is used to measure OGD management. Provision of support to OGD users is measured by the availability of competitions, grants, and other support provided by governments to directly support OGD innovation. These four indicators come from the ODB expert survey and are measured using a scale from 0 to 10.

The independent variable *properties of OGD* is operationalized as a construct consisting of several individual indicators that measure the openness and liquidity of 15 different kinds of OGD. The openness is measured by the availability, the affordability and the shareability of the dataset. The liquidity is measured by the usability (data format and currency), the discoverability and the accessibility of the dataset. In this study, we used the OGD assessment from ODB as the
measurement of each indicator (see Appendix A2 for examples). Based on the common ways in which these datasets are used, these 15 kinds of OGD were grouped into three clusters: OGD related to accountability, OGD related to innovation and OGD related to social policy by the ODB researchers (see Table 8). I have adopted this categorization in this research. The property of each dataset is measured as the total value of these ten indicators of that dataset. The average scores of OGD datasets in each cluster are used as the value of the properties of OGD in that cluster. For example, in our case, the value of OGD related accountability is measured as the average properties of Land Ownership Data, Legislation, National Election Results, Detailed Government Budget, Detailed Government Spend, and Company Register.

Table 8 Cluster of OGD

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Definition</th>
<th>Datasets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accountability</td>
<td>Data central to holding governments and corporations to account.</td>
<td>Land Ownership Data, Legislation, National Election Results, Detailed Government Budget, Detailed Government Spend, Company Register.</td>
</tr>
<tr>
<td>Innovation</td>
<td>Data commonly used in open data applications by entrepreneurs, or with significant value to enterprise.</td>
<td>Map Data, Public Transport Timetables, Crime Statistics, International Trade Data, Public Contracts.</td>
</tr>
<tr>
<td>Social Policy</td>
<td>Data useful in planning, delivering and critiquing social policies &amp; with the potential to support greater inclusion and empowerment.</td>
<td>Health Sector Performance, Primary or Secondary Education, Performance Data, National Environment Statistics, Detailed Census Data.</td>
</tr>
</tbody>
</table>

It is challenging to measure the actual OGD utilization by different users due to the difficulties in identifying the users and tracking the utilization. In this study, the independent variable *user participation* is measured using a proxy indicator, namely the extent to which the

---

13 The cluster was designed by the ODB researchers, based on a qualitative analysis of the common ways in which these categories of data are used.
civil society and information technology professionals engage with OGD, in the form of demanding OGD, participating in open data communities, interaction with open data specialists in civil society or government officials on OGD projects, issues and engagement campaigns. The independent variable OGD training is measured by the availability of training for individuals or businesses to increase their skills to use OGD. Training may be delivered through full-time, part-time or professional development courses, business incubator programs, short-term boot-camp training events. Training may cover topics like general web and data technologies, data science and visualizations, and some other more specific open data issues. These two indicators come from the ODB expert survey and are measured using a scale from 0 to 10.

For the dependent variables, government efficiency and effectiveness, and government transparency and accountability are used as indicators to measure the OGD political impact. The development of the economy and the development of new business are used as indicators to measure the OGD economic impact. The environmental sustainability and inclusion of marginalized groups in policy making and accessing government services are used as indicators to measure the OGD social impact. Measuring impact is notoriously difficult, the ODB used claims made in credible sources concerning possible impact of open data as proxy indicators of areas in which impact may be occurring. OGD experts from different countries were asked to look for credible claims made in academic and scientific publications, use cases, mainstream media and other accredited online sources which explicitly attribute certain impact to open data released by the governments, then they were asked to give a score from 0-10. Highest scores were only given to peer-reviewed studies showing impact and emphasized the importance of sources making a direct connection between open data and observed impact. For scores over 5, participants were asked to cite at least two separate examples in the given category. The average
scores of the indicators of each construct were calculated and used as the value of that construct (see Appendix A3 for an example).

For the control variables that we included, Worldwide Governance Indicators (WGI) in 2015-2017 are used to measure the political governance of each country, GDP per capita 2015-2017 are adopted from the World Bank as the measure of economic development of each country, and Social Progress Index (SPI) in 2015-2017 are used to measure the social development of each country. Table 9 provides a detailed measurement of the constructs included in the models and the origins of the data.

Table 9 Constructs and Data Sources

<table>
<thead>
<tr>
<th>Constructs</th>
<th>Measurements</th>
<th>Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Independent Variables</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2. To what extent is there a well-defined open data policy and/or strategy in the country? [0-10]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. To what extent is there a robust legal or regulatory framework for protection of personal data in the country? [0-10]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2. To what extent are city or regional governments running their own open data initiative? [0-10]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. To what extent is there a consistent open data management and publication approach? [0-10]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4. To what extent is the government directly supporting a culture of innovation with open data through competitions, grants or other support?</td>
<td></td>
</tr>
<tr>
<td>IV2: Government Actions</td>
<td>Data Availability (data openness) a - Does the data exist? (IF a = No THEN 0 ELSE 5)</td>
<td>ODB Expert Survey 2015-2017</td>
</tr>
<tr>
<td></td>
<td>Data Format (data liquidity-usability) b- Is it available online from government in any form? (IF a = No</td>
<td></td>
</tr>
</tbody>
</table>

14 https://opendatabarometer.org
<table>
<thead>
<tr>
<th>Constructs</th>
<th>Measurements</th>
<th>Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Accessibility</td>
<td>THEN 0 ELSE (IF b = Yes THEN 10 ELSE 0))</td>
<td>ODB Expert</td>
</tr>
<tr>
<td></td>
<td>c - Is the dataset provided in machine-readable formats? (IF b = No THEN 0 ELSE (IF c = Yes THEN 15 ELSE 0))</td>
<td>Survey 2015-2017</td>
</tr>
<tr>
<td>Data Affordability</td>
<td>d - Is the dataset provided in machine-readable formats? (IF b = No THEN 0 ELSE (IF c = Yes THEN 15 ELSE 0))</td>
<td>ODB Expert</td>
</tr>
<tr>
<td></td>
<td>e - Is the dataset available free of charge? (IF b = No THEN 0 ELSE (IF c = Yes THEN 15 ELSE 0))</td>
<td>Survey 2015-2017</td>
</tr>
<tr>
<td>Open license</td>
<td>f - Is the data openly licensed? (IF b = No THEN 0 ELSE (IF c = Yes THEN 15 ELSE 0))</td>
<td>ODB Expert</td>
</tr>
<tr>
<td>Data Currency</td>
<td>g - Is the dataset up to date? (IF b = No THEN 0 ELSE (IF c = Yes THEN 15 ELSE 0))</td>
<td>Survey 2015-2017</td>
</tr>
<tr>
<td></td>
<td>h - Is the dataset being kept regularly updated? (IF b = No THEN 0 ELSE (IF c = Yes THEN 15 ELSE 0))</td>
<td>ODB Expert</td>
</tr>
<tr>
<td>Data Discoverability</td>
<td>i - Was it easy to find information about this dataset? (IF b = No THEN 0 ELSE (IF c = Yes THEN 15 ELSE 0))</td>
<td>Survey 2015-2017</td>
</tr>
<tr>
<td></td>
<td>j - Are (linked) data URLs provided for key elements of the data? (IF b = No THEN 0 ELSE (IF c = Yes THEN 15 ELSE 0))</td>
<td>ODB Expert</td>
</tr>
<tr>
<td>IV4: User Participation</td>
<td>To what extent are civil society and information technology professionals engaging with the government regarding open data? [0-10]</td>
<td>Survey 2015-2017</td>
</tr>
<tr>
<td>IV5: OGD Training</td>
<td>To what extent is training available for individuals or businesses wishing to increase their skills or build businesses to use open data? [0-10]</td>
<td>ODB Expert</td>
</tr>
<tr>
<td>Dependent Variables</td>
<td>1. To what extent has open data had a noticeable impact on increasing government efficiency and effectiveness? [0-10]</td>
<td>ODB Expert Survey 2015-2017</td>
</tr>
<tr>
<td></td>
<td>2. To what extent has open data had a noticeable impact on increasing transparency and accountability in the country? [0-10]</td>
<td>ODB Expert Survey 2015-2017</td>
</tr>
</tbody>
</table>
### Constructs

<table>
<thead>
<tr>
<th>Constructs</th>
<th>Measurements</th>
<th>Sources</th>
</tr>
</thead>
</table>
| Economic Impact | 1. To what extent has open data had a noticeable positive impact on the economy? [0-10]  
2. To what extent are entrepreneurs successfully using open data to build new businesses in the country? [0-10] | ODB Expert Survey 2015-2017       |
| Social Impact   | 1. To what extent has open data had a noticeable impact on environmental sustainability in the country? [0-10]  
2. To what extent has open data had a noticeable impact on increasing the inclusion of marginalized groups in policy making and accessing government services? [0-10] | ODB Expert Survey 2015-2017       |

### Control Variables

| CV1: Political Governance (WGI) | Worldwide Governance Index 2015-2017: aggregated indicators based on six board dimensions of governance that measure 1) the process by which governments are selected, monitored and replaced, 2) the capacity of the government to effectively formulate and implement sound policies, and 3) the respect of citizens and the state for the institutions that govern economic and social interactions among them. | Natural Resource Governance Institute, the Brookings Institution, and the World Bank. |
| CV2: Economic Development (GDP per capita) | GDP per capita 2015-2017: Country’s gross domestic product divided by the total number of people. | World Bank |
| CV3: Social Progress (SPI) | The Social Progress Index is the only measurement tool to focus exclusively comprehensively and systematically on the noneconomic dimensions of social performance across the globe. The dimensions include 1) basic human needs, 2) foundations of wellbeing and 3) opportunity. | Social Progress Imperative |

### Results

#### 4.4 Descriptive analysis

Table 10 presents a summary of descriptive analysis of the independent and dependent variables. The low average value of OGD properties in different clusters indicates that even though OGD development has been going on for a long time, there are still many issues related to OGD quality. In the meantime, I also noticed that the value of different types of OGD impact is quite low. This value was given by OGD experts from different countries when they participated in the...
ODB survey. They were asked to look for credible claims made in academic and scientific publications, use cases, mainstream media and other accredited online sources which explicitly attribute certain impact to open data released by the government, then they were asked to give a score from 0-10. Highest scores were only given to peer-reviewed studies showing impact and emphasized the importance of sources making a direct connection between open data and observed impact. For scores over 5, participants were asked to cite at least two separate examples in the given category. These low scores indicate that they did not find examples that attribute certain impact to OGD, it aligns with the finding of previous research studies that there is a lack of empirical evidence of OGD impact and it is difficult to measure the actual impact and value generated by OGD (Zuiderwijk et al., 2018; Jetzek et al., 2019; Meng et al., 2019).

Table 10 Descriptive Analysis

<table>
<thead>
<tr>
<th>Variables</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Government Policies</td>
<td>0</td>
<td>9.00</td>
<td>4.43</td>
<td>2.33</td>
</tr>
<tr>
<td>Government Actions</td>
<td>0</td>
<td>9.25</td>
<td>3.49</td>
<td>2.61</td>
</tr>
<tr>
<td>Properties of OGD related to Accountability</td>
<td>5.00</td>
<td>94.17</td>
<td>29.28</td>
<td>19.40</td>
</tr>
<tr>
<td>Properties of OGD related to Innovation</td>
<td>6.00</td>
<td>95.00</td>
<td>38.86</td>
<td>24.450</td>
</tr>
<tr>
<td>Properties of OGD related to Social Policy</td>
<td>5.00</td>
<td>95.00</td>
<td>46.50</td>
<td>25.37</td>
</tr>
<tr>
<td>User Participation</td>
<td>0</td>
<td>10.00</td>
<td>4.64</td>
<td>2.66</td>
</tr>
<tr>
<td>OGD Training</td>
<td>0</td>
<td>10.00</td>
<td>4.25</td>
<td>1.96</td>
</tr>
<tr>
<td>OGD Political Impact</td>
<td>0</td>
<td>7.50</td>
<td>1.56</td>
<td>1.70</td>
</tr>
<tr>
<td>OGD Economic Impact</td>
<td>0</td>
<td>9.00</td>
<td>1.88</td>
<td>2.23</td>
</tr>
<tr>
<td>OGD Social Impact</td>
<td>0</td>
<td>6.00</td>
<td>.86</td>
<td>1.29</td>
</tr>
<tr>
<td>WGI</td>
<td>-1.85</td>
<td>1.86</td>
<td>.17</td>
<td>.88</td>
</tr>
<tr>
<td>GDP Per Capita ($)</td>
<td>315.78</td>
<td>82081.61</td>
<td>16166.23</td>
<td>18888.21</td>
</tr>
<tr>
<td>SPI</td>
<td>37.52</td>
<td>89.82</td>
<td>68.34</td>
<td>14.72</td>
</tr>
</tbody>
</table>

N= 230
4.4.2 Validity, reliability and assumptions

Before I conducted the regression analysis, the first step was to ascertain the reliability of the measurement for the constructs with more than one indicator. The measurement’s reliability is confirmed as the values of Cronbach’s alpha of all the constructs are greater than 0.7 (see Table 11). The measurement’s validity is assessed using face validity that the model is constructed based on previous research.

Table 11 Cronbach’s alpha

<table>
<thead>
<tr>
<th>Constructs</th>
<th>No. of Items</th>
<th>Cronbach’s alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>Government Policies</td>
<td>3</td>
<td>.822</td>
</tr>
<tr>
<td>Government Actions</td>
<td>4</td>
<td>.934</td>
</tr>
<tr>
<td>OGD Political Impact</td>
<td>2</td>
<td>.860</td>
</tr>
<tr>
<td>OGD Economic Impact</td>
<td>2</td>
<td>.855</td>
</tr>
<tr>
<td>OGD Social Impact</td>
<td>2</td>
<td>.744</td>
</tr>
</tbody>
</table>

The second step is to check the assumptions of multiple regression. Both independent variables and dependent variables are measured on a continuous scale, there are no significant outliers among the independent and dependent variables, the relationship between independent variables and dependent variables is linear, there is no multicollinearity in our data, and the error of the regression is normally distributed and independent. However, the assumption of homoscedasticity of errors is not met (see details in Appendix B). For the multiple regression analysis, I used the Stata command “robust” to get robust standard errors and p values\textsuperscript{15}.

4.4.3 Hypothesis testing

In this study, it was hypothesized that the availability of well-defined OGD related government policies and regulations, the availability of well-resourced and supportive government actions,

\textsuperscript{15} https://www.stata.com/manuals/rregress.pdf
high quality of OGD, extensive user participation and the availability of OGD related training all have a significant positive influence on the generation of OGD political, economic and social impact. Three multiple regression analyses were conducted to test these hypotheses using Stata and the results are presented in Table 12.

Table 12 The Results of Multiple Regression

<table>
<thead>
<tr>
<th>Variables</th>
<th>Political Impact</th>
<th>Economic Impact</th>
<th>Social Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>β</td>
<td>SE B</td>
<td>p</td>
</tr>
<tr>
<td>Government Policies</td>
<td>.207*</td>
<td>.063</td>
<td>.018</td>
</tr>
<tr>
<td>Government Actions</td>
<td>.409***</td>
<td>.065</td>
<td>.000</td>
</tr>
<tr>
<td>Properties of OGD related to</td>
<td>.294***</td>
<td>.006</td>
<td>.000</td>
</tr>
<tr>
<td>Accountability</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Properties of OGD related to</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Innovation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Properties of OGD related to Social Policy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>User Participation</td>
<td>.037</td>
<td>.046</td>
<td>.604</td>
</tr>
<tr>
<td>OGD Training</td>
<td>.138</td>
<td>.063</td>
<td>.057</td>
</tr>
<tr>
<td>WGI</td>
<td>-.074</td>
<td>.139</td>
<td>.307</td>
</tr>
<tr>
<td>GDP Per Capita ($)</td>
<td>-0.001</td>
<td>8.02e-06</td>
<td>.994</td>
</tr>
<tr>
<td>SPI</td>
<td>-.160*</td>
<td>.008</td>
<td>.021</td>
</tr>
<tr>
<td>R²</td>
<td>.708</td>
<td></td>
<td>.767</td>
</tr>
<tr>
<td>Adj R²</td>
<td>.697</td>
<td></td>
<td>.759</td>
</tr>
</tbody>
</table>

*p < .05, ** p < .01, *** p < .001

For the model of OGD political impact, the results show that 69.7% of the variance in the generation of OGD political impact can be accounted for by government policies, government actions, properties of OGD related to accountability, user participation and OGD training while controlling for the diversity of the countries, $F (8, 221) = 67.20, p < .001$. Looking at the unique individual contributions of the predictors, the results show that government policies ($β = .207, p = .018$), government actions ($β = .409, p < .001$), properties of OGD related to accountability
($\beta=.294, p<.001$) positively predict the generation of OGD political impact. Hypotheses 1a, 2a and 3a are supported by the data.

For the model of OGD economic impact, the results show that 75.9% of the variance in the generation of OGD economic impact can be accounted for by government policies, government actions, properties of OGD related to innovation, user participation and OGD training while controlling for the diversity of the countries, $F(8, 21) = 74.04, p < .001$. Looking at the unique individual contributions of the predictors, the results show that government actions ($\beta=.603, p<.001$) and OGD training ($\beta=.216, p=.002$) positively predict the generation of OGD economic impact. Hypotheses 2b and 5b are met.

For the model of OGD social impact, the results show that 56.7% of the variance in the generation of OGD economic impact can be accounted for by government policies, government actions, properties of OGD related to innovation, user participation and OGD training while controlling for the diversity of the countries, $F(8, 221) = 27.06, p < .001$. Looking at the unique individual contributions of the predictors, the results show that government actions ($\beta=.648, p<.001$) and user participation ($\beta=.157, p=.036$) positively predict the generation of OGD social impact. Hypotheses 2c and 4c are met.

Across the three models, the availability of well-resourced and supportive government actions consistently shows a significant positive impact on the generation of different kinds of OGD impact. In the political impact model, 1 unit increase in government actions that support OGD development is predicted to generate .409 units more of OGD political impact. In the economic impact model, 1 unit increase in government actions that support OGD development is predicted to generate .603 units more of OGD economic impact. In the social impact model, 1
unit increase in government actions that support OGD development is predicted to generate .648 units more of OGD social impact.

Besides the independent variables, I also noticed that control variable WGI has a significant positive influence on the generation of OGD social impact, GDP per capita has a significant positive influence on the generation of OGD economic impact. Both findings align what we expected based on previous research. However, unexpectedly, the control variable SPI shows a significant negative influence on the generation of OGD political, economic, and social impact.

4.5 Discussion

As indicated by previous research studies, OGD utilization and impact generation are influenced by various factors, which can be summarized as government related factors, OGD related factors, user related factors and community related factors (see Table 5). Through linear multiple regression analysis using secondary data, a quantitative study was conducted to test some of these factors from each category and their influence on the generation of different types of OGD impact with answering RQ1: what are the most influential factors that affect the generation of OGD impact among different countries over time? Based on the results of multiple regression, following hypotheses were supported (see Table 13).

Table 13 True Hypotheses

<table>
<thead>
<tr>
<th>For the model of OGD political impact:</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1a: Well-defined OGD related government policies and regulations have a positive influence on the generation of OGD political impact.</td>
</tr>
<tr>
<td>H2a: Well-resourced and supportive government actions have a positive influence on the generation of OGD political impact.</td>
</tr>
<tr>
<td>H3a: High-quality OGD related to accountability has a positive influence on the generation of OGD political impact.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>For the model of OGD economic impact:</th>
</tr>
</thead>
<tbody>
<tr>
<td>H2b: Well-resourced and supportive government actions have a positive influence on the generation of OGD economic impact.</td>
</tr>
</tbody>
</table>
H5b: The availability of OGD training has a positive influence on the generation of OGD economic impact.

For the model of OGD social impact:
H2c: Well-resourced and supportive government actions have a positive influence on the generation of OGD social impact.
H4c: User participation in OGD utilization has a positive influence on the generation of OGD social impact.

