Adapting Open
Exploring the Effective Use of Open Data in the Region of Waterloo

by

Cody Skinner

A Thesis
presented to
The University of Guelph

In partial fulfilment of requirements
for the degree of
Master of Science
in
Capacity Development and Extension

Guelph, Ontario, Canada

© Cody Skinner, December, 2014

ABSTRACT

Adapting Open: Exploring the Effective Use of Open Data in the Region of Waterloo

Cody Skinner
University of Guelph, 2014

Advisor:
Professor Helen Hambly Odame

This thesis explored the use of open data in the Region of Waterloo, using assessments of adaptive capacity as a hypothetical use case. Key informant interviews and a systematic review of grey literature were used to survey current perceptions of challenges to the use of open data in small and rural communities. The challenges identified using these methods fell on both the user and supply sides of the open data community of interest, and were presented within three categories: know-how, infrastructure and resources. An inventory of datasets available from relevant levels of government was evaluated for characteristics described in the study's adapted framework for effective use of open data. The three open data portals examined in the thesis were all found to contain data relevant to the case study. However, each portal had strengths and weaknesses when evaluated using the study's framework for effective use of open data.

Table of Contents

1.0 Introduction.....1

1.1 Problem Statement.....2

1.2 Case Study Location.....3

1.3 Climate Change in South Western Ontario.....4

1.3 Open Data.....5

1.4 Research Goal.....7

1.5 Research Objectives.....7

1.6 Significance of the Research.....8

1.7 Overview of the thesis.....8

2.0 Review of Literature.....9

2.1 Adaptive Capacity.....10

2.2 Resources for adaptive capacity.....16

2.2 a) Social resources.....17

2.2 b) Human resources.....18

2.2 c) Institutional resources.....18

2.2 d) Economic resources.....20

2.2 e) Natural resources.....21

2.3 Indicators.....22

2.4 Open data.....24

2.5 Criteria for the effective use of open data.....30

2.6 Summary and Conceptual Framework.....38

3.0 Methodology.....42

3.1 Ontological and Epistemological Approach.....42

3.2 Methodology.....45

3.3 Evaluation of a Data Inventory....46
3.4 Key Informant Interviews....48
3.5 Systematic literature review....50
3.6 Limitations....52
4.0 Findings....55
4.1 Challenges to using open data faced by small and rural municipalities....57
4.2 Challenges: Infrastructure/Technology....61
4.3 Challenges: Know-how....63
4.4 Infomediaries....66
4.6 Know-how on the supply side of open data....70
4.7 Challenges: Resources....72
4.8 Resources at upper and lower tier municipalities in the Region of Waterloo....73
4.9 Savings....75
4.10 The usefulness of open data....76
4.11 General findings on the usefulness of open data....77
4.12 Relevance to the Region of Waterloo....78
4.13 Open data's usefulness for assessing community adaptive capacity....81
4.14 Data content and Format - Granularity....84
4.15 Data Format....86
4.16 Openness....89
4.17 Longitudinal data....92
4.18 Metadata....93
4.19 Metadata: Level of description....95
4.20 Metadata: Data Collection / Creation Methods....97
4.22 Metadata: Time Coverage / Date of collection....98

4.23 Metadata: Update Frequency.....99

5.0 Discussion.....104

5.1 Infrastructure.....104

5.2 Resources.....106

5.3 Know-how.....107

5.4 Evaluating available data.....109

5.5 Available data: Data relevant to the Region of Waterloo.....110

5.6 Available data: Datasets by category of resource determining adaptive capacity.....111

5.7 Format.....114

5.8 Metadata.....117

5.9 For further research.....119

6.0 Final Summary, Conclusion, Further Research and Recommendations.....122

6.1 Final Summary.....122

6.2 Conclusions.....123

6.3 Recommendations.....125

References.....127

List of Figures

1. Nested Hierarchy Model of Vulnerability (Smit & Wandel, 2006).....15

2. Adaptive Capacity according to Wall and Marzall (2006).....38

3. Effective use of open data, adapted from Gurstein (2011).....40

4. Conceptual Framework: Assessments of Adaptive Capacity Using Open Data.....41

List of Tables

1. Relevance of Data on Provincial and Federal Open Data Portals to the Region of Waterloo.....79

2. Resource Category Frequencies for Open Data Relevant to the Region of Waterloo.....82

3. Relevance of Data on Provincial and Federal Open Data Portals to the Region of Waterloo.....85

4. Data formats of available open data.....88

5. Five star open data ratings of available open data.....90

6. Availability of times series open data.....93

7. Level of description for available open data.....96

8. Description of collection / creation methods for available open data.....97

9. Supplied contacts for available open data ….....98

10. Inclusion of time coverage / date of collection for available open data ….....99

11. Update frequencies of available open data.....101

Appendices

Appendix 1 – Sampling.....135

Appendix 2 – Criteria Used in the Evaluation of Available Open Data.....137

Appendix 3 – Qualitative Coding Framework.....140

Appendix 4 – Results of Systematic Review of Literature Search.....141

Appendix 5 – Systematic Review of Literature References.....143

Chapter 1: Introduction

Open data, machine-readable and modifiable data that is free to all to access, re-use and redistribute, is currently a very fashionable topic within Canada. Governments at all levels have been moving toward making their data available to the public to increase transparency, foster innovation and engage their citizens. However, at the municipal level, the focus has been on the tech-savvy population of urban cities. Little thought given to how small and rural municipalities can use open data, and what challenges they may face in doing so.

One of the difficulties in studying open data is that it's potential uses are often hailed as being limitless. However, despite many scholars offering opinions on what is required for open data to be used effectively, very little has been published on actual successful applications of open data. This study attempts to take a specific, albeit hypothetical, use-case and examine its potential through a framework of 'effective use' of open data.

Information and communication technology (ICT) has received much attention for its applications in the field of development, and recently the specific issue of climate change adaptation (Pant & Heeks, 2012). However, it is very difficult to find studies of a similar application of ICT in the developed world, despite it being ubiquitous. Given the researcher's history in the environmental movement, this gap seemed an ideal use-case through which to study the effective use of open data in small and rural communities.

The title of this study, *Adapting Open: Exploring the Effective Use of Open Data in the Region of Waterloo*, in part reflects the challenges specific to small and rural communities that

need to be acknowledged and addressed if open data is truly meant to benefit the widest possible audience. It is also a matter of wordplay on the concept of 'working open', a transparent way of working that invites participation used in the software industry. Successful assessments of adaptive capacity, in the framework employed below, depends on the participation and insight of a community. The use of open data as source material for these assessments leaves open the possibility for anyone to offer their own interpretation and insight in such a process.

1.1 Problem Statement

Access to information is a necessity for climate change adaptation. However, while it has become increasingly common for governments at the federal, provincial and municipal level to make large amounts of data available for public use, the quality of this data and the needs of potential users must be properly considered for the information to be used effectively.

While most of the world has been slow to act on climate change mitigation, those who are most vulnerable are finally making serious moves on anticipatory strategies for adaptation (MacLean & St. Arnaud, 2008). To understand how a community can be affected by climate change, however, is no simple task. Our consideration must reach beyond natural and environmental factors to consider human, social, economic and institutional resources as well (Wall & Marzall, 2006).

Open data is a relatively new concept, but one that is being embraced with great enthusiasm by governments at all levels in Canada. Canada's federal open data pilot project was launched in 2011, followed by a larger open government action plan in 2012 (Government of

Canada, 2012). The federal open data portal now has almost 200,000 datasets from 25 organizations. The Province of Ontario also introduced its own open data portal in 2012, and many local governments, including the Region of Waterloo, are now also making data available for public use.

There are a number of issues that must be addressed for open data to meet its true potential (Gurstein, 2011). Enthusiasts, however, hail anticipated benefits for government accountability, democratic participation and economic, academic and social innovation (OKFN, 2012; Hujboom & Van den Broek, 2011). The research presented here will ask whether open data, used as indicators of human, social, economic and institutional resources, can also contribute to an understanding of a community's adaptive capacity.

1.2 Case Study Location

The Region of Waterloo is a two-tier municipality in South Western Ontario. The region's lower-tier municipalities consist of three cities and four predominantly rural townships. The region is home to two universities, a community college and a large technology industry. The upper-tier municipality has embraced the concept of open data, having opened its own open data portal in 2012.

Work on the Region of Waterloo's open data initiative began as early as 2010, with the portal officially launched at the end of 2011. It was the first upper-tier municipal open data portal in Canada, and at that time contained eight datasets. When examined by this study in June of 2014, the portal's catalogue had reached 32 datasets, each covering all seven lower-tier

municipalities in the region. A report to the upper-tier municipality's Administration and Finance Committee in the Spring of 2013 notes that while offerings on the region's portal are modest, they reflect the services the region provides and interests of the open data community.