Among the results across the three models, some of them fully support the findings of previous research studies. For example, there is a consistent finding across these three models that well-resourced and supportive government action is the most important factor that is predicted to help generate different kinds of OGD impact. As often mentioned in previous research, resources, including budget, IT infrastructure, staff, etc. greatly affect the capacity and capability of government in the implementation of OGD initiatives (Beno, et al., 2017; Maccani, 2016; Ruijer et al., 2017a). During the implementation process, OGD management and standardized OGD publication methods are significant factors that help guide the provision of high quality of OGD for public use (Altayar, 2018; Kučera, 2017). Besides OGD provision, government efforts in supporting OGD utilization and innovation, such as government-led competitions, grants, etc. also help improve user’s awareness and engagement in OGD utilization and impact generation (Maccani, 2016; Khayyat & Bannister, 2017; Susha et al., 2015b; 2015c).

However, some of the factors only show significant influence on a specific type of OGD impact, which partially support the previous research studies. First, in our models, the availability of well-defined OGD related policies and regulations was only found to have a significant positive influence on the generation of OGD political impact, and neither its influence on the generation of OGD economic nor social impact was statistically significant. This result partially supports the findings of previous research that OGD policies and standards are needed to guide the process and guarantee the data quantity and quality (Altayar, 2018; Kučera, 2017).
And well-defined OGD policies and the legislation about privacy and information security is especially important. It helps alleviate people’s concern about privacy and information security, improves the legitimacy of OGD development and provides guidance for its implementation across government departments in a right and effective way, which in the end facilitate OGD impact generation (Susha, et al, 2015c; Attard et al., 2016; Altayar, 2018).

Second, in our models, the high quality of OGD related to accountability was only found to have a positive influence on the generation of OGD political impact. The influence of OGD related to innovation on the generation of OGD economic impact and the influence of OGD related to social policy on the generation of OGD social impact were not significant. It partially supports the argument in previous research that OGD properties, namely the availability, accessibility, usability, etc. of OGD greatly affect the OGD utilization and impact generation. (Jetzek, 2016; Jetzek et al., 2019). It also to some extent indicates that not all OGD is equally valuable for users in achieving OGD impact as indicated by Gurin (2014).

Third, in our models, user participation was only found to have a positive influence on the generation of OGD social impact, and neither its influence on the generation of OGD political nor economic impact was statistically significant. It partially indicates that OGD impact generation not only relies on the provision of OGD, but more importantly relies on OGD utilization. And in this case, the result indicates that the improvement of citizen participation and engagement is expected to generate more social impact. More specifically, the increased citizen engagement helps form a better communication and mutual understanding between governments and citizens, which helps better address community issues and achieve sustainable development (Attard et al., 2016; Craveiro et al., 2016; Gagliardi et al., 2017; Jamieson et al., 2019).
Fourth, in our models, OGD training was only found to have a significant positive impact on the generation of OGD economic impact, and neither its influence on the generation of OGD political nor social impact was statistically significant. It partially supports the findings of previous research studies that OGD tutorials and training not only can help improve users’ awareness, but also improve their skills and capabilities in OGD utilization and impact generation (Susha et al., 2015a; Gascó-Hernández et al., 2018).

Compared to previous research studies, this quantitative study also provided some new insights. Most previous quantitative studies only focused on assessing the relationship between user related factors and intention of OGD use at the individual level. Besides the user related factors, this quantitative study was able to assess the relationship between different types of influential factors and generation of different impact. The research models of this study were also designed in an innovative way to test the influence of properties of a specific type of OGD on a specific type of OGD impact.

In sum, the results of this quantitative study show that OGD impact generation are affected differently by different factors. Following are some practical implications based on this study.

First, the availability of well-resourced and supportive government action is significantly related to generating different types of OGD impact, in this case, political, economic and social impact of OGD. In OGD development, governments are required to take a series of actions and fulfill organizational and culture transformation, such as getting rid of traditional mentalities of holding the data only for internal use, increasing coordination among different agencies and departments, assigning specific staff managing OGD and so on. In order to generate the expected impact of OGD, it is important for governments at different levels to wisely consider the
allocation of their resources, including budget, IT infrastructure and staff time in implementing OGD initiatives. In the process of OGD provision, it is necessary to have a consistent management and publication approach to guide the process and guarantee the data quantity and quality. Besides OGD provision, governments should also devote efforts to increase the citizen engagement and make more OGD to be used by more users. For example, governments should also provide support in OGD based innovation to promote and advocate OGD utilization among general public.

Second, some influential factors are more important than others for the generation of a certain type of OGD impact. In OGD development, governments who are limited in resources should prioritize and focus on the more significant influential factors with considering the type of impact that they expect to generate. This research shows that compared to other factors, the availability of well-defined OGD related government policies and high-quality of OGD related to accountability are more important factors that contribute to increasing government transparency and accountability and improving the effectiveness and efficiency of government service.

The results also imply that OGD in different subject domains have different influences on different types of OGD impact. In this study, I purposely designed the model in a way to test the influence of OGD in a specific subject domain on the generation of a specific type of OGD impact. Among these three models, only OGD related to accountability was proven to have a significant influence on the OGD political impact, and OGD related to innovation and social policy were not proven to be significant influential factors for the OGD economic and social impact.

In order to facilitate economic development and entrepreneurship, the availability of OGD training is a more important factor that needs to be considered. Harvesting useful
information through massive OGD requires professional data analytic skills that the general public may not have. OGD training should be provided on different subjects like general web and data technologies, data science and visualizations, and other specific open data issues through different methods, such as professional development courses, business incubator programs or boot-camp training events. It helps improve users’ awareness, skills and capabilities in OGD utilization and the generation of economic impact.

User participation and engagement is a more important factor to help generate social impact. The increased user participation and engagement in OGD utilization provide more opportunities for the general public to be involved in civic discussion, which helps better address community issues and achieve sustainable development.
CHAPTER FIVE

5 QUALITATIVE STUDY

This chapter illustrates the conceptual framework that was used to guide this qualitative study (section 5.1), the specific research method (section 5.2), the context of the case (section 5.3), the findings of the case study (sections 5.4), and discussions and implications (section 5.5).

5.1 Conceptual Framework

Previous research has discussed different influential factors of OGD utilization (see Table 5), which can be categorized into 1) government related factors, such as law, regulations and policy, resources, management, etc., 2) OGD related factors, such as data availability, usability, accessibility, etc., 3) user related factors, such as awareness, perceived usefulness and ease of use, technological skills, etc., and 4) community related factors, such as training, etc. (Jetzek et al., 2019; Ma & Lam, 2019). In the process of OGD utilization, different stakeholders are involved, such as governments at different levels, non-profit organizations, for-profit organizations, developers, academic institutions, general citizens, etc. and are playing different roles like OGD providers, OGD infomediaries, and OGD users (Lassinantti et al., 2019; Sangiambut & Sieber, 2017). OGD is expected to generate various impacts, including political impact, economic impact and social impact (Alexopoulos et al., 2016; Attard, et al., 2016; Gurin, 2014).

This study was conducted focusing on the case of the New York State (NYS) GIS web services hosted on the NYS GIS Clearinghouse (http://gis.ny.gov/webservices/). The main goal is to better understand the interaction among stakeholders and their perceptions of the facilitators and obstacles in the process of their interaction with the NYS GIS web services. More specifically, who is providing the web services? Who is using the web services? What are they
using it for? How do stakeholders interact with each other in the process of producing and using the web services? How do stakeholders perceive the facilitators and obstacles in the process of their interaction with the web services? What are the benefits and limitations of the web services? Based on literature review and case background research, the following framework (see Figure 14) was designed to guide the data collection through semi-structured interviews and the analysis of this case study.

![Framework of case analysis](image)

*Figure 14. Framework of case analysis*
5.2 Data Collection

The analysis of the case study is based on semi-structured interviews with different stakeholders of the NYS GIS web services, an environmental scan of the NYS GIS Clearinghouse, and reviews of government documents, websites, online news, etc.

For the interview, I used a purposive sampling method to connect with all key stakeholders, including the producers and users of the NYS GIS web services from different fields, such as government agencies at different levels (state and local), nonprofit organizations, private organizations, and academic institutions. Interview invitations were sent out via emails to 1) GIS list server and 2) the contact list of the NYS GIS professionals (http://gis.ny.gov/outreach/whoswho/). I also used a snowball sampling method to connect with more stakeholders who are related to the NYS GIS web services based on the recommendation of the interviewees.

The interviews were conducted from February to December 2020 using different methods: face to face meeting, phone calls, and Zoom meetings due to the influence of the pandemic. The interviews usually lasted no more than 60 minutes. The interview protocol included questions related to 1) informant background information, 2) experience of the NYS GIS web service production and utilization, 3) facilitators, costs and challenges of the NYS GIS web service production and utilization, 4) benefits and limitations of the NYS GIS web service, and 5) plans for future development. Questions about stakeholders’ experience and their perceptions of the influential factors were drafted differently to better interact with different types of stakeholders, namely, the producers of the NYS GIS web service and the users of the NYS GIS web services. Prompts were designed based on findings of previous research to better guide the interviews. The interview protocol is attached in the Appendix C.
In the end, 36 interviews were conducted with stakeholders from different sectors (see Table 14). The interviews were recorded, transcribed and hand coded line by line by me to ensure consistency (Saldaña, 2016). I adopted a grounded theory approach to analyze the transcripts (Glaser and Strauss, 2017).

Table 14 The Type and Number of Stakeholder Interviewed

<table>
<thead>
<tr>
<th>Type of Organizations</th>
<th>Number of Interviews</th>
</tr>
</thead>
<tbody>
<tr>
<td>State government</td>
<td>8</td>
</tr>
<tr>
<td>Local government(^\text{16})</td>
<td>7</td>
</tr>
<tr>
<td>Nonprofit organizations</td>
<td>11</td>
</tr>
<tr>
<td>Private organizations</td>
<td>5</td>
</tr>
<tr>
<td>Academic institutions</td>
<td>5</td>
</tr>
<tr>
<td>Total</td>
<td>36</td>
</tr>
</tbody>
</table>

5.3 Case Description

The NYS Geographic Information System (GIS) Clearinghouse\(^\text{17}\), operated by the NYS GIS Program Office (NYS GPO), was established in the late 90s to disseminate information about New York's Statewide GIS Coordination Program and to provide access to the NYS GIS Data and metadata Repository. It serves as a central location where state agencies and local governments can list the GIS datasets that they have to allow GIS users to browse the data, to find out how to obtain the data, and in some cases directly download the data.

The NYS GIS Clearinghouse provides GIS data resources through its data catalog. Users interact with the catalog to view the lists of all data by owners’ name alphabetically, by sector and by dataset name. Users can also search dataset by owner, by theme, by county, and by sector. However, only some data are open to the public and are available for users to directly download them.

\(^{16}\) Mix of regional, county, city, and municipal governments.
\(^{17}\) http://gis.ny.gov/
In 2015, embarked on the NYS initiative “NY Responds” to build the ShareGIS environment in response to the need for a statewide GIS Platform, the NYS GPO ShareGIS\textsuperscript{18} team has been working on designing an enterprise GIS platform. This GIS platform aims to empower its users with both GIS data services and geoprocessing services through the ShareGIS service catalog.

The NYS GIS web services\textsuperscript{19}, as one of the open GIS data resources, provide users with a convenient access to GIS data via automated means (connecting to the NYS GIS servers through ArcGIS Rest EndPoint). Users can use data from servers directly with desktop GIS software without downloading data as files to their own computers, which makes it possible to always use up-to-date data easily. The NYS GIS web services are currently hosted by different state agencies and are mainly offered as web map service (WMS) and web feature service (WFS). WMS is produced by a map server from data provided by the GIS database. It provides georeferenced map images for users to view geographic information over the Internet. For example, the GPO has been publishing orthoimagery web map services, which are aerial photographs that are geometrically corrected (the scale is uniform). It allows users to directly access and view an accurate representation of the Earth’s surface in the NYS without downloading the data. Whereas WFS offers direct fine-grained access to geographic information at the feature and feature property level. It allows users to retrieve or modify the data they are seeking. For example, the NYS Thruway Authority published the statewide primary roadway route systems, interchanges, and interchange ramps as web feature services. It not only allows users to view the map, but also makes them be able to retrieve and modify the data behind the

\textsuperscript{18} http://gis.ny.gov/shareGIS/

\textsuperscript{19} http://gis.ny.gov/webservices/
maps. Table 15 provides a summary of the number and the types of the web services listed on the NYS GIS Clearinghouse.

Table 15 *The Number of The NYS GIS Web Services by Types and by Agencies (by 4/21/2021)*

<table>
<thead>
<tr>
<th>Agency</th>
<th>The Number of WMS</th>
<th>The Number of WFS</th>
<th>Other services</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adirondack Park Agency</td>
<td>5</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>Department of Environmental Conservation</td>
<td>2</td>
<td>14</td>
<td>0</td>
</tr>
<tr>
<td>Department of State</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Department of Transportation</td>
<td>8</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ITS GIS Program Office</td>
<td>49</td>
<td>7</td>
<td>3 (1 is image service)</td>
</tr>
<tr>
<td>Thruway Authority</td>
<td>0</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>74</td>
<td>36</td>
<td>3 Geocoding services</td>
</tr>
</tbody>
</table>

5.4 Results

The following sections illustrate in detail about the stakeholders related to the NYS GIS web services production and utilization (section 5.4.1), their motivations, experiences, interaction and coordination during the process of producing and utilizing the NYS GIS web services (section 5.4.2), the facilitators and obstacles they have encountered in the process of production and utilization the NYS GIS web services (section 5.4.3), as well as the benefits, limitations and the future development of the NYS GIS web services in their perspectives (section 5.4.4).

5.4.1 Stakeholders

During the case study, I have interviewed different stakeholders related to the NYS GIS web services. Based on the types of organizations they come from, these stakeholders can be categorized as state governments, local governments (governments of counties and towns), non-profit organizations, for-profit organizations and academic institutions. Based on their relationship with the NYS GIS web services, these stakeholders can be categorized as providers of the NYS GIS web services, who are mainly involved in the production process of the NYS
GIS web services, and users of the NYS GIS web services, who mainly use these web services in their work and research. There are also stakeholders that play a dual role, they are not only providers of the web services but also users of the web services (See Figure 15).

Figure 15. Stakeholders interviewed related to the utilization and production of the NYS GIS web services

5.4.1.1 State government agencies

In this case study, I have interviewed eight staff from different NYS government agencies, including one from the NYS GPO, one from the NYS Department of State (DOS), one from the NYS Thruway Authority, one from the NYS Tug Hill Commission, one from the NYS Department of Parks Recreation & Historic Preservation (DOP), one from the NYS Department of Health (DOH) and two from the NYS Department of Environmental Conservation (DEC). The NYS GPO is managing the NYS GIS Clearinghouse that provides free GIS framework data to the public via downloadable data and web services, a more efficient and flexible format of accessing and using open GIS data. The framework data includes orthoimagery, elevation, tax parcels, etc. All the interviewees from these state agencies mainly use these GIS web services for creating maps and conducting analysis for planning and research projects. Besides the NYS
GPO, the NYS DOS, NYS DEC and the NYS Thruway Authority are other state agencies that provide GIS web services in their subject domains through the NYS GIS Clearinghouse to the public. For example, the NYS DOS provides the Offshore Atlantic Ocean Study Boundaries web map service; the NYS DEC provides land, roads and trails web map services, and the NYS Thruway Authority provides thruway roads, ramps and interchanges web feature services.

5.4.1.2 Local governments

Seven GIS staff from local governments were interviewed in this case study. They come from five different counties (Washington, Oneida, Herkimer, Tompkins and Suffolk) and one town (Town of Huntington in Suffolk County). All of them have a certain level of GIS skills. They are frequent users of the NYS GIS Clearinghouse. Six of them have experience using the NYS GIS web services and mainly use it as references to answer public inquiries, and as base maps to create new maps for reports, web applications and discussion about community development. Only one of them mostly uses downloaded data rather than using web services due to the feature of their job tasks, which requires more extensive analysis. Local governments are not directly producing the NYS GIS web services, but they collect, update and share the GIS data (e.g., streets & addresses, tax parcel, etc.) in their governance areas with the state government agencies on a regular basis, and some of data they shared have been turned into web services.

5.4.1.3 Non-profit organizations

I have reached out to ten nonprofit organizations in different domains, including nature conservation, environmental advocacy and urban management and planning. They have greatly taken advantage of state and local government GIS resources in their daily work. Most of them utilize the orthoimagery web services as base maps for creating maps and web applications. Tax parcel data is another GIS resource that they often acquire from different county governments.
Three of them mainly use downloadable GIS data on the NYS GIS Clearinghouse and have limited or no experience using the NYS GIS web services. Besides being the users, three of them also share GIS data with the NYS GIS Clearinghouse. Through the NYS GIS Clearinghouse website, one organization directly make its data (e.g., Hudson River Estuary and Scenic Hudson SLAMM) available to the public, and the other two only provide general information about the datasets and contact information for public to acquire data from them.

5.4.1.4  For-profit organizations

Responses of five interviewees from four profit-making organizations were included in the analysis. They are companies in the field of environmental services, engineering and GIS consulting. All of them have experience directly downloading GIS data and using GIS web services from the NYS GIS Clearinghouse. Together with open GIS data from other sources, they mainly use it for creating maps and conducting analysis for clients.

5.4.1.5  Academic institutions

I also interviewed five researchers from four universities and one museum. Four of them have experience using the NYS GIS web services. Geocoding services and orthoimagery are the most mentioned NYS GIS web services. One of them only has experience using the downloadable GIS data and has no experience using the web services.

In summary, state governments are the main providers of the NYS GIS web services hosted on the NYS GIS Clearinghouse. Local governments also indirectly contribute to the production of the web services by updating and sharing related GIS data to the state. The users of the NYS GIS web services are very diverse, including users from state governments, local governments, non-profit organizations, for-profit organizations and academic institutions. They
mainly use the NYS GIS web services to extract information, create maps, import into web applications and conduct analysis and reports.

5.4.2 Production and utilization

As mentioned in the section above, some stakeholders are the producers of the NYS GIS web services, some are the users, and some are both producers and users. The following sections discuss in detail about their motivation, experience, interaction and coordination during the process of the production and utilization of the NYS GIS Web services.

5.4.2.1 Production of the NYS GIS web services

As one of the focuses of this study, I am interested in knowing who are the providers of these NYS GIS web services? and what motivates them to do so? From the interview, we get to know that in the past decades, the NYS GPO has devoted great efforts in consolidating, managing and publishing GIS data and making it open to the public. Besides the traditional downloadable GIS data, different state agencies have started to push out more web services through the NYS GIS Clearinghouse. The publication of web services is not only one way that OGD is provided but also can be viewed as a way of using OGD: data to data. These state agencies are the providers of the web services, as well as the infomediaries who have transformed the raw open GIS data into GIS web services, a format of open GIS data that provides users with a more convenient access to GIS resources and guarantees them to always access the most up-to-date data.

These government actions in open GIS data and web services provision are driven by internal and external motivations. Externally, by doing so, governments aim to empower the public to utilize these GIS resources for decision making and community development. Internally, they are expected to help remove redundancy in government work and improve their
efficiency in answering public inquiries. As indicated by interviewees from the NYS GPO and the NYS Thruway Authority:

We have a vision of no matter what activities you can spatially do in New York state, we want to make our data available in a way that scales....So, we make that data publicly available and when you have the synergy of all these fairly simple web service interactions and well documented ways for people to get at those, I think we can move the dial on the economy. I think we can make New York state a more competitive place to live and do business ... If we can make those things a tiny bit more efficient, then I can sleep at night knowing that we've done the right thing.

We put out as web services, that's basically to ensure that other entities who want to use the data can get to it quickly, easily, we don't have to get phone calls all the time and provide data and that we know that everybody can be using the same source, and is the most up-to-date source, it is way easier than having people call us all the time.

The NYS GIS Clearinghouse hosts different types of GIS web services produced and managed by different state government agencies (see Table 14). However, due to limited staff and extremely intensive workload, many updates and maintenance of the data used to build the web services are conducted by the local governments (in our case, the counties and towns). As an interviewee from the NYS GPO said:

I remember a bureau director, 20 years ago telling me, we will never have the resources to maintain 24,000 scale street mapping and continuously up to date. And I agree we don't have the resources, but we can have the partnerships to do it. So those 700,000 edits were made with the help of 100 different agencies that have a vast interest in helping us maintain that said to a value proposition large with local governments...

The NYS GPO has created a single web page listing all the available GIS web services for users to access. During the interview, the most mentioned NYS GIS web services are

20 http://gis.ny.gov/webservices/
orthoimagery, streets & addresses and tax parcels. They all come from programs managed by the NYS GPO.

First is the NYS Statewide Digital Orthoimagery Program. Orthoimagery combines the image characteristics of a photograph with the geometric qualities of a map. Orthoimagery web services is one of the products of the New York Statewide Digital Orthoimagery program. Users can access the most current and historical orthoimagery through the web services hosted on the Clearinghouse. The goal of the program is to produce high-resolution digital orthoimagery for NYS on an ongoing annual basis and update the imagery of the entire NYS every 4 to 5 years. The NYS GPO oversees the data collection, data processing, and quality control.

Second is the Street and Address Maintenance (SAM) Program. Streets & Addresses web services is one of the public available products of the SAM Program. The program maintains statewide authoritative street and address point databases (compliant with standards of the National Emergency Number Association) to support Next Generation 911 and other stakeholder and customer activities through extensive state and local government partnerships and the use of data maintenance tools. Some of the data edits are conducted by the counties and reported on a scheduled basis to the SAM Program managed by the NYS GPO. Some edits are conducted by the SAM team on behalf of the counties, and all the updates will be incorporated and assured by the SAM team on a daily basis.

Third is the NYS Statewide Parcel Map Program. Through constant outreach and promotion, the NYS GPO’s Statewide Parcel Map Program is working on collecting GIS tax parcel data from counties and building a statewide tax parcel GIS database that shares a common

21 https://www.911.gov/issue_nextgeneration911.html
data scheme and is open for public use. However, counties, as the collector and steward of the tax parcel data, are not fully on board with sharing their tax parcel data to the public for free due to various concerns. So far, only 26 out of 62 counties are willing to make their tax parcel data available for the public to use for free.

5.4.2.2 Utilization of the NYS GIS web services

As one of the focuses of this qualitative study, I am interested in exploring how do these different stakeholders use these NYS GIS web services? and what do they use it for? Through the interviews, we discover that the NYS GIS web services hosted on the NYS Clearinghouse are used by various users, including state governments, local governments, non-profit organizations, for-profit organizations and academic institutions. Some producers of the web services have a general idea about the usage of their web services and the type of the users, whereas some do not. As indicated by two producers of the web services:

Producer 1: we know, and we can see that as soon as we put data up in web services, people incorporate it in their apps.... And we know which apps are hitting our geocoder and yeah so, we keep track of that through our user communities and through the logs, but it's pretty broad, there's lots of student access, there's people on the private sector use for solar potential or wind farm and you know that sort of industries that need to do a lot of sites selection. And companies that are doing site selection for clients for facilities that have to be located somewhere, those folks are using our web services... We have about 30 to 40 million hits a month... There are some seasonal spikes during the summer vacation season...

Producer 2: I really don't, and that's a little unfortunate. It is mostly because we have not pursued that... I'm sure I could get a better idea if I needed to. At this point, we're just kind of putting the data out there, so people can use it. and it doesn't matter who's using it...So I've not really done any kind of in-depth studies to see who the audience is and how it's used.

Different users have different motivations to use the NYS GIS web services (see Table 16). Most interviewees choose to use web services due to its availability, easy access and high quality. And compared to other formats of GIS data, web services also allow users to access the
GIS data without downloading it, which greatly saves users’ server spaces and efforts in data maintenance and management. The service providers update the data on a regular basis, and users can always access the most up-to-date data. Most web services are GIS framework data that has been used in almost every GIS related project. They are the foundation of all GIS operations and provide the geographical context and structure within which spatially manifested processes operate and on which other data are assembled. Some interviewees choose to use web services because these resources are necessities for their daily work and research. Web services also allow multiple users to simultaneously view and work on the same set of data on different devices at different locations, which is quite convenient for teamwork and collaboration. As some of them said:

An interviewee from a profit organization: better accuracy. They did a very nice job with their address points of finding the locations of all the addresses. But their addresses are being added all the time, so if I was able to use this service, I can have the most up-to-date addresses.

An interviewee from a local government: If I don't have to touch someone else's data and update their data, I don't want to do that. It makes no sense. This is what service is all about. This is why we want to use services…. the fact that they can publish that service and I can consume that service, which means that I never have to worry about updating that data, and they always remain updated, because I'm consuming a service not a static download.

An interviewee from a profit organization: My projects are always all over the place, and I don't want to re-download and merge and create things, it is way easier (to use web services).

Table 16 Users’ Motivations of Using the NYS GIS Web Services

<table>
<thead>
<tr>
<th>Motivation of using the NYS GIS Web Services</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>I can always access accurate and the most up-to-date data.</td>
<td>11</td>
</tr>
<tr>
<td>It is easy and quick to access the data (no need to download the data).</td>
<td>9</td>
</tr>
<tr>
<td>It saves the storage on my server and efforts to maintain the data on our own.</td>
<td>6</td>
</tr>
<tr>
<td>It provides data that I use on a regular basis.</td>
<td>2</td>
</tr>
<tr>
<td>It can be used on different devices and used by different people at the same time.</td>
<td>2</td>
</tr>
<tr>
<td>I was informed and educated by others to use the web services.</td>
<td>2</td>
</tr>
<tr>
<td>I hope to improve my work efficiency through using the web services.</td>
<td>1</td>
</tr>
<tr>
<td>It is more suitable to use web services for what I am trying to do.</td>
<td>1</td>
</tr>
</tbody>
</table>
Different users use the NYS GIS web services in different ways. The interviewers’ utilization of the NYS GIS web services can be categorized into following different ways, based on the five-stage OGD utilization summarized by Davis (2010), which is 1) data to fact, in which, users directly identify a specific fact through searching, browsing and extracting from OGD; 2) data to data, in which, users combine and manipulate datasets to form and share a new dataset; 3) data to information; in which, users create a presentation and interpretation of datasets in texts or graphic; 4) data to interface, in which, users create a method to interactively access and explore the datasets; and 5) data to service, in which OGD is used to provide online and offline services.

1) Data to fact: the NYS GIS web services are used as references for users to look up information. The NYS GIS Web services allow users to easily access the most up-to-date GIS data over the Internet without downloading it first. It is very quick and convenient for users to connect to the web services, obtain the visualization of the GIS data and extract any information that they need for their work and research. For example, state government agencies use the orthoimagery web services to extract information that helps identify the area that they will conduct the field work. Local government staff use the NYS GIS web services as references to answer public inquiries, for example, local residents want to figure out their property boundaries, the location of specific buildings, etc., it is convenient for the local government staff to extract the information from the web services to address residents’ questions. Researchers use the geocoding web services to put in the text version of the address, and then extract the geographic coordinate for those addresses, which is the very first step when they conduct location related research. As it was mentioned by different interviewees:

An interviewee from a state government agency: We have a program...we have a biologist who grew up in the field to survey these plants and animals and so they
use the orthoimagery all the time to help plan where they're gonna do their surveys and what kind of habitat it is that kind of thing. It is used for our day-to-day work plus for specific projects.

An interviewee from a local government: I generally use it for reference purposes, for answering day to day questions that the public might have... Some people just want to see it (aerial photos) for references, sometimes they are trying to solve a boundary dispute, and you can see land features with the aerial photo and that is helpful. And the assessors come in and they want to see where houses are... we you know use the aerials for that.

An interviewee from an academic institution: basically we are doing is we are taking...hospitalizations...which we have individual in a record, so we have this person was admitted to the hospital for this outcome on this date and along with that is the address for that person, so basically what SAM service does and you know to use the GIS to talk with and it's basically you feed a file into ArcGIS and then once it does is, it sends the address to the SAM, and we get our geocoded point, so we get the coordinates, the length of the lawn and the ArcGIS then update the length of the lawn in our data file. We basically get data elements.

2) Data to information: the NYS GIS web services are used to create maps for presentations, reports, web pages, etc. This type of utilization has been mentioned the most by all the interviewees. Among all the web services, orthoimagery web services serve as an excellent base layer for almost every map, upon which other layers of information can be overlaid, such as facilities, land use, species habitat, wetlands, tax parcels, etc. Users from different organizational environments often use it as the base layer to create their own maps. These maps are further used as visualization materials for designing web pages of projects or organizations, conducting reports, and supporting communication among developers, clients, policy makers, etc. These visual presentations help stakeholders better understand the issues, greatly facilitate the discussion and effectively reduce misunderstanding and miscommunication, since everyone is looking at the same presentation that is developed using high quality authoritative data through the web services. As many users indicated in the interview:

An interviewee from a for profit organization: Like I said, some of them are web services, especially orthoimagery, which is a very important layer for what we do... part of my work is creating maps for our department webpage, I also make maps for various project reports that people in my division need... Those
layers...they form the base map for what we put our water system on top of...They are the base map, the foundation and then we add our water service on top of that base layer.
An interviewee from a nonprofit organization: we make use of the legislative district web services that the GIS program office puts out, so you know where the state assembly and Senate districts are. You can use those to overlay it with other web servers such as parks and protected lands. We have a good idea of what conservation resources, what area has fell in with conservation use or have been conserved within each our representatives district, which is useful for us to be able to see, but also be able to show them when we are meeting with representatives and asking questions and lobbying in some cases for support for environmental initiatives.

3) Data to interface: the NYS GIS web services are used to build web applications.

Supported by the technology infrastructure of the organizations, and the programming skills of the staff, some users also bring in the NYS GIS web services to be part of their web application. The most common example is called GIS information viewer. It provides easy access to a preview of GIS based information of the area. The built-in filter function allows people to extract and view the information they are interested in with very low requirements on their digital skills.

For example, the Town of Huntington in Long Island, NY designed a Public GIS viewer with incorporating the NYS GIS web services to provide an interactive map displaying general property information for all properties in the Town of Huntington. It allows general public search by address or tax map number, determine their refuse district, carter providing service, and collection days, view special use district information such as fire districts, school districts, etc., query and display properties that meet specific input criteria: minimum and maximum property area, type of land use (i.e., residential) and zoning (see Figure 16). There are also other interviewees illustrated their experience using the NYS GIS web services in their GIS viewer for the public:

An interviewer from the nonprofit organizations: our web mapping applications, were primarily using web services to represent base mapping layers, so again
aerial imagery, hill shading, political boundaries, roads, streams all those very basic GIS layers.

An interviewer from the local government: We are aware of the many services, I am actually starting to identify them where for instant the NYS wetland services I believe coming out of DEC…so yes, we are pulling in services, we have not begun to build applications other than perhaps maybe some GIS viewers, our public viewers, our internal viewers that we can push in those services from directly from the state. So, for instance, if we want to show DEC wetland, I'm not going to download that information and then publish it, I'm actually going to draw the original services.

![TOWN OF HUNTINGTON](https://geo.huntingtonny.gov/helix/professional/Map.html?user=107&token=bd3DLXgG2L4CoJSgsp8KN5X25SDpcVYAkTtQmtddnptjN3gaHEb2OIPMcjg0Rgj3teOCerQB_X6ntV8BcjNn0rco3t6NAt5pH9c1ym4bf_DyGm4ZulyxHSTyX3cp_QTqRWfczn60w-odUVgIm4GNw7xdeZ3x3Fen56Q_zATUYB-u5k1QCvmEOANs43VK&href=Public/EXT_Public&search=Public/EXT_Search&pCard=Public/EXT_PropertyCardMap)

*Figure 16. Screenshot of Municipal Parking shown in the Public GIS viewer in the town of Huntington*

(https://geo.huntingtonny.gov/helix/professional/Map.html?user=107&token=bd3DLXgG2L4CoJSgsp8KN5X25SDpcVYAkTtQmtddnptjN3gaHEb2OIPMcjg0Rgj3teOCerQB_X6ntV8BcjNn0rco3t6NAt5pH9c1ym4bf_DyGm4ZulyxHSTyX3cp_QTqRWfczn60w-odUVgIm4GNw7xdeZ3x3Fen56Q_zATUYB-u5k1QCvmEOANs43VK&href=Public/EXT_Public&search=Public/EXT_Search&pCard=Public/EXT_PropertyCardMap)

It is worth to mention, besides frequent users of the NYS GIS web services, some interviewees also use web services from other sources, such as Arc GIS online, the NYS DOS gateway, Cornell University Geospatial Information Repository (CUGIR), etc. There are also some interviewees, mainly from non-governmental (nonprofit, private, academic) organizations
that often use downloadable GIS data rather than the web services from the NYS GIS Clearinghouse due to various reasons. One of the main reasons is that GIS web services have limited function to support their work and research. This is illustrated more in the limitation section of the web services.

5.4.2.3 Interaction among stakeholders

As one of the focuses of this study, I am interested in knowing how stakeholders interact with each other in the process of producing and using the web services. Through the interview, we get to know that the NYS has a long history of GIS coordination and collaboration among state agencies, with local governments and other community members, such as non-for-profit organizations, private sectors, academic institutions, etc. The coordination and collaboration among these stakeholders are captured in the following figure (see boxes in yellow in Figure 17), it brings many benefits but also encounters several challenges in the process of producing and using the NYS GIS web services.

![Figure 17. Coordination among stakeholders of the NYS GIS web services.](image)
As mentioned in the section about the production and utilization of the NYS GIS web services, the NYS GPO and other state agencies are the main producers of the web services. The local governments have greatly benefited from utilizing the web services. In the meantime, some local governments are also playing a very important role in updating data, such as traffic counts, and street and addresses on a regular basis and share back to the state for them to produce and update related web services. As one of the interviewees from local governments said:

A good example is that we do traffic counts right, so we have data that they want, so we... it's not the entire web service, but it's part of the process, they take our traffic count data and they imported into their dataset that then becomes a service of traffic counts, so whenever possible, we try to be cooperative you know, we try to get any data that we feel would be use useful at a level higher than a region, we try to give that to people to help agencies who would use that... we also do address points for 911 SAMs, we have an internal data set that began as SAM and then we took SAM enhancement greatly, because SAM had a lot errors in them, and then we give that back to them quarterly, so that then becomes part of the service the address point services.

In the GIS development in the NYS, several councils and workgroups have been established to move the collaboration forward. The Geospatial Advisory Council (GAC) has been serving a critical role as part of the collaborative GIS environment in the NYS. The GAC, led by the NYS GPO, is made up of GIS professionals who represent all sectors (federal, state & local government, non-for-profit & private sectors, academia) of GIS users in the New York State. The council meets quarterly with a mission to advise the state on effective use of geospatial technology resulting in tangible benefits to the NYS. From some interviewees who are members of the GAC, we learn that representatives from different fields, through quarterly meetings, get to know the update on the GIS development in New York State, for example, many interviewees get to know and start to utilize the NYS GIS web services since their involvement in the GAC. In the meantime, these representatives are also able to provide feedback and
suggestions for the future development of GIS in the state. Currently, there is an active
discussion going on about rebuilding the website of the NYS GIS Clearinghouse. As a form of
formal collaboration, GAC also aims to create a deliberative environment to facilitate open
communication about issues related to GIS development that will affect stakeholders in different
domains. As one GAC member said:

We have private not for profit, county government, local government, state and
federal representation at quarterly meetings and we haven’t missed a quarterly
meeting in 20 years you know so. We have an agenda, and we have minutes, and
we have reports and it’s a repeatable format. So that’s the formal collaboration
point, but all of those folks have in many cases less formal relationship to
represent their sector and we try to steer people towards... okay you’re on the
geospatial advisory council, we care about your opinions, but we also care that
you represent your sector. So, it’s not just the issues of your county, how about
your neighboring counties, how about your region you know... So, it is very
collaborative, but it’s kind of open communication and much of it is informal and
much of it deliberate.

The State Agency Advisory Group is a subgroup of the GAC. It is made up of GIS
professionals from different state agencies that work together to address GIS interests specific to
the state government in New York. Their feedback will be adopted by the GAC state government
representatives to present at the GAC quarterly meetings. As one of the members described:

So that's the mechanism... There are a couple of state agency reps on the council
(GAC). Those reps get the feedback from their whole sector through the state
agency advisory group. It's an important part, since not all those folks on the state
agency that come to the Geospatial Advisory Council meetings and would fit in a
room for the Geospatial Advisory Council and that would sway the balance of
representation in the Geospatial Advisory Council.

Besides councils and workgroups, the NYS GPO have also designed different channels to
facilitate the communication among stakeholders. The NYS GIS Email Listserv is the most
mentioned communication channel that is used by all the interviewees. It is an open,
unmoderated list developed to promote discussion of GIS-related subjects, particularly GIS
issues or projects in the NYS. State agencies have been using it to notify any update on the NYS GIS web services and other GIS resources. Users have been using it to report issues in finding and accessing GIS resources they need. There is also much promotional information about GIS webinars, tutorials, software, news, etc. It helps keep all the subscribers up to date about the development in the GIS field. Besides the listserv, there is also contact information of data providers listed on the dataset web page on the NYS GIS Clearinghouse. If users have any issues with the datasets, they can easily get hold of someone that can help them. The NYS GIS Clearinghouse also provides a “who’s who in GIS” directory that contains a list of GIS professionals and their contact information to help them connect with each other. For example, this resource has provided a great help for me to recruit interviewees for this study.

Besides government-led groups and communication channels, there are many state and regional GIS associations and user groups that have been contributing to build an active GIS community in the NYS, such as the NYS GIS Association, the Tug Hill GIS user groups, the Long Island GIS user group, etc. They provide a welcoming environment that allows GIS users to get together, discuss GIS topics in general, GIS development in the NYS and issues related to GIS development in their own regions and create opportunities for collaboration in conducting research projects and addressing community development issues. Many interviewees have mentioned that they get to know about the NYS GIS Clearinghouse and the NYS GIS web services by participating in those associations and user groups.

In sum, in the NYS, there are different government-led and non-government led groups, information sharing between local governments and state governments, and various formal and informal communication channels that help engage different stakeholders and facilitate the coordination and collaboration among them. In general, there is more frequent communication
between state agencies and local governments due to their collaboration in data sharing. Users from nonprofit organizations, private organizations and academic institutions often just go and grab the resources they need from the NYS GIS Clearinghouse. Communication only happens when they have issues with the resources they need; they just directly contact the data providers. The nonprofit organizations also communicate with the local governments in the area that they are in to acquire data.

The coordination and collaboration among these stakeholders greatly facilitate the information sharing to enable the development of web services and other GIS based products. It keeps stakeholders informed and improves their involvement in the development of GIS in the NYS. The shared information also helps reduce redundancy and ultimately improves government efficiency. As interviewees from different state government agencies said:

*The harmony among the professionals and the different agencies, the state and the different stakeholders are a big beneficial factor for the development of the web services in NYS.*

*I think we're having a huge impact on site development in New York state and efficiency of you know your automatic wayfinding apps work better in New York than they do in other states, because you're making all this data available, so we're having a big impact across the board from things like a citizen's directly benefit from to local government efficiencies and state agencies efficiencies which indirectly also benefit the taxpayers, which just allow us to do more work with the resources we have.*

However, there are also many challenges in their coordination and collaboration. Open tax parcel is one of the biggest challenges mentioned by many interviewees. The state has devoted great efforts to get counties in the NYS on board to make their tax parcel data open to the public. Currently, only 26 out of 62 counties have joined the statewide parcel program and are willing to make their tax parcel data open to the public. In our interview, we have covered some GIS staff from the counties that are not willing to share their tax parcel. According to them, they personally are not against the idea of open parcel data, however the management and
leadership of the county are reluctant to do so due to concern about losing revenue from selling tax parcel information and invasion of privacy. As one local government interviewee said:

*I don't have anything to do with that decision, the people in those two counties have the ultimate authority over those datasets. I don't agree with their opinions about why they don't share that data, but they have chosen not to from a political standpoint. I think their concerns are two things. 1) they have a perception that if they publish those parcels to the public, they're gonna somehow wound lose a tiny amounts of revenue that they get from selling that data, that notion is so archaic, in my opinion that you know that the cost benefit analysis of having your parcel data out there is so outweighed that you should have it out there as opposed to charging people money that taxpayers have already paid for, in my opinion it's a misguided opinion to do that, but things are done for that reason 2) then there's some people have a notion that this data is somehow you know invasion of privacy, which you know anything about real property data, it absolutely not, because real property data by state law is public information, it is already out there, so two things money and a pristine invasion of privacy.*

There are some other challenges mentioned by interviewees that are hindering the information sharing between local governments and the state agencies. On the local government side, some counties have experienced technology challenges, lack of staff and limited staff skill to keep them on top of updating and sending the data back to the state. On the state government side, they also indicated the pressure to support the peak time when users access the web services. Some local governments complained about the red tape at the state when data flow from the county to the state and unclear instruction about the formats and kinds of the data that the county should provide, as one interviewee from the local government said:

*They have rules and gag orders, don't tell them this, that I cannot figure out why state agencies are secret about what they do and how they do things. And it's to the point that sometimes I actually experienced that their own bureaucracy has prevented them from giving me data for some times a year. I cannot really explain why the state would not explain or tell us what format they want whatever the data they want in. Because the process they are using on their end is so secret, they do not want to give that clue away.*
5.4.3 Facilitators and obstacles

There are different facilitators and obstacles mentioned by interviewees that have influenced the process of their production and utilization of the NYS GIS web services. In this section, I discuss the facilitators and obstacles for producers and for users separately, and these factors are summarized using the categories I designed in the literature review section, namely, government related factors, OGD related factors, user related factors and community related factors.