1.3 Climate Change in South Western Ontario

While the necessity for adaptation to climate change is now widely recognized in Canada, Dickinson & Burton (2011) depict the overall response as an expanding mosaic made up of initiatives at the provincial, territorial, and municipal levels of government with little overall strategic design. According to the same authors, emphasis at the federal level has not been on developing a national adaptation strategy, but has focused on climate change model and scenario development and providing research to a growing community of adaptation scientists and networks (Dickinson & Burton, 2011).

The provincial government has introduced a number of adaptation initiatives including the Climate Change Action Plan released in 2007 (EPCCA, 2009). The Ministry of the Environment is the lead agency on climate change issues for the province, while the Ministry of Agriculture, Food and Rural Affairs has a role in developing agricultural adaptation strategies to climate change (C-CIARN, 2002). Adaptation plans in the Region of Waterloo have been largely focused on climate change mitigation, including an inventory of greenhouse gas emissions from regional operations. An action plan to reduce those emissions in collaboration with local partners in the community is set to roll out in the near-future (Region of Waterloo, 2013).

Cohen (2011) suggests projected impacts of climate change on ecosystems, natural resources and health may be felt more acutely in rural regions than in urban centres. Rural regions and small towns, already facing a wide range of social, economic, and environmental pressures, have lower capacities to adapt due to lower levels of investment in infrastructure, weaker information systems, and weaker institutions (Cohen, 2011). While Wolf (2011) argues research that focuses on the social process of adaptation in developed countries has been relatively neglected, recent literature suggests residents of rural communities do not necessarily separate climate-related risks from the broader environmental and socioeconomic challenges to which they currently have to adapt (McLeman et al., 2011; Wall & Marzall, 2006).

1.3 Open Data

Open data emerged from the open government movement which, according to Davies (2012) arose as a response to the gap between government practice and evolving articulations of the democratic rights of citizens. Yu and Robinson (2012) trace the beginnings of the movement in the United States, were the term was first used in 1957 by Wallace Parks in work that would influence the eventual passing of the Freedom in Information Act by the US Congress in 1966 (Yu & Robinson, 2012). Canada's own Access to Information legislation would not come for almost another two decades, passing in 1983 under the government of Pierre Trudeau (Hazell, 1989).

The recent global popularity of the movement is seen as having been ignited first in the United States with the election of Barack Obama. Mr. Obama began his presidency by issuing the Memorandum on Transparency and Open Government, setting three explicit goals for open

government: transparency, participation, and collaboration (Obama, 2009). By September 2011, Canada was among 56 other countries that would join eight founding nations in the multilateral Open Government Partnership. The Partnership brings together government and civil society to develop and implement open government reforms, setting a minimum set of standards related to government accountability as conditions for entry (Yu & Robinson, 2012; Open Government Partnership, 2014).

Although open data is but one component of open government strategies, it has become a point of focus for technologists, activists, the private sector, and civil society actors (Davies, 2012). The most popular definition of open data comes from the Open Knowledge Foundation's 'Open Definition' (OKF, 2012). To meet the criteria of this definition, data must be available as a whole and at no more than a reasonable reproduction cost; it must also be available in a convenient, machine-readable and modifiable format; its release must permit re-use and redistribution, including mash ups with other datasets; it must be available for use by all, commercial or non-commercial, and regardless of political or social interest (OKF, 2012). For Yu and Robinson (2012), the term open data combines both the technological and philosophical meaning of the 'open' label, "with a focus on raw, unprocessed information that allows individuals to reach their own conclusions" (pg.12).

The launch of official open data portals in the US and UK were major milestones for the open data movement, though they were actually predated by grassroots transparency websites like GovTrack.us and TheyWorkForYou.co.uk (Hogge, 2010; Janssen, 2012; Yu & Robinson, 2012). Canada's own open data portal was initially launched in March of 2011, and then revamped and relaunched after the federal government signed on to the G8 Open Data Charter

in June 2013. The federal government was well behind the progress made on open data initiatives by a number of Canadian cities and provinces. Ontario's open data portal also enjoyed a new breathe of life in 2013, following the announcement of an Open Government Initiative by Premier Kathleen Wynne. Consultations on open government, including open data, were launched across the province while work began on creating an exhaustive inventory of public data, now available for the public to vote on for release. As noted above, the Region of Waterloo was the first upper-tier municipality in Canada to create an open data initiative, launching its open data portal late in 2011.

1.4 Research Goal

The goal of this research is to explore the use of open data to assess adaptive capacity in rural Ontario communities.

1.5 Research Objectives

The objectives for this research are to:

1) To explore the potential of open data when assessing the adaptive capacity of rural communities.

2) To identify challenges to effective use of open data in rural communities .

1.6 Significance of the Research

This research provides insight into the use of open data at the local level in rural areas. It explores barriers to the effective use of open data for the consideration of initiatives at all three levels of government examined in this research, but is especially important for small municipalities considering implementing an open data initiative.

The study is unique in looking at assessments of adaptive capacity as a use-case for open data and, as such, addresses a gap in studies of climate change in the developed world. The research underscores the importance of a multifaceted view of adaptive capacity when faced with complex issues like climate change. It encourages systemic approaches to policy and programming, informed by data from across a spectrum of perspectives.

1.7 Overview of the thesis

Chapter 2 presents a review of literature in two major parts. First, the concept of adaptive capacity is reviewed with an emphasis of the concept's role in the vulnerability framework. Wall and Marzall's (2006) framework for adaptive capacity, which contributes directly to this study's own framework for adaptive capacity, is introduced. Wall and Marzall (2006) argue the determinants of adaptive capacity as being made up access to five categories of resource is introduced and literature related to each of these categories is explored in greater detail. The second half of the review of literature examines literature related to open data. The origins and perceived benefits of open data are examined prior to moving into a review of the conditions identified in the literature as necessary for open data to be used effectively.

Chapter 3 describes the mixed methods approach employed in this research. Key informant interviews and a systematic review of grey literature are used to identify challenges to the effective use of open data in rural communities, while an inventory of available open data relevant to the Region of Waterloo is evaluated using quantitative methods for its potential to assess local adaptive capacity and for the presence or absence of properties supporting effective use. Chapter 4 reports the findings of the research conducted using those methods. Results from key informant interviews and the systematic review of grey literature related to challenges to open data use in small and rural communities are presented within three themes: technology and infrastructure, resources, and know-how. Findings from the evaluation of available data are presented categorically, with tables reporting frequencies.

Chapter 5 identifies connections between the study's qualitative and quantitative findings and places these within in the context of the literature reviewed in Chapter 2. The discussion chapter closes by identifying areas for further research. Chapter 6 states the study's final conclusions in the face of its original objectives, and closes with a number of recommendations for policy makers, open data stewards and communities seeking to assess their adaptive capacity.

Chapter 2 – Review of Literature

This chapter presents a review of literature in two major parts. The first section examines the concept of adaptive capacity in existing literature, with an emphasis of the concept's role in the vulnerability framework. Five categories of resources contributing to Wall and Marzall's (2006) framework for adaptive capacity are then explored in greater detail. Section II examines the origins, perceived benefits, necessary conditions for the effective use of open data identified in existing literature. Finally, the two sections are brought together into a single a conceptual framework.

2.1 Adaptive Capacity

According to Smit and Wandel (2006) adaptation has grown as a consideration in the field of climate change concurrently with the consideration of climate change itself. These authors trace the term 'adaptation' as having originated in the natural sciences. The concept was first applied to human systems by the anthropologist and cultural ecologist Julian Steward to describe the adjustment of human cultures to the natural environment. The broader social sciences have, in turn, treated adaptation in human systems as being concerned with the success or survival of a culture (Smit & Wandel, 2006). Engle (2011) cites anthropology as providing the origins for understanding human adaptation to the environment, but sees the historical underpinnings of the concept of adaptive capacity as being found in sociology and organizational and business management, where it was viewed as a requisite property for leadership and organizational success. Referring directly to adaptation to climate change, McCarthy (2001) defines adaptive

capacity as “the ability of a system to adjust to climate change (including climate variability and extremes), to moderate potential damages, to take advantage of opportunities, or to cope with the consequences” (McCarthy, 2001, p. 21).

Engle (2011) differentiates between reactive adaptation, responding to a stress that has already occurred, and planned or anticipatory adaptation in which humans have a unique ability to take proactive measures to lessen the impacts of future events. The latter, according to Engle, “depend on one’s ability to understand what the future might resemble, but are also influenced by one’s ability to have learned from past experiences; particularly what worked (and did not) in similar circumstances" (2011, p 648). While there may be little historical precedent for climate change, Engle (2011) states that lessons from previous experience may still be applied to anticipatory adaptation, especially with regard to social mechanisms that contribute to the adaptation process.