5.4.3.1 Facilitators and obstacles for producers of the NYS GIS web services

For producers of the NYS GIS web services, in this case the different state agencies in the NYS, the most mentioned facilitators of producing web services are mainly government related factors and community related factors (see Table 17).

Table 17 Facilitators of Producing the NYS GIS Web Services

<table>
<thead>
<tr>
<th>Facilitators</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Government related factors</td>
<td></td>
</tr>
<tr>
<td>Government’s good awareness and clear vision of OGD</td>
<td>2</td>
</tr>
<tr>
<td>Resources support</td>
<td>3</td>
</tr>
<tr>
<td>Community related factors</td>
<td></td>
</tr>
<tr>
<td>Great communication and collaboration</td>
<td>3</td>
</tr>
</tbody>
</table>

1) Government’s good awareness and clear vision of OGD. New York State, as one of the pioneer states in OGD development, has designed and implemented OGD policies, programs and tools that provide public access to digital data for collaboration and analysis. The NYS GPO also has a good awareness of OGD and firmly believes that by making more and more open GIS data available to the public, it will empower the users and better support activities in addressing community issues. With a comprehensive understanding of the business problems that need to be addressed and the diverse users’ needs, the NYS GPO and other state agencies have been
pushing out more and more GIS web services for users to better utilize the GIS resources in their work and research. As one interviewee from the state government said:

*Having a clarity of purpose. And having a pretty deep understanding of what business problem we're trying to solve. And if we identify that it doesn't have to be comprehensive, we need a deep understanding of a few key business problems, we need to understand the ones that are related to the resources that are put into it, the funding, the few key activities that we understand deeply, the rest just come along for the ride.*

2) **Resources support.** In this case, the resources include technology infrastructures, staff skills and funding. The production and provision of the NYS GIS web services relies on robust technology infrastructures (e.g., servers, software etc.) to be able to host the data, ensure the stability and speed of the connection to access the data. In the meantime, staff with good understanding of GIS and geographic principles, decent skills in GIS related technologies and great communication skills are all necessities to support the production of the GIS web services. Finally, the continuous funding is the foundation that supports the purchase and upgrade of the technologies and recruitment of the skilled staff. As interviewees from the state government said:

*There's lots of technologies that will handle each of those parts of it, as software is the most popular right now. So, we use a lot of that, but we also are looking at alternatives that scale without concern for licensing.
So, a basic understanding of GIS and geographic principles. People that think spatially...if you have the spatial thinking to understand how what you're doing impacts lots of different workflows, that's really the skill that gives you the energy to keep at it.
In my mind the most critical thing is that the funding is there and it's in place to continue the development of the gateway.*

3) **Collaboration.** The NYS has a long and proud history of GIS coordination and collaboration. Workgroups like GAC, different GIS associations and user groups has helped establish a formal but deliberate environment that includes representatives from the federal, state, local governments, nonprofit organizations, private organizations and academic institutions to
facilitate information sharing and collaboratively work on addressing issues related to the GIS development in the NYS. As one interviewee from the state government said:

*I would say in New York state we are extremely lucky, because we have a very well-coordinated program, and the individual players work very well together, there's great communication. I would say there aren't really many problems, we work very well together; the New York state U.S. community is very close and well informed.*

Besides these facilitators, producers also have encountered some obstacles that are changing over time in the process of producing the web services, which are government related, data related and community related factors (see Table 18).

**Table 18 Obstacles Encountered in Producing the NYS GIS Web Services**

<table>
<thead>
<tr>
<th>Obstacles and Challenges</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Government related factors</strong></td>
<td></td>
</tr>
<tr>
<td>Counties’ reluctance to share data (e.g., tax parcel)</td>
<td>2</td>
</tr>
<tr>
<td>Resources (e.g., funding, technology)</td>
<td>5</td>
</tr>
<tr>
<td><strong>Data related factors</strong></td>
<td></td>
</tr>
<tr>
<td>Keep data up to date</td>
<td>1</td>
</tr>
<tr>
<td><strong>Community related factors</strong></td>
<td></td>
</tr>
<tr>
<td>Communication and partnership building (more challenge initially)</td>
<td>1</td>
</tr>
</tbody>
</table>

Initially some producers indicated that communication and partnership building has been a challenge. They physically went to meet with local governments two or three times to understand each other’s needs and how they could collaborate in data sharing and updates that support the production of GIS web services. This is still a challenge nowadays, but the state agencies have gained more confidence, because the benefit of this collaboration has been revealed, which has made more and more partners join the collaboration. The biggest challenge among all the programs development is the tax parcel program, where most county governments
(36 out of 62) are still reluctant to share their parcel data due to the concern of privacy and loss of revenue. As one interviewee mentioned:

*It's a fun problem itself and there's always a challenge with different perspectives and different policies, which I respect, but permissions to make data publicly available is hard for some folks to weigh the value of that over the concerns. Concerns of loss of revenue, concerns of privacy, which are really the two big ones.*

Currently, the bigger challenge that most producers are facing is securing sustainable funding to support the high cost of the technology. Maintaining sustainable funding is always a challenge due to the changing economic environment and leaders in the organizations. Some producers said they had to start the budget application very early to ensure it got approved.

Another challenge is keeping the data update to date. Ensuring the data quality and keeping it up to date is another challenge mentioned by some producers, especially for those who are not the ones that hold the data and update the data. As the producers of the web services, when they receive the data from other entities, usually from local governments, they need to make sure the accuracy of the data before they push it out as web services for the public.

5.4.3.2 Facilitators and obstacles for users of the NYS GIS web services

For the users, in this case, users from the state government agencies, local governments, nonprofit organizations, private organizations and academic institutions, the facilitators during their utilization of the NYS GIS webs services are mainly data related factors, user related factors and community related factors (see Table 19).

| Table 19 Facilitators of Using the NYS GIS Web Services |
|---------------------------------|-----------------|
| Data related factors            | Frequency       |
| Data availability: The availability of the resources that are necessity for doing GIS mapping | 11              |
| Data accessibility: all web services are listed on one web page; easy access and use; stable, fast connection/great performance | 16              |
### Facilitators

<table>
<thead>
<tr>
<th>Facilitators</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data quality: authoritative resources; completeness of the data; accuracy of the data; most up to data</td>
<td>19</td>
</tr>
<tr>
<td><strong>User related factors</strong></td>
<td></td>
</tr>
<tr>
<td>Perceived benefits: save storage space; save time on data managing and downloading</td>
<td>10</td>
</tr>
<tr>
<td>Users’ skills: not much requirement on users’ skills</td>
<td>7</td>
</tr>
<tr>
<td>Supportive technology infrastructure on the users’ side</td>
<td>3</td>
</tr>
<tr>
<td><strong>Community related factors</strong></td>
<td></td>
</tr>
<tr>
<td>Responsive data providers</td>
<td>1</td>
</tr>
<tr>
<td>Channels of communication among stakeholders: GIS user groups, state email listserv, etc.</td>
<td>2</td>
</tr>
</tbody>
</table>

Among these facilitators, data related factors were most mentioned by the interviewees. Many interviewees highlighted the high quality of the NYS GIS web services and indicated that they have had a great experience utilizing this resource. Based on the responses, the data related factors can be categorized into three factors.

1) **Data availability.** GIS data is generally difficult and costly to collect. The NYS GIS Clearinghouse made certain GIS data available through web services to the public at no cost. Users are able to take advantage of these resources that are not available anywhere else, which greatly supports them to conduct GIS maps and develop applications in their daily work and research. As one user said:

   *I am gonna say one thing that the GIS web services are here, even though we are not frequent users, we are only interested in a specific type of data, but just because it's there, it just makes it so much easier than trying to dig out data from other places. And I think it encourages new users, that's what I'm trying to say.*

2) **Data accessibility.** This refers to the ease of access and the stability of the access. About the ease of access, on the NYS GIS Clearinghouse, the NYS GPO has designed a specific page listing all the GIS web services published by different state agencies in the NYS. It is very easy and convenient for users to go to this one central location and access the web services they
need. About the stability of the access, the NYS GIS web services provide data to users through streaming over the Internet. The state agencies have invested a lot in their technology infrastructures and staff time to ensure the stability of the GIS web services and speed of the connection to these GIS web services. Most interviewees indicated that they have never run into issues connecting to the NYS GIS web services and the connecting speed was usually faster than other data sources. As one user mentioned:

_They very rarely go down and they are very rarely unavailable. In fact, the service is very consistent. I run it every day, and I almost never have a situation where I can't do my job because the service is not available, or it is too slow._

3) **Data quality.** It is one of the most mentioned facilitators for users to choose to use the NYS GIS web services. Data quality is a multidimensional concept; in this case, it refers to the authority of the data source, the completeness of the data, the accuracy of the data, and the frequency of the update of the data. Many users choose to use the NYS GIS web services rather than data from other sources, because it is from authoritative state agencies. The completeness and accuracy of the data is more guaranteed compared to GIS data in other platforms where the data are shared by individuals or unknown sources. The producers of the NYS GIS web services also follow a fixed schedule to update the data and push out updated web services. Users will be notified through NYS GIS listserv about any change and update. Also due to the automatic connection of the web services, once the update on the web services gets pushed out, the update will automatically reveal on the applications of users’ ends. It greatly minimizes users’ efforts to keep track and update the data on their own and helps them easily and always get the most up-to-date data. As one of the users said:

_Well, the beneficial factors that you know it's something that's more recent, updated, allegedly, quick and easy, I don't think I have to process much if I was using you know the hill shades for example, that's not something I would have to do myself._
Besides data related factors, user related factors are the other type of facilitators mentioned by many interviewees. Based on the responses, the user related factors can be categorized into three factors.

1) **Users’ perceived benefits.** Users’ perceived benefits is the most mentioned user related factors. Many users indicated that they could easily access the data they need without downloading it, which greatly saved the storage space on their own servers. The automatic connection of the web services also greatly saves user’s time in downloading the data and managing the data, especially the big size data and data that frequently gets updated. Their perceived benefits of the NYS web services have made them choose to use web services rather than downloading the data. As one of the users said:

   *Where I am working, it has an update, I do not need to download to have it updated. It’s just like a quick check. That thing has changed, or I just need to do a quick map like bringing in every other layer from our database just bringing in and just being a quick map. Ideally it is always up-to-date stuff. Our staff needs a map for an event very quickly, and it involves bridges and roads map in some way, it is just a quicker way of doing it.*

2) **Users’ skills:** Many users also indicated that their digital skills such as programming skills and GIS knowledge have helped them adapt to this new method of accessing GIS data. Many users said this new method of accessing data required certain but not too much technical skills, once you knew how to do it, it was very simple and easy.

3) **Supportive technology infrastructure.** This is another facilitator mentioned by a few interviewees. Most of the users are using ArcGIS software in their working and research environment, which allows users to easily connect to the web services and import into the software through the built-in functions. Many organizations indicated that their adoption of the
web services rely on the support from their skilled staff and their technology infrastructure. Like one of them said:

_So, we happen to have an advanced GIS platform, so consuming state services is really a no brainer to us... Yeah you need to have a platform and you need to have people that understand, that is a good point, you need to have human resources, people in your environment and understand how to use a rest service and then how that rest service relates to geospatial web service._

A few interviewees also mentioned community related factors that are facilitating their utilization of the NYS GIS web services. Based on the responses, the community related factors can be categorized into two factors.

1) **Responsive data providers.** Users indicated that when they had issues about certain GIS web services, it was easy for them to get hold of the NYS GPO staff, and they were very responsive in addressing their questions. As one of the users said:

_For the GIS program, the office has always been very quick to respond, they are first to come in and very helpful whenever I've had any questions._

2) **Channels of communication among different stakeholders.** Users are often involved in GIS user groups in different regions of the state and get to learn about the NYS GIS web services and discuss projects, research and issues related to using these web services. The NYS GIS email listserv serves as a popular communication channel, and it provides a great platform for users to get informed about the update and change of the NYS GIS web services, to discuss issues in finding and accessing these web services, and how to incorporate it into their work and research. Like one interviewee mentioned:

_There are 2 lists. There's what we call the GIS listserv, that's been around for 20 years almost and there's 1000 people who have signed up to get all kinds of notifications from job announcements to participating and helping people answer their questions, to understanding that we're making data available. But there's a specific list that people can send their email in and get on the list to be notified of changes to our geo coders and that sort of thing._
In general, users have a satisfactory experience using the NYS GIS web services, especially for frequent users that often use one specific web service. However, many users also mentioned different obstacles that they have encountered at their different stages of using the NYS GIS web services, which are categorized as data related factors and user related factors (see Table 20).

Table 20 *Obstacles Encountered in Using the NYS GIS Web Services*

<table>
<thead>
<tr>
<th>Obstacles and Challenges</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data related factors</td>
<td></td>
</tr>
<tr>
<td>Limited number of web services, especially web feature services, and web services for specific regions</td>
<td>4</td>
</tr>
<tr>
<td>Data quality issues: incomplete data (e.g., incomplete tax parcel); lack of metadata</td>
<td>8</td>
</tr>
<tr>
<td>Data accessibility issues: server down, changing web services link address</td>
<td>3</td>
</tr>
<tr>
<td>User related factors</td>
<td></td>
</tr>
<tr>
<td>Hurdle of connecting the web services due to lack of technology knowledge, especially for new users</td>
<td>5</td>
</tr>
<tr>
<td>Not easy to find the right data (the one that is useful for you), especially for new users</td>
<td>4</td>
</tr>
<tr>
<td>Not aware of the available resources out there, especially for new users</td>
<td>2</td>
</tr>
<tr>
<td>Traditional preference of hard copy of data over connect to web services</td>
<td>2</td>
</tr>
<tr>
<td>Internet connection is not stable on the user end.</td>
<td>3</td>
</tr>
</tbody>
</table>

For frequent users, the most mentioned obstacles for them are mostly data related factors.

1) **Limited number of web services**, many users have mentioned that they would like the state to publish more web services, especially web feature services. As they have experienced the convenience of the web services, they would like to have access to more web services in more subject domains or web services that focus on a specific geographic area. For example, one of the users complained that there was a lack of GIS resources for the Long Island area.

2) **Data quality issues**: Even though most users think that the NYS GIS web services are of good quality, they still run into some quality issues while they are using it. For example, many
users mentioned the incomplete statewide tax parcel data, which has cost them a great effort in obtaining tax parcel data, especially for users who needed tax parcels data from different counties. They had to turn to each of the county governments and went through a lot of paperwork to acquire the data they needed. Some users also complained about incomplete metadata, which has made it difficult for them to understand the dataset. Like some users said:

For the tax parcels layers, they have both a polygon layer and a centroid layer. And the thing is the polygon layers is incomplete, it doesn't have every county, so for our service area, which is about 11 counties in the Hudson Valley, only about half the counties had polygon data available to that service. ... because it was incomplete for our service area, we didn't think that that service would meet our needs, so I had to essentially recreate the data in house and publish it ourselves to include the counties that were missing in the statewide version.

If you have any questions about the data and it's not there in the meta data, that could be a challenge.

I think just the frequency of updating the data, that could be challenging as well.

If I need the latest and greatest, usually I will end up reaching out to someone directly, because it is not right there that's not a function of the people put there for the web services.

3) Data accessibility: The access of the NYS GIS web services heavily relies on the Internet. In general, users have no complaint about the connection of the NYS GIS web services, but some users indicated they had difficulties to connect to the services, either because the server was down or due to the speed constraint on the users’ ends. Another issue with the connection is that, sometime, when the service providers update the web services, the links of the web services change, which affects the users’ connection. They have to refine the links on their ends and re-connect the web services to access the data. As one of the users complained:

For instance, recently NYS they change of the web server address for the imagery, and it ended up breaking my online map to the point that I had to rebuild it, because the link to the ortho changed, basically corrupted it and I couldn't open it to fix it, and have to re-pull into a new server, so that was frustrating. That is not the first time that they changed the address for the ortho, at least the third time since I've been using them. The location of the ortho has changed and I have to change all my links to you know find them again.
New users or non-frequent users also indicated several obstacles that they have experienced in the process of utilizing the NYS GIS web services. When they first got introduced to the NYS GIS web services, many of them did not know what kinds of web services were available and which one was useful for their use. Many of them also complained about the interface of the NYS GIS Clearinghouse, which only had a drop-down menu to filter and see the list of the data, users were not able to search data using keywords. Even though the web services are hosted on one specific page, where it is located is not that intuitive for users who are not familiar with the NYS GIS Clearinghouse to find them. Also, as a new method of accessing GIS data, accessing the web services has created certain technological hurdles for some new users to get familiar with how to connect to the web services. The technological hurdle has greatly disturbed users getting to know about the NYS GIS web services and trying to utilize them. As some of them said:

My general impression is that a lot of people are not even familiar with how GIS web services work. I know that when we were first offering WMS and WFS through Cougar, we were not sure our audience would know what to do. I think in general they are underutilized for many reasons. I think a lot of people are just uninformed about how or when they are useful.
I think there's definitely a bit... like technical GIS acknowledging that you have to know...like how to add and connect the services. You know where to find it. And so, I think that is one of the biggest hurdles.

5.4.4 Benefits, limitations and future development

The NYS GIS web services, as an automatic means to access authoritative and most up-to-date GIS data, has been greatly utilized by users from different types of organizations and has brought a great benefit for their work and research. However, there are some users that still prefer the traditional way of downloading data due to several limitations of the web services. Regarding these limitations, interviewees also mentioned their expectation and suggestions for its future development.
5.4.4.1 Benefits of the NYS GIS web services

Interviewees have mentioned several beneficial impacts revealed through the provision and utilization of the NYS GIS web services. These beneficial impacts can be categorized as political impact, economic impact and social impact.

1) Political impact. With the aim to better satisfy users’ inquiries about GIS data and make GIS data fully reveal its value, the NYS has devoted great efforts in creating a collaborative environment that supports information sharing and the development of open GIS data. The NYS GIS web services, a new form of open GIS data, has been published by multiple state agencies and made available for the public to use for their own purposes. It helps improve government transparency and reduce the general public’s lack of trust in governments. Through our interviews, we discovered that many state government agencies and local governments were frequent users of the NYS GIS web services. The NYS GIS Clearinghouse centralized all the GIS related resources and thus has become many users’ first stop when they search for the data they need for their research and work. The NYS GIS web services have to a great extent supported public service delivery, such as addressing public inquiries, decision making for community development, etc. Easy access and use of high-quality data have helped improve the efficiency and effectiveness of government work. Users from different government agencies can easily access the authoritative and most up-to-date data without downloading it, which is very suitable for them to use as a quick reference and helps reduce their concerns about the storage of their servers and quality of the data. All these feedbacks suggest that the provision and utilization of NYS GIS web services have improved government transparency and increased the efficiency and effectiveness in delivering public services. Like some interviewees from the government sector said:
It eliminates the foil requests; it allows people to have native access to deep data and deep large data. And it allows them to have access to a much more authoritative data source, there is no foil process and anything that they can discover using that access is available to them. And what it does is at least I believe mitigates the unfortunate perception that many of the public have about the government.

So basically, the services gave you high quality of the data serving as a reference to cross check your work and make sure that the information you will use to produce further services for its clients are accurate and high quality. But the fact that the performance is very good and also practical, it is not just a convenient way to have the latest data but also it draws fast enough, so you can work productively with it.

2) Economic Impact. Besides government users of the NYS GIS web services, nonprofit organizations, private organizations, and academic users all benefit from this resource. Since the majority of the interviewees mainly use the NYS GIS web services for their work, most of the utilization is more related to generating economic impact. The availability of different GIS web services makes it possible for users to access this valuable, high-quality, and most up-to-date data that are not available anywhere else for free and with no need to download. It helps users minimize the duplication of efforts and investment in data collection, storage and management. The NYS GIS web services have been used to create maps for presentations, reports, web pages, etc., which serves as tools to facilitate communication and collaboration among governments, developers, clients, policy makers, citizens, etc. It allows them to discuss issues about community development based on authoritative and accurate information with less misunderstanding and better effectiveness in decision making. These web services have also been used to design web and mobile applications for public use and commercial use. In sum, the utilization of the NYS GIS web services has helped reduce the cost of government and non-government users in accessing and utilizing data for their work and projects. It also provides valuable resources that facilitate community development, the emerging of new business/products and economic development in general. For example, some companies have
utilized these resources when they think about expanding their business. Amazon built their warehouse in Schodack, NY and Regeneron, the pharmaceuticals company is building a new warehouse, plus manufacturing, office and lab space in East Greenbush, NY after checking the tax parcel data in Rensselaer County. As some interviewees said:

*It just provides a lot of value with a very low maintenance overhead, so that provides us a good useful tool that we can rely on, and we don't have to spend time managing, so that really gives us a good starting point for a lot of our mapping. So, it's high value, time saving, and high-quality product in general. The primary one that comes to my mind, probably the most important advantage is currency of the data. That's powerful. Trust me. At the county level here on our enterprise services, it is very important for every single agency out there, every single department of Oneida county, every citizen see the same thing every time, you cannot have one version of that data viewed by one group of people here and then a different version that's two years older for a different group. They're gonna come up with different conclusions at some point and that's never good. Having it centralized for the state is a big support for people who do this and there is a lot of economies of scale I think of having that kind of resource available that helps to facilitate the work that other people can do disseminate across the state without the redundancy that every single person doing those same steps to effectively create the same end product.*

3) Social Impact. Through the interviewees, we only identify one explicit example that indicates the availability of the NYS GIS web services has yielded some social impact, namely helping the general public better understand the issues they're trying to address around their communities and improve their engagement in community discussion. In one of the presentations about the NYS open parcel data, the presenter from the NYS GPO mentioned that local citizens have utilized the NYS GIS web services and other open GIS data in Rensselaer County, NY to quickly act on their concerns about the establishment of Amazon warehouse in their community. They conducted a map analysis of the parcel data to remind the town about the law that stated all residents within 1 mile must be informed and allowed to request a pre/post inspection. The blasting of the site for Amazon warehouse was rescheduled as a result. This is a great example showing the social impact of the NYS GIS web services, which is that beyond being informed,
citizens are now enabled to provide feedback and become more engaged in discussing issues related to the development in their communities.

5.4.4.2 Limitations of the NYS GIS web services

Besides the various benefits of the NYS GIS web services, the interviewees also indicated that compared to other forms of OGD, the NYS GIS web services have some limitations. Some limitations are specific about the GIS web services themselves, some are related to the stakeholders who are involved in the production of the web services.

1) Dependency on Internet connection. The access of the NYS GIS web services greatly relies on the quality of Internet connection, both on providers’ and users’ ends. This dependence of the NYS GIS web services on high-quality Internet limits the ability for some users who are in the areas where the Internet connection is not sufficient to utilize the NYS GIS web services. Besides Internet connection, the other technical constraint is the server capacity during peak time. For certain GIS web services, during a certain time of the year, it will run into a peak time of utilization when increased numbers of people access the web services at the same time. It may be challenge for the current capacity of the servers to support the quick connection of the web services, as one of the users said:

One limitation is that it is still going through the network, so your ability to pin, to zoom and to look is only as good as the slowest links. So, there might still be really fast connection internet at your house or office and there’s almost always really good connection in and out of where the state data center is or on the cloud, but somewhere along that list, you’re going through a little bandwidth area or there’s just lots of use, it maybe sort of slow to make use of some of the services.

2) Limitations in the number and function of the web services. The NYS GIS web services has provided users with free and valuable GIS resources for their work and research. However, many users mentioned that for them, there was a lack of web services or incomplete
web services that suited their interests. For example, many users mentioned the tax parcel data; due to the reluctance of the counties, the state is unable to provide statewide tax parcel data as web services to the public. Users who are in need have to go through several hoops to obtain the data from each county; some require filing paperwork, and some cost money. Users also complained that certain web services, especially certain web map services had a low level of usability. They only allowed them to have a view of the data, they could not perform any editing, such as changing symbology or conducting advanced spatial analysis. Web feature services that have a higher level of usability, however, currently are limited published by the providers. As some interviewees mentioned:

The biggest problem I have is if the parcel layer and the parcel polygon layer were complete for our region, that would be a very useful thing for us, so if the data is incomplete, its utility is drastically reduced, and so we end up having to recreate anyway. You can't modify it if you want to create another field and make that field the result of the calculation of other fields, you can't do it with the service, especially you do not have the modification right to do it. In those scenarios, you have to download the data, and pull it locally and modify it.

3) Government bureaucracy and traditional mindset. About the stakeholders who are involved in the production of the NYS GIS web services, interviewees mentioned that bureaucracy and traditional governmental mindset still inhibited data and information sharing among governments at different levels. Some county governments that are unwilling to open up their tax parcel data is a great example. In the meantime, the adoption of consistent data standards is missing during the current data sharing and data related collaboration among governments, which negatively affects the efficiency and effectiveness in government work. In the outreach of the state governments, more frequent communication tends to happen among their known networks, but many other stakeholders are left out in the outreach. It results in that
the providers of the web services do not know their needs and these potential users cannot fully grasp the benefits of the NYS GIS web services. As some interviewee said:

*I think that the counties are most protective of their personal data, but the other services, so for instance like the street layer, I don't think that they are that protective, because the state really manages. I mean I know the locals do their edits, but I think it just depends... I mean I guess that parcels are the most protective for the counties. But, yes, it would be nice to have more feature services. Maybe the adoption of standards between governments. Everyone has their own scheme system in their own field settings in the field, table schema and so forth, maybe the adoption of standards would help. So just outreach into those groups that are kind of around our industry but not core of the industry are the next big challenges.*

5.4.4.3 Future development of the NYS GIS web services

In general, most of the interviewees have a very good experience using the NYS GIS web services, they indicated that they would continue to use and would use more of this valuable GIS resource in the future. To achieve the sustainable development of the NYS GIS web services and also gain more benefits from using it, interviewees also raised several suggestions for its future development. Many of them are related to the obstacles and limitations mentioned above.

Following is a list of suggestions for the further development of the NYS GIS web services, the NYS GIS Clearinghouse and the providers of the NYS GIS web services.

1) Suggestions about the NYS GIS web services:

- Publish more web services, especially more web feature services.
- Have default symbology built in the web services, such as how to mark highways, side streets, etc.
- Improve metadata with adding more description about the data.

2) Suggestions about the NYS GIS Clearinghouse:
• Host all the web services on a state-run website and cloud servers, which helps ensure that different agencies are using the same resources that are hosted in the same environment.

• Better design of the website/interface of the NYS Clearinghouse, especially about the data search function. Fuzzy search, voice search and other search methods should be added.

3) Suggestions for the providers of the NYS GIS web services:

• Advocate more local governments to be on board to help make state-wide open tax parcel data to the public.

• Get every county to use standard tax parcel format.

• Provide support to help improve the GIS capabilities of local governments, mainly refers to improving technological infrastructure, increasing the number of staff and staff capabilities of utilizing GIS resources.

• Better outreach to promote web services and improve people’s awareness of it, including GIS professionals in government at different levels, non-profit organizations, private sectors, academic institutions, and the general public.

• Provide training and education to improve people’s capabilities to utilize these open GIS resources in different organizations, including governments at different levels, non-profit organizations, private sectors, academic institutions, and the general public.

5.5 Discussion

Focusing on the case of the NYS GIS web services hosted on the NYS GIS Clearinghouse, this qualitative study was conducted to address two research questions related to the process of OGD utilization: RQ2: How do stakeholders interact with each other in the process of OGD
utilization? RQ3: How do different stakeholders perceive the facilitators and obstacles in the process of OGD utilization? More specifically, who is providing the web services? Who is using the web services? What are they using it for? How do stakeholders interact with each other in the process of producing and using the web services? How do stakeholders perceive the facilitators and obstacles in the process of using the web services? What are the benefits and limitations of the web services?

Among previous research studies, very few of them have looked into the process of OGD production and utilization. This case study has provided detailed empirical evidence for us to better understand the production and utilization of one specific OGD: the NYS GIS web services. The findings of this study, to some extent, support the findings of previous research studies. They also add new insights that help us better understand the process of OGD production and utilization by illustrating the interaction among stakeholders and how different factors affect OGD utilization from the perspectives of different stakeholders. In the following paragraphs, I compared the findings of this case study with previous research studies.

First, about the roles and motivations of stakeholders. The case study shows that various stakeholders are involved in the process of producing and utilizing the NYS GIS web services and can be roughly categorized as providers, users/infomediaries and end users using the scheme designed by previous studies (Lassinantti et al., 2019). As indicated by previous studies, governments and their departments act as dominant control points in OGD development that oversee publishing, modification, update, access and management of OGD (Sangiambut & Sieber, 2017). In this case, different state agencies are the main providers of the NYS GIS web services and the NYS GPO are managing the NYS GIS Clearinghouse where most of the web services are hosted. However, they are not the only ones that contribute to the provision of OGD.
Local governments also indirectly contribute to the production of the web services by updating and sharing related GIS data to the state.

Just as Wang & Lo (2016) stated, these government actions in open GIS data and web services provision are driven by internal and external motivations. In this case, externally, governments aim to empower the users to utilize these GIS resources for decision making and community development. Internally, they are expected to help remove redundancy in government work and improve their efficiency in answering public inquiries, which aligns with the previous research that indicated governments aim to achieve political and economic value through OGD (Janssen et al., 2012).

Most of the stakeholders included in this case study are users/infomediaries of the NYS GIS web services from different organizational environments (governments, non-profit organizations, for-profit organizations and academia). They extract, aggregate, process and transform the data into a format that is more useful and beneficial for the end users just as previous studies described (Magalhães et al., 2013; Chan et al., 2016; Johnson & Greene, 2017). We also noticed that some stakeholders are playing a dual role in this process which has not been extensively discussed by previous research studies. In this case, some state government agencies are not only the producers but also the users of the NYS GIS web services.

Different from Brugger et al (2016) said that different types of stakeholders are playing the role of infomediaries with different motivations, among the responses that I collected in this study, I did manage to find a common theme among users’ motivations, that is the availability of the NYS GIS web services provides the users with easy access, high quality and most up-to-date data which are necessities for their daily work and research. And compared to other formats of
open GIS data, web services also allow users to access the GIS data without downloading it, which greatly saves users’ server spaces and efforts in data maintenance and management.

Second, besides these findings of roles and motivations of stakeholders in OGD production and utilization that have been mentioned by previous study, this study also took a step further to explore the interaction among these stakeholders, which was mentioned that would affect the processes of co-production and co-creation using OGD (Hivon & Titah, 2017; Kučera, 2017; Lee et al., 2015; Maccani, 2016), but were discussed in a limited way in previous studies. In the NYS, there are different government-led and non-government-led groups, and various communication channels that help engage different stakeholders and facilitate the communication among them. However, in reality, two-way communication happens more frequently between state agencies and local governments during the process of producing the web services due to their collaboration in data sharing. Users from non-profit organizations, for-profit organizations and academia mainly just go and grab the resources they need. The communication among them rarely happens and only happens when they directly contact the data providers for data related issues. Instead with the state government, the nonprofit organizations often communicate with the local governments in the area that they are in to acquire data. The coordination between state agencies and local governments facilitates information sharing, helps reduce redundancy and improves government efficiency. However, it also encounters diverse challenges, such as the unwillingness of local governments in sharing data with the public, the costs and increasing demands for new technology infrastructures, the lack of staff and limited staff skill, etc.

Third, about the types of utilization. In this case, the utilization of the NYS GIS web services fits into the stages of OGD utilization defined by Davis (2010), namely 1) data to fact:
the NYS GIS web services are used as references for users to look up information. 2) data to information: the NYS GIS web services are used to create maps for presentations, reports, web pages, etc. 3) data to interface: the NYS GIS web services are used to build web applications.

And similar to the summary of previous examples of OGD utilization, most of the utilization falls into the category of data to information. It is worth mentioning that the production of the NYS GIS web services in this case can also be viewed as a way of utilizing OGD: data to data. These state agencies have transformed the raw open GIS data into GIS web services, a format of open GIS data that provides users with a more convenient access to GIS resources and guarantees them to always access the most up-to-date data.

Fourth, during the production and utilization of the NYS GIS web services, different stakeholders have indicated different facilitators and obstacles that they encountered in the past and/or they are currently experiencing. In this case, the providers, namely the NYS state agencies, especially the NYS GPO, is a great example of government that perceived the benefit of OGD and committed to improve serving the community through the NYS GIS web services (Altayar, 2018; Chorley, 2017; Kučera, 2017; Wang & Lo, 2016). Whereas some local governments are concerned about the unpredictable impact of OGD and privacy issues related to OGD development and are unwilling to participate in OGD development (Ingrams, 2017; Kučera, 2017; Ruijer et al., 2017a; 2017b), which in this case is that some county governments are not willing to join the open tax parcel data program. As the main providers of the web services, interviewees from the governments also emphasized on the importance of technical infrastructure, skilled staff and funding in OGD provision, which was also frequently mentioned as factors that affect the capacity and capability for OGD development in previous research (Beno, et al., 2017; Maccani, 2016; Ruijer et al., 2017a).
In this study, from the users’ perspectives, data related factors are the most important factors that affect their experience of using the NYS GIS web services, which aligns with the finding of my literature review. The availability, accessibility and quality of the NYS GIS web services are mentioned by most of the interviewees as the facilitators of their utilization. These are also indicators of data openness and data liquidity mentioned by Jetzek (2016) and Jetzek et al (2019) as factors affect OGD utilization and impact generation.

Users’ perceived benefits of the web services is another factor related to users that makes many interviewees choose to use the NYS GIS web services. This supports and extends the finding of the previous research that utilizes the TAM and the UTAUT models to assess users’ intentions to use OGD as an approximation to the actual OGD utilization (Anneke & Martijn, 2016; Maccani, 2016; Saxena & Janssen, 2017; Susha et al., 2015a; Talukder et al., 2018; Weerakkody et al., 2017b; Wirtz et al., 2017; Wirtz et al., 2018; Zuiderwijk et al., 2015b). Similar to previous findings (Dechman & Syms, 2014; Smith & Sandberg, 2018; Wirtz et al., 2018), a certain level of digital and GIS skills is also proven to help with the utilization of the web services. Compared to previous studies, a new finding from our study is that most of the users are users within organizational environments, the technology infrastructures within the organization also affect users’ experience using the web services. This has not been much mentioned by previous research.

The community factors mentioned in this case refers to responsive data providers that help address users’ questions in their utilization of web services, which has not reached to the level of increased engagement among stakeholders that lead to co-production and co-creation as it indicated by previous research (Hivon & Titah, 2017; Kučera, 2017; Lee et al., 2015; Maccani, 2016). However, the popularity and frequent use of different kinds of communication channels
indeed help keep users informed about the development of the web services and provide additional support (webinar, workshop, etc.) to improve their skills in using the web services, which aligns with the statement made by Susha et al (2015a) and Gascó-Hernández et al (2018) that tutorials and training are methods that help people improve their awareness and skills in OGD utilization.

Fifth, about the benefits, unlike previous research more talk about the potential OGD impact, the findings of the study reveal the benefits of using the NYS GIS web services, which actually fit the categories mentioned in previous research studies, namely the OGD political impact, economic impact and social impact (Alexopoulos et al., 2016; Attard, et al., 2016; Gurin, 2014). On top of this, this study has provided more tangible evidence of the generation of these different types of OGD impact. The political impact mainly shows that the availability of this open GIS resource helps improve citizen’s trust to government and improve government transparency. Due to the easy access and high quality of the web services, using it in government work also helps improve the efficiency and effectiveness in delivering public services. The economic impact mainly shows in reduced costs through avoiding redundancy in government work using shared GIS data. Using these web services as communication tools also helps facilitate discussion and decision making in community development and economic development. The social impact mainly shows as the increased citizen engagement in civic discussion about community issues.

Besides the beneficial impacts, interviewees also mentioned the limitations of the web services that constrain their utilization and prevent maximizing the benefits returned through the efforts that have been devoted in developing this resource. They also provide several suggestions for its future development. Based on all these findings, we propose the following suggestions for
stakeholders to take into consideration to help achieve the sustainable development of the NYS GIS web services or the NYS GIS resources in general:

First, for the future development of the NYS GIS web services, it is necessary to continue to improve the quantity and quality of the data. About the quantity, it refers to having more GIS web feature services available, which can give users more usability and flexibility to utilize the data for their own purposes. About the quality, it means that more work is needed to address issues related to data completeness and metadata, which are some of the most mentioned obstacles by users that negatively affect their perceptions of the usability of the data.

Second, for the future development of the NYS GIS Clearinghouse, the state has been collaboratively discussing with stakeholders from other organizations about the rebuilding of the NYS Clearinghouse. Two things mentioned a lot by the interviewees need extra attention. First, is to refine the limited search ability of the Clearinghouse with adding direct search, fuzzy search and voice search function for users to better find the data they need. The other one is that, since the NYS GIS web services are limited in conducting extensive research, it will be nice to build a bridge between the NYS GIS web services and the related downloadable data for users to choose from based on their purposes. If the web services are not sufficient, they can easily turn to use the downloadable data instead.

Third, for governments at different levels, there is a need to continue refining the constraints of their bureaucracy and traditional mindsets. It is also important to have more governments, especially local governments, to get on board with opening up their data to the public. From this case, we get to know that collaboration among different stakeholders provides great benefits in OGD development but also encounters great challenges. In order to form a more effective collaboration among the stakeholders, the state governments in this case as the leader in
the development of the NYS GIS web services, should continue outreach to more stakeholders that have not been included, to understand their needs, their limitations and challenges in OGD development, and provide assistance and support to improve their awareness of the NYS GIS web services. Financial and technical support in building and updating infrastructures should be provided to improve certain local government agencies’ capabilities in OGD provision. Tutorial and training should be provided to help improve users’ awareness and skills in utilizing the NYS GIS web services and other open GIS resources.
CHAPTER SIX

6 DISCUSSION AND CONCLUSION

This chapter offers some concluding remarks. It summarizes the conclusion of this dissertation through comparatively discussing the results of the quantitative and the qualitative study. Based on the findings, seven practical implications are raised to help guide the OGD development in the future. Limitations and directions for future research are also discussed.

6.1 Comparative Discussion of Quantitative and Qualitative Study and Conclusion

This dissertation was conducted to better understand the process of OGD utilization, as well as the roles of different influential factors and stakeholders in OGD utilization and impact generation. Based on a comprehensive review on current research studies, a quantitative study was designed to answer the question about “what” by assessing what are the most influential factors that affect the generation of OGD impact (RQ1), which have been mentioned often at the theoretical level, but much less examined using empirical evidence. Simultaneously, a qualitative study was conducted to answer the question about “how” by exploring how these factors affect the process of OGD impact generation. Focusing on a specific case in a specific country, the qualitative study investigated the process of OGD utilization by capturing the interaction among stakeholders (RQ2) and their perceptions of the facilitators and obstacles in the process of OGD utilization (RQ3). The qualitative study also revealed the benefits that have been generated through stakeholders’ different interaction (production and utilization) with the OGD. All of these have been limited explored by previous studies.