Yohe and Tol (2002) note systems’ environments are inherently variable, but that most variability over short periods of time fall within a 'coping range' - “a range of circumstances within which, by virtue of the underlying resilience of the system, significant consequences are not observed” (p. 26). It is important to understand the boundaries of these coping ranges, beyond which the consequences become more significant. Smit and Wandel (2006) define a system's coping range as the intensity or duration of conditions that a system can accommodate, adapt to or recover from. They note that some authors apply coping ability to shorter term ability for a system to survive, and use the term adaptive capacity to describe long-term, sustainable adjustments. Brooks and Adger (2004) use adaptive capacity to refer to the property of a system to to expand its coping range under current variability, or in light of future

conditions. “In practical terms, adaptive capacity is the ability to design and implement effective adaptation strategies, or to react to evolving hazards and stresses so as to reduce the likelihood of the occurrence and/or the magnitude of harmful outcomes resulting from climate-related hazards" (Brooks & Adger, 2004, p. 173).

The vulnerability framework for climate change adaptation suggests the vulnerability of a system, at any scale, is a function of the exposure to variation in external factors influencing the system; the sensitivity of that system to be affected, either adversely or beneficially, by those stimuli; and, the capacity of the system to adapt to the effects of those conditions by moderating negative effects and taking advantage of new opportunities (Smit & Pilifosova, 2003; Yohe & Tol, 2002; Smit & Wandel, 2006; Wall & Marzall, 2006; Gupta et al., 2010; Pant & Heeks, 2012; Holsten & Kropp, 2012; Schneiderbauer et al., 2013). A system can be exposed to many different stresses simultaneously, with different sectors and people sensitive to the same exposure (Yohe & Tol, 2002). “[A] system (e.g. a community) that is more exposed and sensitive to a climate stimulus, condition or hazard will be more vulnerable, Ceteris paribus, and a system that has more adaptive capacity will tend to be less vulnerable, Ceteris paribus" (Smit & Wandel, 2006, p. 285).

Wall and Marzall (2006) identify three cornerstones of adaptation to climate change using the vulnerability framework: reducing a system's sensitivity; altering the system's exposure; and increasing the capacity of the system to adapt to the changes. However, a community's power over the first two of these options may be limited. Exposure and sensitivity are the result of interaction between the characteristics of the system and the attributes of the climate stimulus (Smit & Wandel, 2006). Adaptive capacity moderates exposure and sensitivity,

and represents a system's ability to prepare for and adjust to stress, reducing negative impacts and taking advantage of any new opportunities that arise (Yohe & Tol, 2002, Engle, 2011, Holsten & Kropp, 2012).

While the vulnerability framework is employed in this study, the concept of adaptive capacity is also found in another widely used approach. The resilience framework focuses on the ability of a dynamic system to fluctuate within the boundaries of a stable state (or multiple stable states) without being pushed beyond those limits to a point where the system must transform itself (Engle, 2011). While resilience theory originates in the natural sciences, it increasingly includes human contributions to system dynamics, as evidenced by the growing study of social-ecological systems (Engle, 2011). Engle (2011), citing Walker et al., (2006), characterizes the concept of adaptive capacity in resilience literature as the capacity for actors to manage and influence resilience through interactions between the human and environmental components of a system. Ahmed describes the concept as "an aspect of resilience that reflects learning, flexibility to experiment and adopt novel solutions, and development of generalized responses to broad classes of challenges" (2006, p.21).

For Engle (2011) the concept of adaptive capacity in resilience literature is limited in that it is difficult to translate into practice, and thus is difficult to measure or characterize. Meanwhile the concept of resilience is often interpreted differently by practitioners and scholars, with the former seeking an engineered return to a stable equilibrium while the latter are interested in reorganization, renewal and learning (Engle, 2011). But while Engle (2011) sees the vulnerability approach as more practical, the author warns it can exclude key variables present in the resilience approach that more accurately capture the dynamics of a system, and is

susceptible to focusing on a limited time frame and spatial scale. The author advocates for assessments of adaptive capacity to use an approach that takes advantage of the strengths of the two frameworks.

Smit and Wandel (2006) list four categories of analysis focused on adaptation to climate change: those assessing the degree to which adaptation can moderate or reduce negative impacts of climate change, or realize positive effects; those with a focus on specific adaptation options for a particular system; those focused on the relative adaptive capacity (or vulnerability) of specific geographic locations, with comparative evaluations using a select set of indices and variables; and, finally, those focusing on the implementation of adaptation processes. Smit and Wandel (2006, p. 285) suggest this last category is not often explicitly called adaptation research, but can include studies "in the fields of resource management, community development, risk management, planning, food security, livelihood security, and sustainable development".

The motivation for such an approach is to identify what measures can be taken, in a practical way, to increase adaptive capacity. The findings are specific to the community under investigation, and are not meant to be scalable. This approach avoids presumptions of the specific variables that represent exposures, sensitivities, or aspects of adaptive capacity. Instead it seeks to identify these empirically from a community. "It employs the experience and knowledge of community members to characterize pertinent conditions, community sensitivities, adaptive strategies, and decision-making process related to adaptive capacity or resilience. It identifies and documents the decision-making processes into which adaptations to climate change can be integrated" (Smit & Wandel, 2006, p. 285).

Holsten and Kropp (2012) note that while sectoral vulnerability assessments have become common, transferable approaches needed for wider adaptation strategies are rare. Yohe and Tol's (2002) widely-cited work attempted to find methods of assessing adaptive capacity that are “sufficiently flexible to accommodate diverse applications whose contexts are location specific and path dependent without imposing the straightjacket constraints of a 'one size fits all' cookbook approach” (2002, p1). However, Brooks and Adger (2004) argue the factors constituting adaptive capacity are determined by the internal characteristics of a system as well as the nature of external influences affect it. Such a complex and context-dependent mix of factors make it impossible to develop a list of standard indicators that characterize a universal set of determinants for adaptive capacity. The authors instead suggest appropriate indicators must be selected on a case by case basis (Brooks & Adger, 2004).

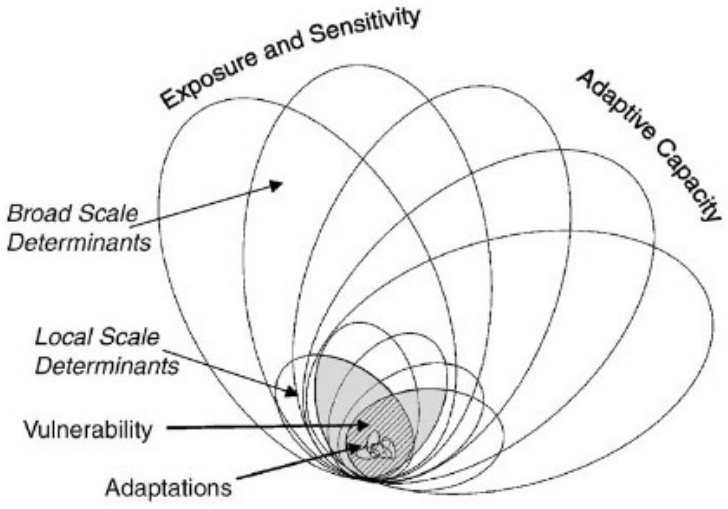

Figure 1. Nested Hierarchy Model of Vulnerability (Smit & Wandel, 2006, p. 286)

For Smit and Wandel (2006), the determinants of adaptive capacity are interdependent and can occur at a variety of scales. Some determinants are mainly local while others reflect larger scale

socio-economic and political systems. "Individual determinants, thus, cannot be isolated: adaptive capacity is generated by the interaction of determinants which vary in space and time. The determinants of adaptive capacity exist and function differently in different contexts" (Smit & Wandel, 2006, p. 288). As visualized in Figure 1, the different sizes of loop represent exposure, sensitivity and adaptive capacity at various scales, with the smaller loops representing the local or community level. As such, local initiatives to enhance adaptive capacity can be distorted or nullified by broader forces with the power to shape local vulnerabilities (Smit & Wandel, 2006). Their overlap shows the interdependence and interaction of the various forces that determine adaptive capacity (Smit & Wandel, 2006).

2.2 Resources for adaptive capacity

Schneiderbauer et al. (2013) distinguish three conceptual levels of adaptive capacity to guide assessments, each with a decreasing degree of specificity: impact-specific adaptive capacity, sector-specific adaptive capacity and regional-generic adaptive capacity. For each level, the authors state various dimensions (political, legal, institutional, social and economic) aspects must be identified and measurable indicators selected to describe them.

Brooks and Adger identify resources required for adaptation, including “financial capital, social capital (e.g., strong institutions, transparent decision-making systems, formal and informal networks that promote collective action), human resources (e.g., labour, skills, knowledge and expertise) and natural resources (e.g., land, water, raw materials, biodiversity)” (2004, p. 173). Wall and Marzall (2006) formalize this resource-based approach to adaptive capacity into a framework founded on a community's access to resources within five general

categories: social, human, institutional, natural and economic resources. Access, as they define it, includes both the availability of resources and the ability to mobilize them. A general discussion of the importance of each category to adaptive capacity follows.

2.2 a) Social resources

Climate change adaptation is a social process, shaped by a community's historical trajectory and influenced by the social and cultural context in which it occurs (Wolf, 2011; McLeman et al., 2011). Local adaptive capacity is strongly influenced by social networks and the social capital embedded within them which in turn influences a community's ability to act collectively (Engle, 2011; Wolf, 2011; McLeman et al., 2011). Wall & Marzall (2006) argue that communities with strong social resources have more adaptive capacity. Folke et al. (2005) suggest informal social networks can promote innovation and enhance flexibility. Simpson et al. (2005) describe social capital as an outcome of interaction and participation in networks and the formation of new social ties. These can vary from intimate relationships with like-minded people to more diverse and outward-looking ties across groups with different characteristics. Strong ties are important for trust and reciprocity, while diverse, albeit weaker, ties are important for the exposure to new ideas, skills and resources (Simpson et al., 2005). Social capital can be polarized by virtuous and vicious cycles. Communities with high social capital grow even stronger from self-reinforcing virtuous cycles, while vicious cycles trap dysfunctional communities into diminishing social capital (Putnam in Simpson et al., 2005).

Social capital is sensitive to changes in the demographics of a community, especially in terms of its age structure and cultural makeup. High levels of social capital come from

diversified leadership representative of the age, gender and cultural composition of the community (Simpson et al., 2005). Over-reliance on the same individuals or groups to lead community initiatives can lead to problems (McLeman et al., 2011). In Wolf's (2011) view, examining the psychological, social, and cultural aspects of adaptation can expose dimensions of power and their effect on those strategies. Cohen (2011) suggests it may be challenging to initiate dialogue on adaptation within an environment of transparency and trust when different interests produce different visions of adaptation.

2.2 b) Human resources

Wall and Marzall's (2006) definition of human resources includes the community's collective physical capabilities, skills, knowledge and experience (Wall & Marzall, 2006). Davidson et al. (2003) suggest the limited human capital and highly specialized skill sets characteristic of employment in resource sectors has increased the vulnerability of rural communities to climate change. The success of future capacity building and planning will also be influenced by the ability to adjust to and take advantage of rapidly changing demographic patterns in rural communities (McLeman et al., 2011).

2.2 c) Institutional resources

The capacity of social groups to act depends on the quality of the formal institutions within which they must operate (Wall & Marzall, 2006). Brooks and Adger (2004) state adaptation strategies will fail without a community's willingness to adapt and a degree of consensus on the action to be taken. Both of these requirements depend on the ability to act collectively, which is

itself heavily influenced by institutions (Brooks & Adger, 2004, 173). Meanwhile, the socio-cultural context in which institutional structures operate also influence decision-making (Wolf, 2011). Folke et al. (2005) point out that processes of change can also challenge governance systems themselves.

Gupta et al. (2010) argue the challenge of climate change calls for institutions that both promote the adaptive capacity of society and allow society to modify its institutions to face those changes, striking a proper balance between its own rigidity and flexibility (Gupta et al., 2010). The authors argue that while institutions are traditionally conservative, creating stability and predictability in society, successful contributions to adaptive capacity will be made by institutions that support social actors to respond to system stresses both through planned processes and autonomous, spontaneous change (Gupta et al., 2010). For Gupta et al., institutions that promote adaptive capacity are those that

> 1) encourage the involvement of a variety of perspectives, actors and solutions; 2) enable social actors to continuously learn and improve their institutions; 3) allow and motivate social actors to adjust their behaviour; 4) can mobilize leadership qualities, 5) can mobilize resources for implementing adaptation measures, and 6) enhance principles of fair governance. (2010, p.464)

McLeman et al. (2011) found rural communities had inadequate resources for planning and coordination with higher levels of government, while generalized, ill-fitting information and regulation were challenges to increasing local adaptive capacity. Brooks and Adger (2004) agree that regulations or policies from higher levels of government can make certain adaptation strategies unviable or hinder the ability for a community to take action locally. Some authors have suggested ensuring new policies do not undermine others, both within and across sectors, is key to effective adaptation strategies (Urwin & Jordan, 2008; Adger et al., 2005). However,

Folke et al. (2005) argue the redundancy of institutions and their overlapping functions across organizational levels can spread risks and increase the diversity of response options.

2.2 d) Economic resources

Wall and Marzall (2006) describe economic resources as "financial resources to draw upon should a climate or weather event disturb income generation" (pp. 385). This can be in the form of savings or investments, or diversity of employment options should any there be an interruption to the operations of a specific industry (Wall & Marzall, 2006). Brooks and Adger (2004) point out economic factors and vested interests can undermine adaptive capacity if key actors refuse to accept the existence of the risks associated with climate change, or to accept responsibility for adaptation.

Changing commodity prices, trade agreements, resource use rights, government subsidies and support programs and globalizing flows of capital and labour all affect the context within which adaptation occurs (Adger et al, 2005; Smit & Skinner, 2002). In the view of Adger et al. (2005), "any assessment of the economic efficiency of adaptation actions requires consideration of, first, the distribution of the costs and benefits of the actions, second, of the costs and benefits of changes in those goods that cannot be expressed in market values, and, third, the timing on adaptation actions" (pp. 82).

Smit and Wandel (2006) suggest adaptation can be considered as adjustments to deal with changing conditions within the constraints of broader economic or social–political arrangements, but where those constraints are particularly binding, it may be a more effective

strategy to change those broad structures themselves. McLeman et al. (2011) point out maintaining community identity when undergoing economic diversification can be a challenge.

2.2 e) Natural resources

Natural resources, for Wall and Marzall (2006) are not only the resources needed to sustain life, but also economic activity in a community. Kulig, Edge and Joyce (2008) point out that rural communities are usually resource-reliant or single-industry towns, and are thus already particularly sensitive to negative external forces such as downturns in global markets. Smit and Skinner (2002) suggest agriculture in Canada could benefit from climate change if the correct steps for adaptation are taken, but will suffer if no measures are taken at all. While most impact studies consider changes in average climate conditions, analyses of agricultural vulnerability indicate that the key attributes of climate change for farmers are those related to climatic variability and the frequency of extreme conditions (Smit & Skinner, 2002; Bryant et al., 2000). Ahmed (2006) describes natural resource management approaches as often being based on a presumed ability to predict external changes as well as the responses to management measures. However, this becomes especially difficult when, as with climate change, the uncertainties are many. Furthermore, Ahmed (2006) argues sustainability is a matter of maintaining the functionality of a system when it is under pressure, or of maintaining the elements needed to reorganize should radical alterations to the structure and function of the system be necessary.

2.3 Indicators

Wall and Marzall's (2006) framework for adaptive capacity is an attempt to balance contributions from scientists and other specialists with information from the stakeholders whose livelihoods are in the balance by providing an assessment of adaptive capacity using indicators specific to a given community. Cohen (2011) argues planning for adaptation requires cooperation between researchers, practitioners and stakeholders, facilitated by quantitative analysis. While Smit and Wandel (2006) state the determinants of adaptive capacity should be identified by the community in question through the active engagement of stakeholders and decision-makers, they agree that an assessment should also include information collection on community relevant phenomena and processes, and the integration of information from multiple sources (p. 285). “Once relevant conditions have been identified, and future livelihoods considered, information from other scientists, policy analysts, and decision-makers, are integrated into the analysis to identify potential future exposures and sensitivities ... and future adaptive capacity ... to determine future vulnerability” (Smit & Wandel, 2006, p. 289).

Adaptive capacity is regarded as the most challenging component to determine in the vulnerability formula, no agreed-upon applied concept of what constitutes adaptive capacity in practice (Yohe & Tol 2002; Schneiderbauer et al., 2013). Brooks and Adger (2004) suggest adaptive capacity cannot be directly measured, and must be characterized by examining potential changes in the sensitivity of systems. As such, the authors argue against falling into “the potential pitfall of trying to identify a comprehensive list of quantitative capacity indicators”, stating “it is more important to understand and to characterize the adaptation process in a pragmatic manner” (Brooks & Adger, 2004, 176). For the authors, this includes

assessing the willingness and resources necessary to adapt.

While it may be difficult to accurately characterize the above resources, Wall and Marzall (2006) argue that the assessment of local capacity to handle stress from a host of external and internal factors is necessary in order to promote long-term sustainability in rural areas. Their framework is tailored to local conditions through input by specialists balanced by engagement with stakeholders in the community in order to define the composition and importance of each category of resource and to select appropriate indicators for each (Wall & Marzall, 2006, p. 379). The specifics and relative importance of each category of resource will depend on the context of the community, on the nature of the hazards faced, and on the adaptation strategy pursued (Brooks & Adger, 2004; Wall & Marzall, 2006).