The results of these two studies have helped enrich the original conceptual framework (See Figure 9) with more empirical evidence (See Figure 18). It not only revealed the importance but also the diversity of the influential factors on OGD utilization and impact generation from
the perspectives of different stakeholders (boxes in yellow). It also captured the interaction among different types of stakeholders (producers, users/infomediaries, users/infomediaries and end users) in their different interaction with the OGD (production and utilization) (boxes in green).

![Diagram showing the perspectives of different stakeholders and their interactions with the OGD](image)

**Figure 18.** Enriched framework based on the quantitative and qualitative Study.

---

22 The texts in blue indicate the similarity of the findings in the quantitative study and the qualitative study.
Focusing on the RQ1, the results of the quantitative data analysis suggested that the well-resourced and supportive government actions was the most significant factor that affected the generation of different types of OGD impact. The results also indicated that some influential factors were only significant for the generation of a certain type of OGD impact. For example, the well-defined OGD policies and high quality of OGD related to accountability were only found to have a significant positive impact on the OGD political impact, the user participation was only found to have a significant positive impact on the OGD social impact and the OGD training was only found to have a significant positive influence on the OGD economic impact. The qualitative study also showed that some factors tested in the quantitative study also affected the OGD utilization in general, such as the well-resourced and supportive government action, OGD quality and OGD training. However, unlike the general influence of the factors on OGD impact identified from the quantitative study, the findings of the qualitative study showed the influence of these factors are more diverse as that different OGD stakeholders (the NYS GIS web services producers vs. users/infomediaries) had different opinions on the factors that affected their different interaction (production vs. utilization) with the OGD.

Focusing on Q2 and Q3, the findings of the qualitative study indicated that different stakeholders were involved in the production and utilization of the NYS GIS web services. Their perceptions of influential factors, including facilitators and obstacles on OGD production and utilization were different due to their different roles (producers vs. users/infomediaries) during their different interaction (production vs. utilization) with the NYS GIS web services.

First, during the production of the NYS GIS web services, governments aimed to empower public utilization of GIS resources for better decision making in community development and improve the efficiency in public service delivery. As the main providers, state
government agencies in New York were motivated to publish the NYS GIS web services and made it freely available to the public. With the help from the local governments in data sharing and data updating, the state government agencies consolidated the regional GIS data into statewide GIS web services and kept it up to date, which could also be viewed as a type of OGD utilization: data to data. As indicated in section 5.4.2.3, formal government-led work groups, GIS associations and user groups have enabled different users/infomediaries from governments, non-profits, for-profits and academic institutions to frequently participate in the discussion about the design the NYS GIS Clearinghouse, the provision of the NYS GIS web services and the GIS development in the NYS in general. There was also frequent informal communication that facilitated the collaboration between the providers and users/infomediaries in the production of the NYS GIS web services. The coordination and collaboration among these stakeholders facilitated the information sharing, which helped enable the development of NYS GIS web services and other GIS based products, reduce redundancy in government work on data collection and management, and ultimately improve government efficiency. However, the collaboration also faced challenges like uncooperative stakeholders in information sharing, the lack of fundings, the lack of staff and imitated staff in data management and bureaucracy that constrained the collaborations across departments or organizations.

Based on the findings in section 5.4.3.1, we can see that for the providers, they more noticed the factors related to themselves that facilitated or disturbed the production of the NYS GIS web services. About the facilitators, the state governments, especially the NYS GPO had a great awareness of the potential benefits of OGD, a clear vision of its OGD development and has devoted great resources (e.g., fundings, technology infrastructure, staff time and skills) to support the development of the NYS GIS web services. The frequent and effective
communication and collaboration with other stakeholders also helped better address issues related to the production of the NYS GIS web services or GIS development in the NYS in general. In the meantime, the providers also encountered challenges that were changing over time during the process of producing the NYS GIS web services, such as unwillingness of some local governments in data sharing, securing sustainable funding to support upgrades in technology infrastructure and keeping data up to date.

Second, during the process of utilizing the NYS GIS web services, various users, including users from the state governments, local governments, nonprofit organizations, profit organizations and academic institutions have been actively involved in using the NYS GIS web services for their work and research. Most users were motivated to use the web services due to its availability, easy access and high quality. And compared to other formats of GIS data, web services also allowed users to access the GIS data without downloading it, which greatly saved users’ server spaces and efforts in data maintenance and management. Different users have used the NYS GIS web services in different ways. In most of the cases, the NYS GIS web services were used to create maps for presentations, reports and web pages (data to information). And the users/infomediaries often used them as tools to facilitate the communication and collaboration with the end users (e.g., clients, policy makers, citizens) in the discussions related to urban planning and community development. These web services were also used as references for users to quickly look up information (data to fact) or as the basic element to build web applications (data to application). During the utilization process, there was limited communication between the users/infomediaries and the providers, the users just went and grabbed what they needed from the catalogue of the web services. Occasionally, the users would reach out to the providers
to acquire data they needed but were not listed on the Clearinghouse or asked questions related to the dataset they were interested in.

Based on the content in section 5.4.2, we can see that for the users/infomediaries, they more mentioned data-related factors and user-related factors that affected their utilization of the NYS GIS web services. The most mentioned data-related factors were the availability, accessibility and the quality of the data. Due to the high quality of the data, more and more users/infomediaries were attracted to use these resources in their work and research. However, any issues related to the data quality also made them change to use other types of GIS data instead, such as the downloadable datasets. The most mentioned user-related factors were the user perceived benefits and user skills. If users recognized the advantages of the NYS GIS web services compared to the traditional downloadable GIS data, and they had the skills to utilize this new method of accessing data, they were more likely to choose using this resource. However, if users felt the NYS GIS web services were limited for their purposes of use, or they were intimidated by the initial learning hurdle of using this resource, they would stay with the traditional method of accessing the GIS data: downloading and managing the data on their own.

It is worth mentioning that since the users in this case were all organizational users, the availability of suitable technology infrastructures also affected users’ utilization of the NYS GIS web services. For example, the speed and stableness of the Internet was mentioned as one of the obstacles that negatively affected the accessibility of the NYS GIS web services.

The qualitative study also showed that different benefits have been generated through utilizing the NYS GIS web services. First, political impact, which was most mentioned by interviewees from governments at the different levels. The NYS GIS web services allowed users to directly access GIS data and to get to know about government related information at first
hand. It helped improve government transparency and reduce the general publics’ lack of trust in governments. In the meantime, different government agencies have been greatly utilizing this GIS data without collecting or managing the data on their own or concern about the data quality. It greatly improved the effectiveness and efficiency in delivering public services. Second, economic impact, which was mentioned the most by users from different types of organizations. The availability of different GIS web services made it possible for users to access this valuable, high-quality, and most up-to-date data that were not available anywhere else for free, and with no need to download them. It helped users minimize the duplication of efforts and investment in data collection, storage and management. It enabled the creation of tools to facilitate communication and collaboration among community stakeholders to discuss issues about community development. This authoritative and accurate information also helped reduce misunderstanding among stakeholders and improve the effectiveness in decision making. Third, social impact, which was least mentioned in this study. The availability of the NYS GIS web services has empowered the public through putting high quality GIS data in their hands to help them better understand the issues they’re trying to address around their communities and improve their engagement in community discussion.

Based on the feedback, we also noticed that the same type of open GIS data has been used by different users and generated different types of Impact. For example, tax parcel web services in Rensselaer County were used by companies and local residents and have generated economic and social impact as I discussed in section 5.4.4.1. Compared to the results of the quantitative studies, that OGD related to accountability only showed significant influence on OGD political impact, in this case, open GIS data, as an example of OGD related to innovation showed a connection with the generation of OGD economic and social impact.
In summary, through a multi-method approach, this dissertation has conducted two research studies to better understand the OGD utilization and impact generation, but with different focuses. The quantitative study focused on assessing the influence of different factors on the generation of different types of OGD impact in general, whereas the qualitative study focused on exploring the process of OGD utilization in a specific case to better understand how these factors affect the OGD utilization from the perspectives of different stakeholders. This study investigated the interaction among stakeholders, the stakeholders’ perceived facilitators and obstacles in the process of producing the utilizing the NYS GIS web services, the types of utilization and the benefits of this open GIS resources that has been generated through their utilization. These two studies together have provided general and specific empirical evidence to help us better understand the OGD utilization and impact, practical implications are raised to help guide OGD development in the future.

6.2 Implications

Based on the results of the quantitative and qualitative study, the following are seven practical implications to be considered for the future development of OGD.

Implication 1: Challenges in collaboration among different levels of governments in OGD provision. The quantity and quality of data provided through OGD projects depends frequently on collaboration across levels of government, which will further affect the OGD utilization and the generation of the expected OGD impact. Governments as the main stakeholders in OGD provision, have experienced many organizational and cultural challenges in the process of opening up data. From the qualitative study, we can see that governments at different levels were involved in OGD provision but to different extents. The state government agencies were the main leaders in publishing the NYS GIS web services, whereas the local
governments mainly assisted this process through updating the data regularly and sharing data to the state. However, due to the difference in the perceptions and capabilities of the governments, especially local governments, the cooperation of relevant units cannot be assumed. Time, efforts, and outreach strategies are needed to address these challenges, such as frequent discussion and physically visiting to better understand potential collaborators’ expertise, limitations and needs, and presenting the benefits of the collaboration that have been generated to motivate more collaborators to be involved.

Implication 2: Well-resourced and supportive government actions are important to increase OGD utilization and impact generation. In the quantitative study, the well-resourced and supportive government actions were operationalized as the extent of implementation of OGD initiative at different levels of government, the availability of consistent OGD management and publication approach and the availability of government support in OGD competitions and grants. It was predicted to have a significant positive influence on the generation of OGD political, economic and social impact. And from the qualitative study, we also get to know about the importance of government resources and support in OGD provision and utilization. Many state agencies have been collaborating with the local governments to open up high quality GIS data to the public and make sure the data is always up to date through regular maintenance and update. The high quality of OGD was mentioned by the users as one of the facilitators in their utilization of the web services. All these findings support the findings that government resources and efforts in implementation of OGD initiative at different levels of governments through OGD provision, and the availability of consistent OGD management and publication approach are facilitators for OGD utilization and impact generation. During the process of utilization, users also indicated that the lack of awareness or the initial learning hurdle has negatively affected
their utilization of the web services, especially for new users. Based on this, I believe that more government actions in supporting OGD utilization are needed, especially in the form of government-led webinars, workshops, training, and other supportive activities to help improve users’ awareness and alleviate the learning hurdle they have in OGD utilization. In summary, in order to facilitate OGD utilization and impact generation, governments should realize that their responsibilities are more than just making the data available, they also share the responsibilities in improving user engagement through better promoting the value of OGD and helping improving users’ awareness and capabilities to use the OGD.

Implication 3: Improved communication between OGD providers and OGD users helps improve OGD utilization. From the qualitative study, we get to know that the availability of informal and formal communication channel has greatly facilitated the coordination and collaboration between OGD providers and OGD users during the process of OGD provision, more specifically, it helped improve the information sharing, reduce redundancy in government work on data collection and management, and ultimately improve government efficiency. However, during the process of OGD utilization, the interaction and communication among stakeholders was quite limited even with the different communication channels being available, such as the email listserv and contact information of the data providers. Based on this, I believe that in order to improve the communication between OGD providers and users, it is not only necessary to have the communication channels, but more importantly, think about how to improve the communication experience and quality to help providers better promote the OGD, better understand and address users’ needs. On the other side, governments also need to take actions to advocate users to express their issues or needs about OGD, such as adding more interactive components on the platform of the NYS Clearinghouse to allow users to grade or
comment on datasets, share their experience of using these open GIS resources, and discuss their issues in utilization and inquiries of other resources. The improved information exchange will help guide governments to provide more useful OGD that can be more utilized by more users, and ultimately lead to generate more beneficial impact.

Implication 4: The important role of community organizations in facilitating OGD utilization. As indicated in the qualitative study, at the current stage, compared to governments who are the main providers of the OGD, community organizations, such as user groups, associations in different subject domains (in this case, GIS associations) are playing an active role in engaging users to improve their awareness and skills in accessing and using OGD. They have created a welcoming environment that can engage with a bigger group of users in learning and discussing the development of OGD in their subject of interest. In order to improve OGD utilization and facilitate the generation of OGD impact, I believe that more communication and collaboration between the governments and community organizations are needed. The governments are the leading efforts in providing open data resources and other supportive resources, such as fundings, grants to enable OGD utilization and innovation. With collaborating with community organizations, governments can take advantage of the popularity of the community organizations among OGD users to attract more diverse users. Their joint efforts will provide a great support to improve users’ awareness of OGD and their capabilities of using OGD.

Implication 5: The influence of organizational infrastructure on OGD utilization. Different from previous studies that mostly discussed the influential factors of OGD utilization related to users themselves, from this qualitative study, we got to know many users were using the NYS GIS web services within a specific organizational environment and mainly used it for
their jobs. Besides their awareness and skills in accessing and using the OGD, the technology infrastructure of the organization was also raised as an influential factor that affected the users’ experience in using the NYS GIS web services. Organizations with limited technology infrastructures should also actively look for partners, and through collaborative efforts to help create opportunities and provide support that enable themselves to improve their technology infrastructures, so that they can better support the OGD utilization for their work. It is also important to have higher levels of governments (e.g., federal governments, state governments) who have abundant resources to step in and provide supportive fundings and grants to support local governments’ transformation and infrastructure upgrade.

Implication 6: Different factors are recognized by different stakeholders at different stages of OGD development. The influential factors of OGD utilization and impact mentioned in previous research studies can be generally categorized as government related factors, OGD related factors, user related factors and community related factors. However, little is known about how different stakeholders perceive the roles of these different influential factors based on their interaction with the OGD. From the qualitative study, we can see that producers and users had quite different opinions on the factors that affected their production and utilization of OGD. In the case of the production of the NYS GIS web services, state government agencies, the main provider of the web services, recognized the key factors that are most related to their own capacities: such as funding, staff skills, communication, and collaboration with other stakeholders. In terms of the utilization of the NYS GIS web services, users have paid more attention to data related factors such as the availability, accessibility and the quality of the data, as well as user related factors, such as the perceived benefits and user skills. Understanding the different perceptions of influential factors by stakeholders at different stages of OGD
development has the potential to support the allocation of efforts and resources when trying to improve stakeholders’ interaction with OGD at a specific stage.

Implication 7: It is important to better track OGD utilization and measure OGD impact.

About OGD utilization, the qualitative study has captured different types of OGD utilization by different types of users. About the OGD impact, the quantitative study used expert opinions to evaluate it, whereas the qualitative study used stakeholders’ feedback to measure it. By doing so, they provide more empirical evidence that helps us better understand who are the users of OGD? How are OGD often used by different types of users? What limitations or issues are there in the process of utilization? What benefits or impacts are generated through using OGD? Through collecting and presenting this evidence, it helps guide governments to provide useful OGD that more suits users’ needs, which helps improve the effectiveness in OGD utilization and impact generation. The increased realization of OGD impact also helps advocate and encourage more governments and other organizations to participate in OGD initiatives. This increased government involvement in OGD development will better leverage the OGD to improve government transparency, enable innovation in public services and business products, and to facilitate the development of the economy and the improvement of the quality of life. From the qualitative study, we get to know that most of the governments OGD providers had a general idea of the amount of usage of their data through tracking and recording the traffic of their OGD platform. However, there was lack of further analysis of the users’ activities, mostly likely due to the lack of resources, especially staff time to dedicate on this matter. In order to better track OGD utilization and impact generation, governments should not only improve their awareness of this issue, but more importantly, investments in technologies and staff time are needed for
governments to better understand the users’ utilization experience, which can help guide them to improve their efficiency and effectiveness in OGD provision.

6.3 Limitations and Future Research

This dissertation has provided general and specific empirical evidence that help us better understand the OGD utilization with focusing on different influential factors and stakeholders. However, there are some limitations in each of the studies.

The quantitative research has provided new empirical evidence about influential factors of OGD impact generation. The research model was also designed in an innovative way to test the influence of different types of OGD on the generation of different types of OGD impacts. However, it has several limitations. First, the research model was constructed based on extensive literature review but was limited by the availability of data. There were many other influential factors (See Table 5) that have not yet been included in the model, since there are no appropriate indicators for them, and the analysis was also conducted based on limited sample size. Second, among previous research, some factors were mentioned related to different aspects of the OGD development, such as the adoption of OGD initiatives, the development of OGD initiative, the utilization of OGD, and the impact of OGD generation. In this case, we constructed the model in a simplest way by listing all the factors included as predictors in our models without testing other possible paths of the influence, for example, the user participation as a mediator in the models. Third, in this study, we only focused on testing the influence of the properties of OGD in different subject domains on one specific type of OGD impact without testing their influence on different types of OGD impact.

For future study, measurement should be designed, data should be collected for other factors and more factors should be included in the model to test and provide more new insights.
to the findings. A test using a larger sample is needed to yield more generalizable findings. The refinement on the construct of the model is also needed to test other possible paths of the influence. Last but not least, some variables are notoriously difficult to measure. For example, in this case, OGD experts’ evaluation based on claims made in credible sources concerning possible impact of open data was used as proxy indicators of areas in which impact may be occurring. For future study, more discussion and research should be conducted to better track and measure OGD impact.

Focusing on a specific case, the qualitative case study has explored the process of OGD utilization by capturing the interaction among different stakeholders and the facilitators, obstacles, and benefits in OGD utilization from their perspectives. However, it also has some limitations. First, I tried to include as many and as diverse stakeholders as possible in this study, however, I was not able to find and directly interview the end users of the web services, who were usually using web services-based products, such as web pages, reports, information viewer, etc. Second, this study only focused on a specific type of OGD, the NYS GIS web services. All the findings are based on the interviewees in NYS that share their experience specifically related to the NYS GIS web services. Ideally, to truly reveal the benefits of the multi-method approach used in this dissertation, more case studies should be conducted in each of the countries that were included in the quantitative study.

For future research, interviews directly with the end users should be conducted to help add more insights about their perceptions of the facilitators and obstacles in utilizing OGD or OGD based products. More studies should be done to explore the utilization of open GIS data in different contexts or on other types of OGD to help test and refine the findings of this study, and
to make it more useful to guide the future of OGD utilization, impact generation and development of OGD initiatives in general.
REFERENCES


Determinants of Continuance Intention to Use Open Data Website: An Insight from Indonesia. *Pacific Asia Journal of the Association for Information Systems, 11*(2), 96-120. https://doi.org/10.17705/1pais.11205


Jiang, Shao, Liou, Shao, & Shi. (2019). Improving the Sustainability of Open Government Data. *Sustainability, 11*(8), 2388. [https://doi.org/10.3390/su11082388](https://doi.org/10.3390/su11082388)


https://doi.org/10.1016/j.giq.2018.11.001

https://doi.org/10.1145/2756542


Torres-Reyna, O. (2007). Panel data analysis fixed and random effects using Stata (v. 4.2). Data & Statistical Services, Princeton University.