Frazier, Thompson and Dezzani (2014) criticize many assessments of adaptive capacity for failing to weight indicators based on their influence on vulnerability or for the use of unquantifiable determinants. Brooks and Adger (2004) suggest indicators may be quantitative "such as population density or average income, or qualitative, representing factors such as the principal type of economic activity in a region, or people's perceptions of risk" (Brooks & Adger, 2004, p. 176). Yohe and Tol (2002) argue qualitative analysis can help to uncover practical insights related to adaptive capacity, such as the scale at which various elements of adaptive capacity are needed but ultimately suggest all determinants of adaptive capacity can, in principle, be depicted in multiple ways.

The intended use and audience of an assessment will also likely influence the selection of indicators. Juhola and Kruse (2013) focus on assessments used to make recommendations for policy makers on how to enhance adaptive capacity. The authors state that in order for results to

be relevant and useful policy makers, adaptive capacity assessments need to consider the target audience for the results, how results will be presented to that audience and how they can be put into action. The authors suggest the translation of results to stakeholders requires careful consideration of what information is necessary and how it is communicated and can benefit from the use of knowledge brokers between researchers and policy makers. Results should be presented in a format specific to the target audience, who will in turn be an important influence on how results are used.

2.4 Open data

Open data is one of many 'open' movements, along with open source, open access, open innovation, open education and open knowledge. These movements, Davies (2012) suggests, are meant to challenge a set of previously existing 'closed' arrangements. Beyond its use in the world of software development, the term 'open source' is now taken to denote a collaborative innovation strategy, using the power of many to collaborate on the development and testing of projects aimed at creating freely available end products (Yu & Robinson, 2012).

Yu and Robinson (2012) state the common thread across 'open' labels lies in the ability for users to access a certain class of information and interact with it on their own terms, usually through the use of technology. "The label 'open', as applied to various kinds of information, thus inherits both a technological and a philosophical meaning" (Yu & Robinson, 2012, pg.11). At a technological level, the term suggests using technology to handle information efficiently, thereby extending the range of logistically feasible ways in which that information can be used.

Philosophically, the term suggests participation and engagement by enabling sharing and reuse via an absence of legal barriers and an enthusiasm for innovation (Yu & Robinson, 2012)

The combination of government reform efforts and the emergence of advanced technology tools for information access has renewed the idea that democratic governments should be open, accessible, and transparent (Dawes, 2010). When it comes to open data, Worthy (2013) distinguishes between three kinds of transparency. Accountability transparency refers to the use of open data to hold authorities to account for both their political and financial performance, while participative transparency is information that serves to stimulate increased participation. Lastly, informing transparency is information on services and operations needed to judge performance.

Peled (2013) cautions open data enthusiasts to temper their expectations for transparency, warning that expectations coming up short could have negative impacts on the willingness to invest in future transparency programming. Instead, Peled (2013) argues, open data should be one component in a wider transparency program focused on improving services and general information. Meanwhile, issues of data quality must be addressed, with citizens and government collaborating to identify errors and add context.

Oliver (in Meijer, 2009) argues that the use of information and communication technology fuels an 'Information-Transparency Cycle', in which technology facilitates the collection and analysis of a wide range of information. This information can be directly and cheaply redistributed, allowing interested parties not only to react in a timely manner, but indicate what information is missing. The demand for new information re-initiates the cycle.

Meijer (2009) argues that scholars who identify as premodernists, modernists and postmodernists attach different meanings to computer-mediated transparency. For Meijer (2009), modern transparency is computer-mediated transparency. It is unidirectional, with information structured and decontextualized for use by computers. Janssen, Charalabidis and Zuiderwijk (2012) warn that more information can result in less understanding, more confusion, and less trust. "The reality is that finding the right data might be difficult, there might be a huge information overload, and large differences in the way open-data analysts and how an individual might analyze the same data may draw different conclusions" (Janssen et al., 2012, p. 266).

Meijer suggests a premodern perspective on transparency gives preference to traditional ways of producing trust, such as through direct contact and contextualized interactions. The opportunity for direct transparency, however, declines when societies move to a larger scale necessitating the need for computer-mediated alternatives. The post-modern view, on the other hand, frames computer-mediated transparency as a technology of representation, used to generate powerful images to dominate assessments of performance. Postmodernists argue that the criterion for evaluating transparency thus moves from an emphasis on effectiveness to esthetics. Meijer (2009) argues that we live in all three of these realities at the same time and recommends an approach of diversity to create a rich array of mechanisms of trust.

Open data is also hailed for its potential to enable greater civic participation. With government transparency comes heightened awareness levels bringing increased support for the implementation of projects and access to community energy and knowledge for the design of policy (Eversole, 2011; Evans-Cowley & Hollander, 2010). However, Evans-Cowley and Hollander (2010) point out that the extent to which participation is possible will be, in part, a

result of how well information has been presented using these tools. Technically savvy citizens may appreciate open data, but others will find it to be an inaccessible option.

Eversole (2012) suggests expert knowledge has often been the justification for limiting community participation in development processes and decisions. But, the author asserts, tacit knowledge must also be acknowledged as important. Unlike expert knowledge, it is familiar with the particular constraints and possibilities of a given physical ecosystem or cultural value system. The literature on participation and open data has yet to address how the local or tacit knowledge can be integrated with the structured, decontextualized information provided by government.

Scholars argue the success of any technology used for participatory ends will ultimately depend on shifts in cultural practices and policy (Maier-Rabler and Huber, 2011; Janssen et al 2012; Dawes, 2010). "In the open approach", state Maier-Rabler and Huber (2011), "socio-technical aspects meet political demands for co-creation by citizens and authorities in the online and offline world that set the frame for the discussion of open government" (p. 185). Addressing open data specifically, the authors contest the idea that data flows unidirectionally from government to citizens, while engagement goes vice versa from citizens to government. Noveck (2010) distinguishes between deliberation and collaboration as distinct form of democratic participation, and argues a collaborative approach to open government would advance participation beyond public debate. Noveck believes advances in technology have made large scale collaboration possible, and that the future of public institutions will depend on a collaborative ecosystem.

Open data is also said to have the potential to promote economic growth by contributing to the development of new products and services, improving existing ones, and supporting competitive markets by informing investments and business strategies (Janssen et al., 2012; Davies, 2012). Government data is valued for reliability, standardization, consistency and comprehensiveness (Lakomaa & Kallberg, 2013; Davies, 2013). Innovators are much more likely to invest time and resources in projects using data when they are assured about its continued availability, while social applications of data are often reliant on frequent updates and timely access (Davies, 2013).

Prior to open data, private actors had already demonstrated their ability to create valuable services using government information, even if it meant having to do extra work to collect and clean that data (Yu & Robinson, 2010). Ideally, open data would be released in a way that lowers the practical costs of reuse and lead to a boom in third party services built to meet the needs of specific communities. The outstanding issue, state Janssen et al. (2012), is for governments to know which data holds the most value and merits release.

However, Huijboom and Van den Broek (2011) argue that the precise economic impact of open data remains largely unclear. There is even less evidence of the effects of open data on participation, accountability and transparency, as social or political effects would require impact analysis using non-economic or non-quantitative means (Halonen, 2013). While it is clear that more data is available to citizens, whether or not that data is relevant to the issues the public is interested in is another matter (Halonen, 2013). While the costs of producing and publishing data can be measured, there is as of yet no widely accepted model for evaluating or forecasting the accrued benefits of open data (Shakespeare, 2013; Hujboom & Van den Broek, 2011).

Although various methods have been suggested to generate indicators of the impacts of open data, “none of the proposed methods are comprehensive or rigorous enough to encompass all the key aspects, and potential impacts, of assessing an open data initiative—nor to provide effective coverage of developed and developing nations” (Davies, Perini and Alonso, 2013, p. 27). This is problematic, as measures of the impact of open data should be central to informing decisions around the design and implementation of open data initiatives (Ubaldi, 2013; Jetzek, Avital, & Bjørn-Andersen, 2012; Helbig et al., 2012).

Halonen (2013) points out that the overall number of open data users may even be irrelevant, with the more important indicators found among those actually making use of open data with relatively high impact (Halonen, 2013). In the author's view, measuring transparency is not simply a question of the actual amount of information available, but also a question of equality in the accessibility and usability of that information. An assessment of open data's effects on participation must assess whether improved access to an increased amount of data will change citizens’ behavioral patterns.

Jetzek et al. (2012) suggest open data initiatives should assess the potential value open data may generate. The authors contend that open data use can be conceptualized as a value network, in which value is co-produced through a process that offers due consideration for the different and divergent interests of all collaborating partners (Jetzek et al., 2012). From the perspective of the private sector, this value can be accounted for in cost-savings, both direct and indirect, or opportunities to generate new revenue. Ubaldi (2013) suggests impact measurements are key to providing practitioners the incentive needed to move from proof-of-concept to production applications of open data to begin with.

On the 'social' side of open data, Jetzek et al. (2012) suggest looking for effects on existing indicators of equality or life expectancy. It would, however, be a long time before significant effects are seen using these measures. Nor do they account for new or unexpected benefits from open data use, with some researchers arguing that there "is no linear progression from data, to decision-making, to impact – but rather that these are in on-going interaction" (Davies, Perini and Alonso, 2013, p. 13).

Davies, Perini and Alonso (2013) argue a socio-technical approach to assessing open data impacts would provide an understanding of the flow of data from open data initiatives to potential users through a range of technical and social intermediaries. It would reveal how global standards, platforms, infrastructure and 'eco-systems' of open data affect local contexts and how the benefits from open data initiatives are distributed. The authors recommend researchers make use of case studies towards these ends but contend that it is also important to move beyond local findings to assess open data initiatives at the macro level (Davies, Perini and Alonso, 2013). By doing so, it will become possible to "understand commonalities between cases of open data publication and use, and to uncover common mechanisms through which open data may be involved in bringing about impacts" (Davies, Perini and Alonso, 2013, p. 12).

2.5 Criteria for the effective use of open data

Open data researchers and practitioners seem much more certain about the necessary preconditions for the success of open data than they are of how to measure it. Gurstein (2011) lays out a number of requirements for what he calls the 'effective use' of open data. These include an accessible and sufficient internet connection, hardware and software; the skills to use

those tools; the data in a useful format; sufficient knowledge to make sense of the data; resources sufficient for translating data into activities for local benefit; and the legal, regulatory or policy regime needed to use the data (Gurstein, 2011). Meeting the criteria of this framework, in the author's view, would ensure opportunities and resources for translating open data into useful outcomes are made available for the widest possible range of users (Gurstein, 2011).

Many scholars now contend that open data presents a potential risk of widening the gap in access it is, in part, meant to close. Despite claims that open data will provide new actors with the means to participate in public debate, it can be argued those most likely to take advantage of open data are already empowered and engaged individuals with a particular technical skill set and sufficient resources (Davies, 2010; Halonen, 2013; Peled, 2013; Janssen, 2012; Gurstein 2011). With a strong focus on making data available to the public, the open data movement has not done well to address the matter of data's intellectual accessibility (Janssen, 2012). While "technology platforms, common standards and open licences all play a key role in making data easier to access and simpler to process, they operate against a wider backdrop of organization and social arrangements, power dynamics, and market conditions which may or may not be favorable to the use of data by different agents" (Davies, 2012, p. 1). Some go as far as to compare open data to a subsidy for private interests that no longer have to pay for access to government data (Bates, 2012; Davies, 2011, 2013; Peled, 2013). For others, the limited time and resources of government open data programs may become completely focused on meeting the needs of developers and activists "because their cry for data is louder and the immediate rewards for government in terms of reputation and goodwill from the public will be better" (Janssen, 2012, p.3).

Furthermore, a public that is arguably already apathetic to the information currently available to them (Halonen, 2013), may be even more put off by the challenge of accessing and analyzing raw data (Halonen, 2013; Peled, 2013). Even before the advent of open data, Sawicki and Craig (1996) had begun to recognize major challenges in the use of data for social innovation and change. They identify a lack of technical sophistication among community groups, and the eventual transformation of data into analysis that can affect policy. "[E]ncyclopedic data dumps do not have much impact. They are not issue- or policy-oriented, and often do not serve as even basic references" (Sawicki & Craig, 1996, p. 518). The authors view the adoption of data-enabled strategies within community organizations as a further challenge: "The degree to which the products of information technology are produced for community organizations as opposed to being produced by community organizations themselves is an empowerment issue" (Sawicki & Craig, 1996, p. 518).

Many advocates are calling on open data curators to make open data inclusive beyond app developers and data scientists (Meijer et al., 2012; Davies, 2010, 2012b; Desouza & Bhagwatar, 2012; Janssen 2012). Many have called for the release of government open data to be accompanied by complementary programming to increase data literacy (Davies, 2012; Shakespeare, 2013; Janssen, 2012; Peled 2013; Maier-Rabler & Huber, 2011).

Writing from a community informatics approach to developing individual capacity for new technologies, Simpson (2005) recommends formal programs and informal activities that increase awareness of the potential and benefits of ICTs; formal and informal education and training programs that equip community members with the skills and knowledge required to become competent users of ICTs; building leadership across community groups and

organizations, with ‘local champions’ encouraging others to engage; building technical expertise for supporting and maintaining information and communication technologies. The author highlights the importance of access to the new technology and the ability to experiment with it; exposure to the new technology in a social setting that caters to different interest groups; recognition of the role of local opinion leaders and existing social networks; and, an understanding that time is needed for a new technology to take hold.

For others, the goal of inclusiveness can be met through the use of accessible tools to easily manipulate data into summaries and visualizations in a customized way that meet a user's needs (Halonen, 2013). While intermediaries are able to process data into platforms and products with social and economic value for others (Davies, 2013). Janssen (2012) warns against an over-reliance on such services as any representation of data will contain value judgments. This puts government, or the party contracted to provide intermediary services, in a position of power that risks being misused to advance their own interpretations of the data (Janssen, 2012). There also exists a risk that commercial companies may profit disproportionately for services interpreting data that is meant to be freely accessible (Janssen, 2012).

While Gurstein's (2011) own work refers mostly to what he calls the 'demand' side, he admits that aspects of effective use are dependent on the supply side of open data. The elements Gurstein refers to are often matters of how the data is preprocessed and presented for discovery and reuse. This includes formatting, tagging and metadata, but is also a matter of how presentation can affect its interpretation, for example by choice of interval in time series data, or the choice to include certain measurements over others. The impacts of these characteristics will

vary between users and how they intend to use the data, with different requirements needed of the data depending on the desired outcome (Davies, 2013; Gurstein, 2011).

Dawes (2010) also warns against assumptions that government information is neutral, and must be evaluated for comprehensiveness, authoritativeness, and trustworthiness. Helbig et al. (2012) remind readers that "data does not exist in the wild; it is deliberately created by socio-technical processes" (Helbig et al., 2012, p. 13). While data quality is often considered solely in terms of accuracy, Dawes (2010) urges the need for data to be accessible to users, clearly represented, and contextually appropriate for a task in terms of temporal, security, granularity and other requirements.

Context, as described by Helbig et al. (2012), is information related to the environment from which data is acquired or extracted, encoded, and typically used to impact government and public life. Failure to include context, the authors argue, can lead to conflicts in overall meaning, misunderstanding of data elements and a lack of use due to uncertainty over the value of the information a dataset holds (Helbig et al., 2012). Different users will require different information, and so future-proofing data by contextualizing it for a diverse set of interests is important. By ensuring data has sufficient context for use by various audiences and users, open data initiatives can contribute to overall public value creation (Helbig et al., 2012).

Zuiderwijk et al. (2012) argue that very little attention is paid to metadata in the process of opening data. According to Dawes (2012), one needs to understand the processes of data collection, management, access and dissemination in order to fully understand data and it potential use. Often given low priority, the metadata provided with open data is regularly insufficient for the needs of data reusers, including just enough detail to be used by the original

users of the data (Dawes, 2012; Peled 2013). Common challenges include poor documentation, a lack of commonly agreed standards and ambiguous data definitions (Zuiderwijk et al., 2012)

Dawes (2010) uses the term stewardship to describe the responsibility for assuring the accuracy, validity, security, management, and preservation of information. The author suggests all public officials and government organizations should be stewards of public information, regardless of its original purpose or source. Others scholars suggest open data disrupts the government's traditional role as the data owner (Helbig, Cresswell, Burke and Luna-Reyes, 2012; Davies, 2010). For Helbig et al. (2012), this means all stakeholders should be regarded as stewards of open data, with a joint responsibility to assure data quality.

The quantity of data made available by governments does not amount to much if that data is not actually in demand by users (Davies, 2013). The 2013 Open Data Barometer global report laments that in most countries key data sets for entrepreneurship and policy analysis are not available as open data, and when published are in non-standard formats (Davies, 2013). Janssen (2012) suggests some public servants may simply be going through the motions, 'ticking the open data box' with little consideration of the data's value.

Citing the U.S. National Institute of Standards and Technology, Dawes (2010) describes useful data as containing content that is helpful, beneficial or serviceable to the intended end-user, or that supports the usefulness of other information by making it more accessible. The author goes on to state that useful data should benefit a wide variety of public and private users, enhancing public access to government data and facilitating the combination and reuse of information for new purposes. To enable this, Dawes encourages the implementation of policy and practices that promote investment in information management, analysis and presentation

while providing incentives and guidelines to share data and related information resources.

Dawes (2010) suggests looking at lessons learned from inter-agency data sharing to inform those policies and practices, while others (Bates, 2012; Helbig et al., 2012) recommend engaging with developers, civil society, communities and domain experts to understand data demand, as well as to provide insight on the collection and analysis of that data. Helbig et al. (2012) warn each stakeholder will have their own distinct requirements regarding data quality, file formats, metadata and regularity of updates. The demands of one set of users may conflict with another and, without proper consideration, lead to the pursuit of competing goals within the overall strategies of data agencies (Helbig et al., 2012).