Weckström, C., Kujala, R., Mladenović, M. N., & Saramäki, J. (2019). Assessment of large-scale
https://doi.org/10.1016/j.jtrangeo.2019.102470
https://doi.org/10.1177/0020852317719996
World Bank: Open Data in 60 Seconds. (2019). Retrieved from


government data objectives: Is there a mismatch between objectives and results?
https://doi.org/10.1177/0020852317739115.

https://doi.org/10.1016/j.ijinfomgt.2019.05.024

APPENDIX

Appendix A1. Example of Measurement of Independent Variables adopted from the ODB

Expert Survey

<table>
<thead>
<tr>
<th>Indicators</th>
<th>United States (2017)</th>
</tr>
</thead>
<tbody>
<tr>
<td>To what extent does the country have a functioning Right to Information (RTI) / Freedom of Information (FoI) law?</td>
<td>Score: 5</td>
</tr>
<tr>
<td></td>
<td><strong>Explanation by experts:</strong> The Freedom of Information Act (1) states that any person &quot;has the right to request access to federal agency records or information except to the extent the records are protected from disclosure by any of nine exemptions contained in the law or by one of three special law enforcement record exclusions&quot; (2). Its various exemptions are listed on the US Department of State website (2). More information on submitting a request, and FOIA appeals can be found on the Department of Justice website (3). Users can also look at detailed data on each agency's FOIA activities on FOIA.gov (4). In 2015, the Obama administration has been noted to withhold a large number of FOI requests (5). In addition, the Obama administration sets record for unfulfilled FOIA requests, as the &quot;U.S. government in 2015 came up empty-handed when asked for public records more than 100,000 times, The Associated Press disclosed.&quot; (6). In 2015, the White House announced it was exempting itself from records requests (7). Since 2009, the US government has confronted whistleblowers and official leakers. There have been eight prosecutions under the 1917 Espionage Act â€“ more than double those under all previous US presidents combined (8).</td>
</tr>
</tbody>
</table>

## Sources
(2): [http://foia.state.gov/Learn/FOIA.aspx](http://foia.state.gov/Learn/FOIA.aspx)
(7): Josh Hicks, The Washington Post, A Sunshine Week surprise:
<table>
<thead>
<tr>
<th>Score: 7</th>
<th><strong>To what extent is there an active and well-resourced open government data initiative in the country?</strong></th>
</tr>
</thead>
</table>

**Explanation by experts:**

There is an open government data initiative at the federal level. The offices involved are Office of Management and Budget (Office of the Federal Chief Information Officer), and the U.S. General Services Administration (Data.gov Program Management Office), as well as open data leads at the participating federal agencies (11).

Under the current administration, as of 30 October 2017, there has been a delay in releasing an Open Government Partnership (OGP) National Action Plan (NAP) to articulate goals for increasing government transparency and accountability from the US Office of the Federal Chief Information Officer (1), even though President Trump committed participating in OGP in September 2017 (12).

The NAP was scheduled for release in October and required as part of the country's participation in the OGP. The US, a founding member of OGP, has submitted three national actions plans so far (2).

In the US, a top-level commitment towards open data and a very strongly documented open data initiative was made and was ongoing since 2014, when President Obama signed an executive order (3) that made "open and machine-readable data the new default for government information" (4) (5) (6). The offices involved are Office of Management and Budget (E-Government), Office of Science and Technology Policy, and the U.S. General Services Administration (Data.gov Program Management Office), as well as open data leads at the participating federal agencies. In 2009, Open Government Directive required government data be made available in an accessible, timely manner (7).

The open government initiative includes a robust open data policy (8) and used to include an open government working group (9). The government has acted on the initiative having launched data.gov as well as a number of open data initiatives across different sectors such as health, energy and education (10).

The US Government participates in the Open Data 500 Global Network (11). The network has six members, including Australia, Mexico, the United States, Italy, Korea, and Canada. It meets
To what extent are civil society and/or information technology professionals engaging with the government?

**Score:** 9

**Explanation by experts:**

On 29 June 2017, the fourth annual DATA Act Summit explored the transformation of federal spending from disconnected documents into useful data. The Summit brought together the Congressional allies and administration officials who are driving these changes; the agency leaders who are using newly-standardized data to derive new insights; the nonprofit advocates who are supporting data-driven accountability; and the technology companies whose solutions are doing the work (1).

On 28 September 2016, the White House Open Data Innovation Summit brought together government trailblazers, entrepreneurs,
companies, advocates, and civic innovators who are using Federal open data across all sectors (2). The Innovation Summit included highlights of open data helping cancer patients, including sub-national efforts such as Washington DC's open data portal public safety and aerial imagery datasets, and of Maryland’s Montgomery County data uploads to CountyStat in real-time for crime, food inspection, and parking information (3).

The Code for America Summit (4) draws experts, citizens and government officials to discuss open data topics in an often interactive setting. On a sub-national level, there are a number of ongoing open data community groups that host regular meetups in Washington D.C. (5), Chicago (6), New York City (7), Delaware (8), San Francisco (9), and Boston (10).

The Knight Foundation, a non-profit foundation, has funded “Open data to Open Knowledge” (2015-2017), a project dedicated to transforming the city of Boston’s open data collection into an accessible resource for the public - engaging with a number of stakeholders, including local libraries, as well as residents, researchers and public employees (11).

GovLab is promotes engagement of civil society and open data experts with Government at their 2015 programme of open data projects includes: Network of Innovators, GovLab Academy, Data Collaboratives, Open Data 500 Global Network, and Open Gov Research Repository (12). Its Network of Innovators app is allowing users to search through government innovators on topics such as opening data, challenges and crowdsourcing for public good (13), encouraging knowledge sharing and networking in the open data space.

There is also feedback on Data.gov through multiple channels (14). Technical requests, such as requesting new features on Data.gov, are tracked through GitHub (15).

## Sources
(4): https://www.codeforamerica.org/summit
(6): http://www.smartchicagocollaborative.org/contact-us/meetups/
(7): http://www.meetup.com/NYC-Open-Data/
(8): http://www.meetup.com/Open-Data-Delaware/
To what extent is training about open data available for individuals or businesses who want to increase their technical skills or develop businesses to use open data?

Score: 9

Explanation by experts:
On 25 July 2017, the Office of Management and Budget and the Center for Open Data Enterprise co-hosted a Roundtable on Open Data for Economic Growth that included approximately 80 government and business leaders. The goal of the Roundtable was to connect government and private-sector leaders to use open data as a strategic resource for better government and growing the American economy, by developing a better enabling environment for the use of open data and identifying high-value data sets for business use (1), with a methodology that the Center for Open Data Enterprise has developed over three years to help the White House, federal agencies and their data users work together to improve open data, although this methodology is not open per se (2).

Sunlight Academy offers via GitHub a collection of interactive tutorials for journalists, activists, researchers and students to learn about tools by the Sunlight Foundation and others to unlock government data (3). Note that their Sunlight Labs has retired (4).

GovLab Academy is a training programme hosted by the GovLab, which includes coaching programs, workshops, courses, project clinics, and master classes (coming soon). It offers teams and individuals, inside and outside of government: Access to high-quality, curated content aimed at busy practitioners; Mentoring designed to help take public interest projects from idea to implementation; Connections to subject matter experts, who know whatâ€™s been tried and what does and does not work; and Tools for peer-to-peer learning and skill sharing (5). In addition, GovLab launched a Network of Innovators app is a skill-sharing network for government and civic innovators worldwide, which allows individuals to build their own network of innovators and find the right collaborators for projects they are working on (6).

American universities also offer a range of courses in different areas of open data. University of Washington offers Data Science at Scale Specialization through Coursera (7). Database Design & Implementation courses are available at the University of
Washington (8) and Berkeley (9). Berkeley’s School of Information (10) offers courses includes Data Visualization and Communication (11). Harvard co-hosts a MOOC on Big Data: Mathematical Modelling (12) (13).

GovEx and the Data Points Podcast: GovEx is part of Johns Hopkins University’s 21st Century Cities Initiative and partner in What Works Cities. Launched by Bloomberg Philanthropies in April 2015, What Works Cities is a national initiative to help 100 mid-sized American cities enhance their use of data and evidence to improve services, inform local decision-making, and engage residents (14) (15).

## Sources

(2): http://opendataenterprise.org/what-we-do
(4): https://sunlightfoundation.com/2016/10/12/how-were-wrapping-up-sunlight-labs-in-the-open/
(5): http://govlabacademy.org/
(6): https://networkofinnovators.org/
(7): https://www.coursera.org/course/dataset
(8): http://www.pce.uw.edu/courses/database-design-implementation.html
(9):
http://extension.berkeley.edu/search/publicCourseSearchDetails.do?method=load&courseId=40918
(10): http://www.ischool.berkeley.edu/
(11): http://www.ischool.berkeley.edu/courses/dataset
(12): https://www.class-central.com/university/harvard
(14): In-person interview with Eric Reese, GovEx - Johns Hopkins. 8 October 2016.
(15): http://govex.jhu.edu/resource-type/podcast/
Appendix A2. Examples of OGD Assessment

| Dimensions of OGD Assessment | United States (2017) Dataset – Budget
|                            | Total score: 60 | United States (2017) Dataset – Health Sector Performance
|                            | Score:5         | Total score: 90 |
| a - Does the data exist?     | Score:10        | Score:15        |
| b - Is it available online from government in any form? | Score:0 | Score:15 |
| c - Is the dataset provided in machine-readable formats? | Score:5 | Supportive evidence: No, the dataset format is pdf, zip. |
| d - Is the machine-readable data available in bulk? | Score:10 | Score:10 |
| e - Is the dataset available free of charge? | Score:0 | Score:15 |
| f - Is the data openly licensed? | Score:15 | Score:15 |
| g - Is the dataset up to date? | Score:10 | Score:10 |
| h - Is the publication of the dataset sustainable? | Score:5 | Score:5 |
| i - Was it easy to find information about this dataset? | Score:0 | Supportive evidence: GPO --> Document: Analytical Perspectives BUDGET OF THE U. S. GOVERNMENT Fiscal Year 2018 --> link in doc to White House www.budget.gov/budget which has been renamed https://www.whitehouse.gov/omb Descriptive metadata: MODS, Authenticity Metadata: PREMIS, All Format & Metadata Files: ZIP file; guidance documentation: |
| j - Are (linked) data URIs provided for key elements of the data? | Score:0 | Score:0 |
Appendix A3. Examples of Measurement of Dependent Variables adopted from the ODB

Expert Survey

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Political Impact 1:</strong></td>
<td></td>
</tr>
<tr>
<td>Increasing government efficiency</td>
<td>Score: 2</td>
</tr>
<tr>
<td>and effectiveness</td>
<td><strong>Explanation by experts:</strong></td>
</tr>
<tr>
<td></td>
<td>In July 2017, The Sunlight Foundation has been critical of federal agencies, noting that they have been slipping on their open data responsibilities - where the Foundation had &quot;to almost sue&quot; the Obama administration to release enterprise data inventories. In addition, The Department of Treasury, The Department of the Interior, The Department of Housing and Urban Development, The Department of Veteran Affairs and The Department of Homeland Security have yet to release their plans for a more open government (1). In its impact section, data.gov provides examples of citizens leveraging open data, however it does not endorse them (2). &quot;At the local and national level, agencies and government departments share their data with other departments, across departments, and have realized the value of that. For example, at the local level, 311 data to publish and use across the government and unify systems across geographies. Open data is used by internal users much more than in the past&quot;, stated Eric Reese (GovEx) (9). The Center for Open Data Enterprise has published a nonpartisan informational Open Data Transition Report (3) that draws upon experts from business, the public sector, and civil society to provide an action plan for continuity and further improvements in open government data that shows how open data can fuel national initiatives in healthcare, medical research, energy, criminal justice reform, education, labor, veteran benefits, for example (4). In 2016, the State of the Union of Open Data report (5) was published by Data Foundation, US first industry focused open data research organization (6). Centers for Medicare and Medicaid Services (CMS) launched in 2015. Currently (September 2017), users can explore and access more than 28 million records of $17 billion in payments (from 2013 to 2015) in order to make more informed healthcare decisions. US Agency for International Development (USAID) and the Federal Communications Commission (FCC) are also freely sharing data both to improve the efficiency and management of their programs and to provide the public with better information for decision-making (5). In 2015, Pew Research Center's Internet &amp; American Life Project</td>
</tr>
</tbody>
</table>
published research on Americans' views on open government data (7). The research found that overall, Americans are optimistic that open data initiatives can make government more accountable, however many Americans remain skeptical about open data's ability to improve government performance (7). The research finds that overall, although open data initiatives have reached many Americans due to their active online lives, most Americans "have yet to delve too deeply into government data and its possibilities to closely monitor government performance" (7). No follow up study on the topic has been conducted since.

It is worth highlighting that the US Government Accountability Office has conducted research on federal data transparency and have found "persistent challenges related to the quality and completeness of the spending data agencies report to USAspending.gov". For example, GAO reported in June 2014 that roughly $619 billion in assistance awards were not properly reported. The GAO has found that effective implementation of the Digital Accountability and Transparency Act of 2014 (DATA Act), which was enacted to help address these challenges, would "help promote transparency to the public and address ongoing government management challenges by expanding the quality and availability of federal spending data" (8).

Several datasets provided by the General Services Administration (GSA) such as the Per Diem API (lodging and per diem rates for government travel to 332 U.S. locations) and the .gov (the list of all .gov domains, including state, local, and Tribal .gov domains) are frequently used and requested. These datasets are valuable mainly for other government agency use.

Overall, there is still limited evidence of open data having a noticeable impact on increasing government efficiency and effectiveness.

### Sources
(2): http://www.data.gov/impact/
(3): http://opendataenterprise.org/transition-report
(5): https://static1.squarespace.com/static/56534df0e4b0c2babdb6644d/t/5818c71d1b631b707adc8763/1478018860740/the-state-of-the-union-of-open-data.pdf
(7): http://www.pewinternet.org/2015/04/21/open-government-data/
Poetical Impact
2: Increasing transparency and accountability

Score: 2:
Explanation by experts:
The Trump administration has been under scrutiny regarding their transparency and accountability efforts, especially by civil society actors in the country, such as Sunlight Foundation: "there is a regression to secrecy in both Congress and the White House. The change has not gone unnoticed around the globe, as our national standing to defend democracy and our government’s ability to advocate for anti-corruption efforts has been precipitously eroded." (1).
According to Eric Reese (GovEx), "Sunlight Foundation and others are being hyper aware of changes, where they might not have been in the past."

Data.gov has an ethics datasets catalog to access the full version of ethics datasets from across the government, dedicated to government accountability (7). However, there has been much criticism around the ethical responsibility of the Administration and President Trump (e.g., tax returns disclosure, never fully divested from his businesses, White House issued secret waivers issued for administration officials, embraced nepotism at the highest reaches of the White House, White House has frequently at odds with the nonpartisan Office of Government Ethics) (1).
The US opted out of the Extractive Industries Transparency Initiative (EITI) (2) (3), three years after originally joining it in 2014 (4). The Interior Department said the State Department, now headed by former ExxonMobil CEO Rex Tillerson, will continue to "lead the U.S. commitment" as a "supporting country," but all its decisions with regard to the pact's standards will be done unilaterally (2)(3).

There is progress on releasing spending data and federal data transparency in terms of ongoing transparency and accountability commitments with the DATA Act (5) (6). On 9 May 2017, federal agencies officially began reporting data in compliance with the open standards created under the DATA Act, the landmark legislation that cleared Congress in 2014. The data has transferred to beta.usaspending.gov and will eventually be transitioned to USAspending.gov, the long-term online home of federal spending data. In doing so, government will give citizens, watchdogs, Congress and federal workers unprecedented public access to structured information about spending and open up new horizons for oversight, accountability, activism and innovation. As previously acknowledged by the US Government Accountability Office (GAO), there are "persistent challenges related to the quality and completeness of the spending data agencies report to USAspending.gov". GAO reported in June 2014 that roughly $619 billion in assistance awards were not properly reported.
GAO has found that effective implementation of the Digital Accountability and Transparency Act of 2014 (DATA Act), which was enacted to help address these challenges (13). On 29 January 2016 it was noted that in order to “improve the transparency and quality of the federal spending data made available to the public, the DATA Act directs the Office of Management and Budget (OMB) and the Department of the Treasury (Treasury) to establish government-wide data standards that include common data elements for reporting financial and payment information by May 2015. In general, the act requires agencies to begin reporting financial spending data using these standards by May 2017 and publicly post spending data in machine-readable formats by May 2018.” (14)

In 2017, the Police Data Initiative (PDI) represents 130 law enforcement agencies who have released more than 200 datasets and originated as a result of several recommendations in the Task Force on 21st Century Policing that focused on technology and transparency. In 2016, the Obama Administration announced that 53 jurisdictions, covering more than 41 million people, have committed to the PDI, with over 90 datasets released to date. This follows the 2015 commitment, where the White House announced that it will be releasing new datasets relating to police officers via its PDI, including the public release of 101 datasets (e.g., information on policy stops, use of force, officer-involved shootings) (8).

The FOIA app allows citizens to submit requests more easily (9). As a result of these efforts, there has been a steady reduction in the FOIA backlog. For example, since the beginning of Fiscal Year 2015, DHS has reduced its FOIA backlog by 20 percent, from 103,480 to 82,324 as of July 1, 2015 (9).

However, as of 4 May 2017, the Trump administration refused to release to the Asbury Park Press $783 million in payroll data for more than 10,000 federal employees working outside the United States, including employees working at U.S. territories like Puerto Rico and Guam, in federal records request the Press filed in January information that has been provided to the public for the last decade (16). The federal agency also refused to release most performance-based bonuses given to federal employees in fiscal year 2016 (16).

In addition, the Obama administration has been noted to withhold a large number of FOI requests (10), and sets record for unfulfilled FOIA requests, as the government in 2015 did not release public records for more than 100,000 requests (11). In 2015, the White House announced it was exempting itself from records requests (12). Moreover, in 2016, there were $36.2 million in legal costs fighting lawsuits filed by news organizations under the Freedom of Information Act, with the Justice Department accounted for $12 million, the Homeland Security Department for $6.3 million and the Pentagon for $4.8 million. The three departments
accounted for more than half the governments total records request (15).

### Sources
(7): https://ethics.data.gov/browse/
(9): http://www.dhs.gov/blog/2015/07/07/dhs-launches-efoia-app
(17): Skype interview with Eric Reese (GovEx, Johns Hopkins University), 11 January 2018.

<table>
<thead>
<tr>
<th>Social Impact 1: Noticeable impact on environmental sustainability</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Score:</strong> 3</td>
</tr>
<tr>
<td><strong>Explanation by experts:</strong></td>
</tr>
</tbody>
</table>
According to Eric Reese (GovEx), "even though the US government pulled out of the Paris climate change agreement at national level, sub-national governments still signed up; Bloomberg also funded these efforts (15) (16) (17).

Open Source Applications using US Environmental Protection Agency (EPA) Linked Data, including apps to: Locate Facilities by Zip Code, Maps of Electricity Generators, and Nuclear Power Plants. Note that all apps use the Callimachus platform (11). However, note that we can no longer download Daily Data from EPA (12).

Open Data 500 Companies lists several research institutions and companies in the US that have used open data to improve environmental sustainability (4).

CoolClimate: The CoolClimate Network (CCN) is a division of UC Berkeley's Renewable and Appropriate Energy Laboratory (RAEL). CCN carries out research on carbon footprint mitigation, develops carbon management tools for public, business and government users and implements best practice programs grounded in behavior research to engage business and individuals in voluntary carbon reductions. All calculations and data are fully transparent, published on our website and in peer-reviewed research (4) (5).

iRecycle: tells you how, where and when to recycle just about anything. Using your current location, ZIP code, address or city, get access to vital details for collection points, such as Web site, phone number, directions, hours of operation and other materials collected (4) (6).

WaterSmart Software helps water utilities engage with their customers to save water and money. WaterSmart deploys a turnkey, cloud-based platform for water utilities that employs sophisticated data analytics and behavioral science tools to reduce residential water demand. The solution is proven to reduce water consumption by up to 5% and improve customer satisfaction by nearly 200% (4) (7).

U.S. Climate Resilience Toolkit is quite involved in open data and list on their website tools to help you manage your climate-related risks and opportunities (8).

The North Dakota Department of Health, Environmental Health Section receives reports of environmental incidents, and has an up-to-date database includes incidents at oil production sites and general incidents that occurred during the transportation of oil (9), which was covered in the NYTimes (10).

Energy IQ (13) is an action-oriented benchmarking tool for non-residential buildings. Energy managers, building owners, architects and engineers use it to identify energy efficiency opportunities, save money, and reduce carbon emissions. The app uses the Energy and Consumption Efficiency Data and Statistics dataset (14).
## Sources

5. http://coolclimate.berkeley.edu
15. Skype interview with Eric Reese (GovEx, Johns Hopkins University), 11 January 2018.