Selecting data sets for release must be recognized as a matter of balancing resources, time and effort. Pre-existing good data management practices at agencies supplying data will reduce the cost and effort needed to make data available and increasing the probability it will be easy to reuse (Helbig et al., 2012). Yu and Robinson (2012) emphasize two dimensions of data quality. Related to the issues presented above, the first describes actual or anticipated benefits of the released data, be it service delivery or transparency. The second is a matter of how the data is structured, organized, and published, determining how difficult it will be for users to use it and thus whether the data is adaptable or inert (Yu & Robinson, 2012). Adaptability, the authors argue, is independent of the subject matter of datasets. Machine readable data available for bulk download are more adaptable to new formats and uses, including combination with other datasets. “Adaptability may depend on not only the format of the data itself but also on the prevalence and cost of the human and technological capital necessary to take advantage of it" (Yu & Robinson, 2012, p. 30).

Tim Berners-Lee has set out a model for open data publication called the 'Five Stars of Linked Data' ("Openness Rating", 2013). Each level on Berners-Lee's scale represents a step toward making it easier to use data with other data. Interoperability is a key factor to innovation through the use of open data, and denotes the ability to combine different datasets in 'mash-ups' to reveal new or more comprehensive results. In addition to standard licenses that allow data from different sources to be used together, interoperability is also affected by file format, metadata, and common data definitions (Helbig et al., 2012; OKF, 2012; Peristeras et al., 2009; Uhlir & Schroder, 2007). The idea behind the model is for governments to post data online even if the above determinants are less than ideal, making improvements as time goes on (Davies, 2013). The lowest level simply requires the data to be covered by an open license, the next level that the data be machine-readable, and the third that the machine-readable format be non-proprietary. The four and five star levels of the model incorporate the use of linked data, which allows users to connect disparate data across the web using unique identifiers ("Openness Rating", 2013).

In the view of some scholars, it's possible that too much emphasis is being placed on these requirements, at the expense of releasing useful data. Dawes (2012) argues the demand for raw, machine-readable datasets has led to open data portals releasing "low hanging fruit" rather than potentially more valuable datasets that require extra work to be fit for release. Others suggest governments may first move to release politically favourable data, serving their own goals while avoiding any uncomfortable increases in public scrutiny (Yu & Robinson, 2012; Worthy, 2013). Some even argue open data initiatives could potentially hinder the transparency movement if initiatives choose to prioritize the release of certain types of data (Janssen, 2012;

Bates, 2012), although Peled (2013) suggests there is little hard evidence to support these claims. Bates (2012) argues that while open data is being hailed for politically progressive goals like transparency, participation and innovation, multinational corporations and financial market players are also interested in open data. The author warns the open data movement risks being perceived as, or actually becoming, a subsidy for these commercial and market interests.

While critiques of open data initiatives are emerging, many advocates see it as more important to make data open and deal with resulting issues afterward. "It is difficult to make a (non-reactionary) argument against the free flow of data and information, however, it is crucial to understand the implications of co-opted partial approaches to openness, particularly if they risk weakening public institutions already under significant political and market pressure" (Bates, 2012, p.6).

2.6 Summary and Conceptual Framework

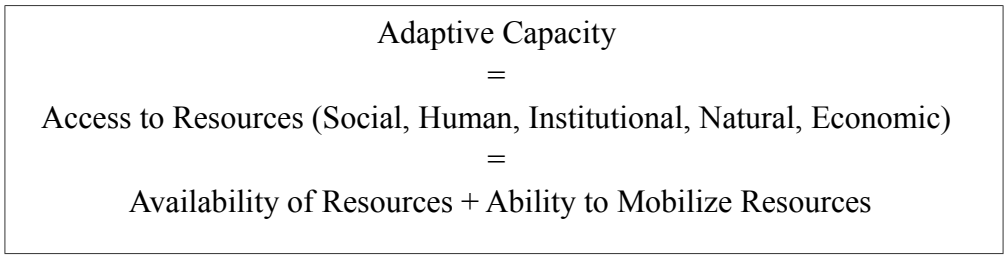

Figure 2: Adaptive Capacity according to Wall and Marzall (2006)

Wall and Marzall (2006) argue adaptive capacity, at a variety of scales, is dependent on access to social, human, institutional, natural and economic resources. Access, as they define it, includes both the availability of resources and the ability to mobilize them. While the authors stress that

appropriate indicators of access to each resource will vary by community, the review of literature above attempts a general exploration of each of these resources in the context of adaptation.

Wall and Marzall (2006) make no suggestion as to how one may assess the level of access to each category of resource except to state the need to balance input by specialists with input from the community. This is not so much a gap in the literature as it is a necessary omission given their argument that the make up of each category of resource will vary by community. As such, this research explores the general possibility of using open data for such an assessment. Because it is outside of the scope of this study to determine which indicators are of value to the community used as a case study, the research instead explores the potential for the effective use of available open data relevant to the community and, generally, to each category of resource.

Much has been written about possible benefits from open data applications although the literature shows little consensus on how those benefits should be measured. Without established indicators measuring the success or failure of open data initiatives, it seems difficult to provide exemplary case study of open data being used effectively, let alone any evidence-based formula for success.

Despite this, the existing literature does offer many recommendations around the necessary conditions for the effective use of open data. Many of these conditions are brought together in a single framework by Gurstein (2011). The review of literature explores and expands on issues identified in this framework, including data format, quality and standards; the user's ability to manipulate and interpret data; and, the policies of the government supplying

data to the public. These considerations form three general categories in the adapted framework for effective use of open data presented below and employed throughout this study.

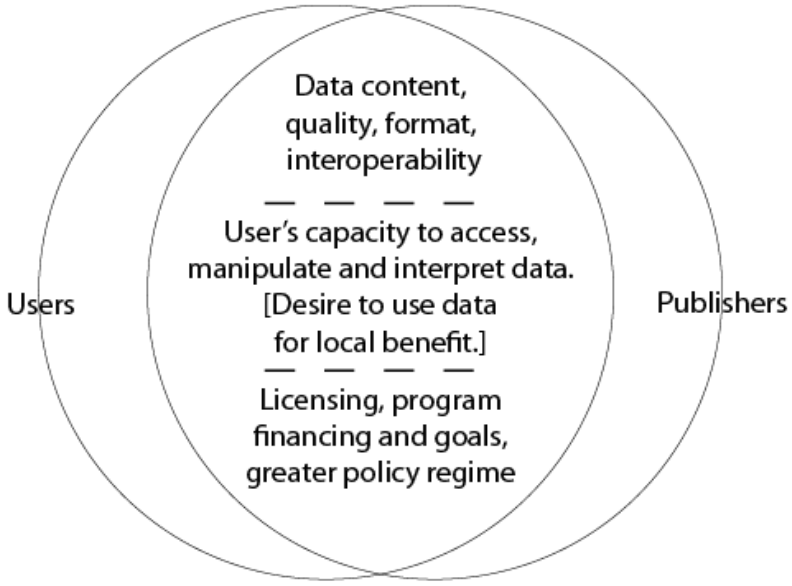

Figure 3: Effective use of open data, adapted from Gurstein (2011)

The study presents these elements as falling somewhere along a spectrum between the 'user' and 'curator' sides of the open data community of interest. The duality of user/curator is, of course, a simplification as public servants are also identified as users of open data, while users in the community of interest often contribute to the process of releasing open data through consultation. This research also acknowledges the possibility that a mosaic of interests outside of the community of interest can also influence the three categories of elements.

The literature related to the elements that make up this framework tend to generalize the contextual factors that may affect the effective use of open data. This research will attempt to address one aspect of this contextual gap by looking specifically at challenges to the effective

use of open data in small and rural communities, on both the user and supply sides of the open data community of interest.

Seeking to explore the potential for open data to be used in assessments of adaptive capacity, this study brings together the frameworks of Wall and Marzall (2006) and Gurstein (2011) into a single conceptual framework, presented in Figure 4 below. The conceptual framework posits open data as a potential source of the measurements needed in community assessments of adaptive capacity. However, this is only possible if the conditions necessary for effective use of the data are met.