<table>
<thead>
<tr>
<th>Social Impact 2: Increased inclusion of marginalized group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score:3</td>
</tr>
<tr>
<td>Explanation by experts:</td>
</tr>
<tr>
<td>Open data is helping cancer patients (1). Representations from various agencies spoke on panels and break-out discussions throughout the day. One audience favorite was the Cancer Moonshot panel on how data can fight cancer. (The Cancer Moonshot is a $1 billion dollar project of Vice President Joe Biden to end cancer.) During the White House Open Data Innovation Summit Panel, Presidential Innovation Fellow Michael Balint shared an API prototype created with the National Cancer Institute to help cancer patients learn about clinical trials they could qualify for (2). According to Eric Reese (GovEx) (10), &quot;in some ways, at the national level, we are going backwards. However, local government efforts foster. For example, in Austin, Texas by Chief Equity Officer are working towards equity as their goals (11). Another example is in Minneapolis, which published all of their spending contracts that women and minority businesses can bid for, and as a result were awarded more contracts&quot; (12).</td>
</tr>
</tbody>
</table>
Girls Who Code (3) has gone from 20 girls in New York to 10,000 girls in 42 states. Women Who Code a not-for-profit organization has chapters in various US cities (4). Data.gov has an entire page devoted to accessibility, where it fosters/safeguards accessibility in a number of ways: routinely reviewed for alignment with the latest Web Accessibility Initiative Guidelines for W3C; routinely tested for compliance with Section 508 of the Rehabilitation Act using a technical standards checklist, in-depth testing with screen readers, policy experts, and people with disabilities (5). In addition to accessibility standards, different data sources are available and provide information on different marginalized groups in the US, such as US Disability statistics (6). The US Department of Veteran Affairs has its own open data website with Featured Data includes VA Facilities Locations, Homeless Resources and Family Caregiver Resources available in JSON (7). The NCAI Policy Research Center recently received a grant from the National Science Foundation to improve the quality of data for American Indian and Alaska Native populations by modeling exciting new ways in which tribal nations are collecting important citizen data (8). The Anti-Eviction Mapping Project is a data-visualization, data analysis, and storytelling collective documenting the dispossession of San Francisco Bay Area residents in the wake of the Tech Boom 2.0. Through digital maps, oral history work, film, murals, and community events, the project renders connections between the nodes and effects of new entanglements of global capital, real estate, high tech, and political economy. It studies the displacement of people but also of complex social worlds as certain spaces become desirable to such entanglements. The project uses open source and open data maps (9). However, there are no peer reviewed studies or compelling evidence that open data is widely cited to have made a significant contribution to the inclusion of marginalized groups.

## Sources
(2): https://www.cancer.gov/research/key-initiatives/moonshot-cancer-initiative
(4): https://www.womenwhocode.com/networks
(5): http://www.data.gov/accessibility
(6): https://www.disabilitystatistics.org/
(7): http://www.va.gov/data/
(9): http://www.antievictionmap.com/
(10):
| Economic Impact 1: Noticeable positive impact on the economy | Score: 4  
Explanation by experts:  
Open data for economic growth is a priority for the Trump administration. On 25 July 2017, the Office of Management and Budget and the Center for Open Data Enterprise co-hosted a Roundtable on Open Data for Economic Growth that included approximately 80 government and business leaders. The goal of the Roundtable was to connect government and private-sector leaders to use open data as a strategic resource for better government and growing the American economy, by developing a better enabling environment for the use of open data and identifying high-value data sets for business use (1).  
Data.gov states that open data impacts include "cost savings, efficiency, fuel for business, improved civic services, informed policy, performance planning", amongst others (2).  
Open Data 500 (3) lists 529 companies using open data to make business. Examples include Washington-based company called Porch.com (4), which uses city work permits, licenses, and other residential construction information to create a searchable database and iTriage (5), a health care technology company which offers a portfolio of location-aware mobile apps designed to give consumers instant access to information about healthcare providers and facilities right in their neighborhood.  
## Sources  
(2): http://www.data.gov/impact/  
(4): Porch.com  
| Economic Impact 2: Increasing new businesses | Score: 8  
Explanation by experts:  
To a great extent, entrepreneurs are successfully using open data to build new businesses in the US. New businesses are launched based on U.S. Government open data. From weather and satellite data that powers Global GPS and consumer mapping applications to information on housing and demographics by city, state, and zip code, entrepreneurs have developed products and services fueled by free and open government data. Data.gov showcases many of these |
achievements on (9).
Moreover, a range of commercial apps that have been produced due to
data.gov (1). In addition, according to Open Data 500, there are quite a few
successful businesses based around open data (2), although some of these
companies are actually research centres (6).
Open Knowledge US created a business listings page that showcases the
state of open data for Business Listings in all the places for which they
have information (3).
On 17 May 2016 International Data Corporation (IDC) has published its
2016 IDC Innovators report recognizing pioneering players in the smart
city open data platforms market. IDC Innovators are companies with under
$50 million in revenue that offer an inventive technology and/or
groundbreaking new business model. Junar, NuCivic (a GovDelivery
compartment), OpenDataSoft, and Socrata were all named as IDC Innovators
in the smart city open data platforms market for 2016. (7)
NYC Business Atlas: "Welcome to the NYC Business Atlas. Use this tool
to explore business conditions in your neighborhood. Maps layers display
data for the entire neighborhood tabulation area. Population and
Demographic data are based on the closest census tract. Business
conditions data include all businesses in the neighborhood tabulation area.
To get started, choose a map layer to see citywide economic trends. Enter
an address or pick a location to see more detail." (8)
A November 2014 Information Week listed innovative healthcare startups
that have been generated by the open data movement (4). A similar article
about promising healthcare startups was published on 'Your Story', 20
companies to watch in the new wave of data-driven healthcare (5).
## Sources
(1): http://www.data.gov/applications
(3): http://us-city.census.okfn.org/dataset/business-listings
(4): Joel Gurin, Information Week, 9 Healthcare Innovations Driven by
Open Data, November 18th, 2014,
http://www.informationweek.com/government/open-government/9-
healthcare-innovations-driven-by-open-data/d/d-id/1317530.
(5): Joel Gurin, Your Story, 20 companies to watch in the new wave of
driven-healthcare/
(6): http://www.opendata500.com/us/list/ - go to research and consulting
(7): https://www.idc.com/getdoc.jsp?containerId=prUS41288716
(8): https://maps.nyc.gov/businessatlas/
(9): https://www.data.gov/impact/#business
Appendix B. Assumptions Check

Assumption 1. The independent variables are quantitative or dichotomous, the dependent variables should be measured on a continuous scale.
In our models, all the independent variables are quantitative, and the dependent variables are measured on a continuous scale.

Assumption 2. There should be no significant outliers.
In our models, after checking the cook’s d values (see Table B1), no value is bigger than 1, so there is no sign of significant outliers in our data.

Table B1

<table>
<thead>
<tr>
<th>Model</th>
<th>Cook’s D</th>
</tr>
</thead>
<tbody>
<tr>
<td>OGD Political Impact</td>
<td>.100</td>
</tr>
<tr>
<td>OGD Economic Impact</td>
<td>.062</td>
</tr>
<tr>
<td>OGD Social Impact</td>
<td>.056</td>
</tr>
</tbody>
</table>

Assumption 3. The assumption of a linear relationship between each of the independent variables and dependent variables.
After checking the Pearson Correlation, the correlation between each independent variable and the three dependent variables are statistically significant (see Table B2). I conclude that there is a linear relationship between each independent variable and the dependent variables.

Table B2

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>OGD Political Impact</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OGD Economic Impact</td>
<td>.767***</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OGD Social Impact</td>
<td>.747***</td>
<td>.769***</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OGD Government Policies</td>
<td>.741***</td>
<td>.728***</td>
<td>.600***</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OGD Government Actions</td>
<td>.810***</td>
<td>.852***</td>
<td>.734***</td>
<td>.863***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Properties of OGD related to Accountability</td>
<td>.765***</td>
<td>.762***</td>
<td>.704***</td>
<td>.781***</td>
<td>.826***</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Properties of OGD related to Innovation</td>
<td>.762***</td>
<td>.773***</td>
<td>.673***</td>
<td>.826***</td>
<td>.856***</td>
<td>.831***</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Properties of OGD related to Social Policy</td>
<td>.630***</td>
<td>.655***</td>
<td>.524***</td>
<td>.748***</td>
<td>.746***</td>
<td>.712***</td>
<td>.792***</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>User Participation</td>
<td>.708***</td>
<td>.716***</td>
<td>.650***</td>
<td>.782***</td>
<td>.836***</td>
<td>.718***</td>
<td>.730***</td>
<td>.632***</td>
<td>1</td>
</tr>
<tr>
<td>OGD Training</td>
<td>.729***</td>
<td>.806***</td>
<td>.670***</td>
<td>.740***</td>
<td>.850***</td>
<td>.751***</td>
<td>.780***</td>
<td>.710***</td>
<td>.768***</td>
</tr>
</tbody>
</table>
Assumption 4. The data shouldn’t show multicollinearity.
I used Variance Inflation Factor (VIF) to check whether there is a multicollinearity issue. The VIF value of each independent variable in each model is less than 10 (see Table B3), which indicates that there is no multicollinearity issue in our model.

Table B3
**Variance Inflation Factor**

<table>
<thead>
<tr>
<th>Variables</th>
<th>Political Impact</th>
<th>Economic Impact</th>
<th>Social Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Government Policies</td>
<td>5.39</td>
<td>5.49</td>
<td>5.45</td>
</tr>
<tr>
<td>Government Actions</td>
<td>7.92</td>
<td>7.91</td>
<td>7.31</td>
</tr>
<tr>
<td>Properties of OGD related to Accountability</td>
<td>3.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Properties of OGD related to Innovation</td>
<td></td>
<td>5.18</td>
<td></td>
</tr>
<tr>
<td>Properties of OGD related to Social Policy</td>
<td></td>
<td></td>
<td>2.95</td>
</tr>
<tr>
<td>User Participation</td>
<td>3.72</td>
<td>3.72</td>
<td>3.75</td>
</tr>
<tr>
<td>OGD Training</td>
<td>4.35</td>
<td>4.31</td>
<td>4.35</td>
</tr>
<tr>
<td>WGI</td>
<td>5.27</td>
<td>5.26</td>
<td>5.26</td>
</tr>
<tr>
<td>GDP Per Capita ($)</td>
<td>3.85</td>
<td>3.94</td>
<td>3.84</td>
</tr>
<tr>
<td>SPI</td>
<td>5.37</td>
<td>5.48</td>
<td>5.51</td>
</tr>
</tbody>
</table>

Assumption 5. The assumption of independence of errors.
I used the Breusch-Godfrey test, all the p-value is bigger than .05 (see Table B4), I failed to reject the null hypothesis, which indicates the independence of the errors in our models.

Table B4
**Breusch-Godfrey Test**

<table>
<thead>
<tr>
<th>Prob &gt; chi2</th>
<th>Prob &gt; chi2</th>
<th>Prob &gt; chi2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lag (1)</td>
<td>.108</td>
<td>.135</td>
</tr>
<tr>
<td>Lag (3)</td>
<td>.335</td>
<td>.500</td>
</tr>
</tbody>
</table>

Assumption 6. The assumption of normal distributed errors.
This assumption means the residuals (errors) of the regression line are approximately normally distributed. The normal PP plots of residuals of each model (Figure B1) conclude that the residuals are approximately normally distributed.
Figure B1. Normal PP plots of model residuals

Assumption 7. The assumption of homoscedasticity of errors. According to the scatter plots of the predicted value and model residuals, we can see that there is a sign of heteroscedasticity issues in our models. I used the Stata command “robust” to get robust standard errors and p values to fix this issue.

Figure B2. Scatter plot of predicted value and model residuals
Appendix C1. Interview Consent

The Utilization of Open Government Data: NYS GIS Web Services Case Study

INTERVIEW CONSENT

Purpose: This interview is conducted for the interviewer’s dissertation that focuses on the utilization of Open Government Data. Specifically, as part of the case study of New York State GIS Web services, this interview aims to answer two research questions: 1) how stakeholders interact with each other in the process of OGD utilization? And 2) what are the facilitators and obstacles in the process of OGD utilization and impact generation in the points of view of different stakeholders?

Participation: Interviews are being conducted with stakeholders who have experience using OGD and are related to NYS GIS Web services. We request your permission to interview you about your experiences and perceptions in this regard, and if necessary to contact you by phone or email for clarification. The interview will last about an hour.

Rights: Your cooperation in this study is completely voluntary. If you prefer not to participate or if you choose to withdraw from the study at any time, you may do so without penalty. You may refuse to answer any question and may leave the interview at any time. This interview will be recorded using audiotape equipment and fieldnotes. The information you provide will be used for research purposes only.
A copy of this form will be given to you. For questions about the research, please feel free to contact the interviewer via email (xzhao6@albany.edu). For questions about your rights as a research participant, contact the Compliance Office, Office for Sponsored Programs, University at Albany, at (518) 437-4569.

Confidentiality: To protect your privacy, all transcribing of tapes will be done by the interviewer. Fieldnotes, audiotapes, transcripts, and analysis will be kept by the interviewer. Presentations, reports, and publications will focus attention on general findings. Interviewees’ contributions will be referred to by type of informant, and, to the extent possible, these contributions will be reported in ways that avoid identification of those individuals. When a point can be made only by identifying a person, or when concealing identity is not possible due to the small sample size of persons or organizations involved or the uniqueness of the subject under study, those individuals’ prior permission would be requested.

__________________________________________________________________________
Interviewer’s Signature  Interviewee’s Signature

__________________________________________________________________________
Date  Date
Appendix C2. Interview Protocol

The Utilization of Open Government Data: NYS GIS Web Services Case Study

INTERVIEW PROTOCOL

Definition of OGD: In this research, we define Open Government Data (OGD) as data released by governments in machine readable formats and with licensing that allows anyone to have free access and re-use for any purpose.

Part 1: Informant’s background information

1. Please briefly introduce yourself, your professional background and your current position and responsibility.

<table>
<thead>
<tr>
<th>Prompts</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>✓ What is your professional background?</td>
<td></td>
</tr>
<tr>
<td>✓ What is your current position?</td>
<td></td>
</tr>
<tr>
<td>✓ What are your major responsibilities?</td>
<td></td>
</tr>
</tbody>
</table>

Part 2: GIS Web Services

In this section, we would like to know your experience with the GIS Web Services hosted by NYS ITS GIS Program Office, and discuss the roles of the key stakeholders, the facilitators and obstacles, and benefits and limitations of using this OGD based service.

2. Could you please describe how are the NYS GIS Web Services relate to the vision and goals of your organization?

<table>
<thead>
<tr>
<th>Prompts</th>
<th>For producer:</th>
</tr>
</thead>
</table>
   | ✓ What motivated the department to provide this service? | • development of OGD initiatives
   |               | • perceived benefits of OGD
   |               | • innovative organizational culture
   |               | • increasing inquiries from the users
   |               | • …

<table>
<thead>
<tr>
<th></th>
<th>For user:</th>
</tr>
</thead>
</table>
   | ✓ What motivated you/your organization to use this service? | • perceived benefit of the service: good data, no cost, etc.
   |               | • perceived ease of use of the service
   |               | • business development, research need
   |               | • …

3. Could you please describe the NYS GIS Web Services, what are its goals and how is it produced?
### Prompts

#### For producer:
- ✓ What is the production process of this service?
- ✓ Is there any regulation or policy that guides the production process?
- ✓ How many staff are involved in the production process?
- ✓ What types of staff skills are required to produce this service?
- ✓ What types of technology are used in the production process?
- ✓ What types of OGD are used to produce this service?
- ✓ How does your organization collect and manage the data to ensure its quality?
- ✓ What is your organization trying to achieve through providing this service?

#### For user:
- ✓ What do you know about the GIS Web Services hosted by NYS ITS GIS program office?
- ✓ What are you/your organization trying to achieve through using this service?

### 4. In your opinion, who are the main actors collaborating in the production and utilization of this service?

#### Prompts

**For producer:**
- ✓ Who are the main collaborators in the development of GIS Web Services and what are their roles?
- ✓ What is the benefit of this collaboration?
- ✓ What challenges have you encountered during the collaboration?
- ✓ What methods have you taken to cope with it?

**For user:**
- ✓ Who are the main collaborators in the utilization of GIS Web Services and what are their roles?
- ✓ What is the benefit of this collaboration?
- ✓ What challenges have you encountered during the collaboration?
- ✓ What methods have you taken to cope with it?

### 5. In your opinion, who are the main users of this service and how has it been used?

#### Prompts

**For producer:**
- ✓ Who are the users?
- ✓ How can users access this service?
- ✓ What is the amount of usage?
- ✓ What do they use this service for?
- ✓ What are the main strategies that the department has used to promote this service and with what results?
- ✓ Is there any communication channel/method between the producers and users? And how is the communication?
For user:
✓ How did you hear about this service?
  - social media
  - newsletter
  - online GIS forum
  - GIS association
  - word of mouth
  - …
✓ What do you usually use this service for?
  - data analysis for research
  - data analysis for work report
  - creating new interface or product for individual business
  - …

6. In your opinion, what are the key factors that facilitate your work related to the GIS Web Services?

Prompts
- For producer:
  ✓ What are the key factors that facilitate your production of GIS Web Services?
    - government initiative
    - budget/funding support
    - leadership
    - standards and protocols of production
    - familiarity in GIS background knowledge
    - skill in utilizing GIS related software
    - collaboration with other stakeholders
    - …

- For user:
  ✓ What are the key factors that facilitate your utilization of GIS Web Services?
    - familiarity in GIS background knowledge
    - skill in utilizing GIS related software
    - knowledge in OGD utilization
    - …

7. In your opinion, what have been the main challenges in your work related to the GIS Web Services?

Prompts
- For producer:
  ✓ What have been the main challenges in the development of this service during different stages of their development (initial VS now)?
• limited budget
• lack of technology and infrastructure
• poor data management and quality control
• difficulties in designing and implementation standardized protocols or guidelines
• lack of staff (e.g., lack of staff time, lack of skills, etc.)
• issues during collaboration with other government agencies or external organizations
• ...

For user:
✓ What have been the main challenges in the utilization of this service (initial VS now)?
• limited types of services available
• low quality of the service
• lack of technological skill in using the service
• lack of communication channel to report issues to service providers
• ...

8. What are the main strategies that you or the organization has adopted to cope with these challenges?

<table>
<thead>
<tr>
<th>Prompts</th>
</tr>
</thead>
<tbody>
<tr>
<td>apply for funding</td>
</tr>
<tr>
<td>hiring skilled staff</td>
</tr>
<tr>
<td>staff training</td>
</tr>
<tr>
<td>consultation with experts</td>
</tr>
<tr>
<td>improve communication with others during collaboration</td>
</tr>
</tbody>
</table>
| ...

9. In your opinion, what are the benefits of this service?

<table>
<thead>
<tr>
<th>Prompts</th>
</tr>
</thead>
<tbody>
<tr>
<td>easy and free access to high quality data</td>
</tr>
<tr>
<td>save time and effort on data processing</td>
</tr>
<tr>
<td>improve citizen engagement</td>
</tr>
<tr>
<td>enable OGD reuse to improve current products and services and create new ones</td>
</tr>
</tbody>
</table>
| ...

10. In your opinion, what are the main limitations of this service?

<table>
<thead>
<tr>
<th>Prompts</th>
</tr>
</thead>
<tbody>
<tr>
<td>lack of channel for user feedback</td>
</tr>
<tr>
<td>limited types of services</td>
</tr>
<tr>
<td>technology requirement on users</td>
</tr>
</tbody>
</table>
11. What is the plan of your work related to the GIS Web Services in the future?

<table>
<thead>
<tr>
<th>For producer:</th>
<th>✓ What is your/the department’s plan for the development of this GIS Web Services in the future?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• better understand user and user needs</td>
</tr>
<tr>
<td></td>
<td>• track its utilization and help to improve the service</td>
</tr>
<tr>
<td></td>
<td>• workshop and training on how to use this service</td>
</tr>
<tr>
<td></td>
<td>• OGD contest to facilitate use and innovation using this service</td>
</tr>
<tr>
<td></td>
<td>• …</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>For user:</th>
<th>✓ What is your/the organization’s plan for the utilization of this GIS Web Services in the future?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• increase utilization in business development</td>
</tr>
<tr>
<td></td>
<td>• increase utilization in academic research</td>
</tr>
<tr>
<td></td>
<td>• …</td>
</tr>
</tbody>
</table>

**Part 3. Wrap up**

12. Besides what we have discussed, are there any additional comments you would like to add?

13. About the case of NYS GIS Web Services, are there any other interviewees you would recommend me to talk to?

**Thanks for your help!**