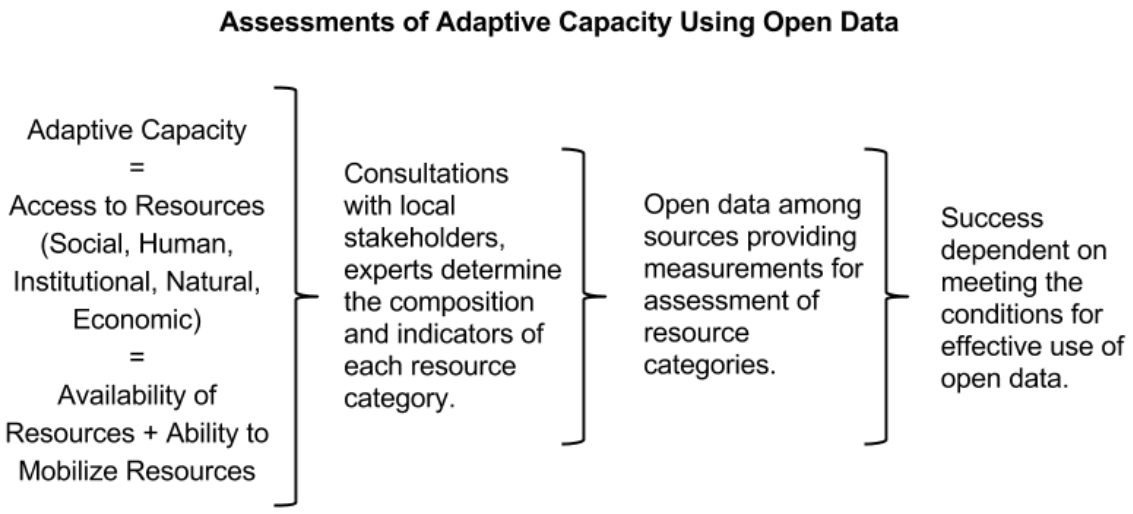

Figure 4. Assessments of Adaptive Capacity using Open Data

Chapter 3 – Methodology

This chapter presents the researcher's ontological and epistemological approach before going on to explain the study's mixed methods approach. The study uses a combination of key informant interviews, a systematic review of grey literature and an evaluation of available open data to explore challenges to open data use in small and rural communities and the suitableness of open data for assessments of adaptive capacity. The chapter closes with a discussion of the study's limitations and the researcher's attempts to address them.

3.1 Ontological and Epistemological Approach

The researcher's ontological approach lays in systems theory, a worldview of causal networks composed of relationships between actors or entities each with it's own set of properties dictating behaviour. This approach is appropriate given the goal of increasing adaptive capacity of communities in the face of complex problems, like climate change. Systems theory posits that the networks that make up a system are organized at a variety of levels, with higher level networks encompassing lower ones, but not merely a direct result of their sum. Such thinking is visible not only in this study's discussion of adaptive capacity, but also in the importance it places on data supplied by multiple agencies at three levels of government, each with its own policies and practices affecting the potential for public to make effective use of that data.

Wimsatt (1994) describes the entities that act as nodes in causal networks as having properties which are knowable 'robustly'. Something is robust, according to the author, if it is "accessible (detectable, measurable, derivable, definable, produceable, or the like) in a variety of

independent ways" (Wimsatt, 1994, p.208). Wimsatt reasons that the greater the variety of means of access yielding the same results, the greater the probability of being wrong can be minimized. "Robustness has the right kind of properties as a criterion for the real, and has features which naturally generate plausible results. Furthermore, it works reliably as a criterion in the face of real world complexities, where we are judging the operational goodness of the criterion--not its goodness under idealized circumstances" (Wimsatt, 1994, p. 209).

This provides the premise for one of this study's research questions. Adaptive capacity, according to scholars (Smit & Wandel, 2006; Wall & Marzall, 2006; McCarthy, 2001), is a property of a system produced by relationships between various entities or nodes within that system. This study asks whether adaptive capacity can be assessed using open data: recorded observations of things, events, activities and transactions which may or may not coincide with the nodes deemed important for adaptive capacity.

The same premise can also be explained using a critical realist's approach. As described by Sayer (2000) critical realism sees room for both interpretive understanding and causal explanation in the social sciences. The author differentiates between intransitive and transitive dimensions, the former being 'real world' phenomena and the latter being the way we come to know that phenomena. This not only supports the idea that open data is meaningful, but that it can reused, in ways possibly different from their original purpose, for novel applications.

Furthermore, critical realism also distinguishes between "the real, the actual and the empirical" Sayer (2000), The 'real' in this definition refers to the structures and powers of objects, the actual refers to what happens if and when those powers are activated. Using critical realism as an epistemological approach, this study looks at a selection of open datasets as 'real'

objects and attempts to distinguish whether or not they have the properties needed to be used effectively.

A second epistemological approach figures largely in this study. Social constructivism is employed in the exploration of barriers to effective use of open data conducted through the analysis of a systematic review of grey literature and key informant interviews conducted with the open data community of interest relevant to the case study. Such an approach is well-suited for this particular topic as the concept of open data is in many aspects one that is still being defined, with different and sometimes competing interests influencing the implementation of government open data initiatives.

Communities of practice, as described by Wenger (2000), are the basic building blocks of a social learning system. Built through mutual engagement, they are bound together by a collectively developed understanding of the community's identity and shared communal resources in the form of language, routines, tools, styles and more. A variety of communities of practice play a role in this study, including public servants, civic activists, technologists and data librarians. It is necessary to understand the needs of each of these communities to construct a larger image of effective open data use.

Communities of Interest can be thought of as “communities-of-communities”, making use of multiple knowledge systems associated with a variety of communities of practice (Fischer, 2001). Community's of interest are often more temporary than communities of practice, forming around a specific project. Given their composition, communities of interest can face challenges in building a shared understanding of that project, as multiple knowledge systems and interests are negotiated. Shared references, or 'boundary objects', meaningful across

those systems allow theme to interact (Fischer, 2001). Open data itself may be viewed as a boundary object within the community of interest examined in this study.

3.2 Methodology

Wall and Marzall (2006) argue the five categories of resource that make up their framework for adaptive capacity will differ among individual communities according to their unique context. Many of the 'supply-side' elements in this study's limited framework for effective use of open data also depend upon data management practices that can differ according to the organization releasing data. As such, this research will use the Region of Waterloo as a case study. A number of local open data projects in the area make this an ideal case study to explore the potential of open data in rural communities. The Region of Waterloo's open data program is one of three regional open data programs in the province covering rural areas. Separate open data programs at the cities of Waterloo and Kitchener and University of Waterloo will not explicitly factor into this project, but their existence may contribute to the number and capacity of open data users in the Region of Waterloo. While the region is home to a vibrant tech sector and two universities, it also contains four townships containing a mix of small settlements and rural areas.

This study employed a mixed methods approach, with procedures undertaken concurrently. Key informant interviews and a systematic review of grey literature have been used to identify challenges to the effective use of open data in rural communities. A comprehensive inventory of available open data sets relevant to the Waterloo Region has been compiled and evaluated both for the potential to assess local adaptive capacity to climate change and the presence or absence of properties supporting effective use.

3.3 Evaluation of a Data Inventory

In order to evaluate the use of currently available open data to assess adaptive capacity in rural communities, the study has created an inventory of all open data geographically relevant to the Region of Waterloo available through regional, provincial and federal government open data portals. Lists of all data sets available at the three open data portals under examination were downloaded on June 3rd, 2014. On this date, 191,759 were available on the federal open data portal, 212 on Ontario's portal and 32 on the Region of Waterloo's portal. With commitments made to continue adding to available data at each level, it is possible that the number of data sets available at these portals will have changed by the completion of this study. Because such a large number of datasets were available from the federal portal, a statistically significant sample of the total population was used for this study. The sample was stratified to accurately reflect the proportion of datasets originating from various ministries in the total population. The sampling process is documented in Appendix 1.

Once full catalogues had been downloaded and sampling completed, datasets from each portal were assessed for their relevance to the Region of Waterloo based on the geographic coverage and granularity of the data (ie: are provincial or federal datasets broken down in such a way that data specific to the Region of Waterloo is discernible from other areas covered?).

The collection of datasets were categorized as being related to one of the five categories of resource presented in Wall and Marzall's (2006) framework of adaptive capacity: social, human, institutional, natural, economic resources, or assigned to a sixth category, Other. The criteria used for the categorization, based on descriptions found in the framework, is presented in Appendix 2.

This study recognizes that community stakeholder engagement is needed to truly evaluate these datasets for their potential as indicators of adaptive capacity, as the composition and importance of each resource category is defined by the priorities of the community itself (Wall & Marzall, 2006; Smit & Wandel, 2006). Such a process, however, would require time and resources beyond the means of this project. Instead, the study will be limited to describing the datasets available in the broadest definition of each category.

The findings section (Chapter 4) presents descriptive statistics of the regionally relevant datasets by resource category and data type. Tests of difference will be conducted to examine whether the proportion of datasets in each category relevant to the Region of Waterloo are statistically different to the proportions present in the general population of datasets.

Next, each dataset was evaluated at the ordinal or interval level for its belonging to one of the five resource categories for adaptive capacity presented by Wall and Marzall (2006), their relevance to the Region of Waterloo and the presence of conditions for usability found in the expanded framework for effective use presented in this paper's review of literature. These include the dataset's level of openness, the availability of contact information for data stewards, the level of description for a dataset, information on collection/creation methods, the collection date or period of coverage, the frequency of data updates and the availability of other increments for time-series data. Despite some datasets included in the evaluation being listed on one of the three central government open data portals while actually hosted elsewhere (eg. a ministry's own data portal), the evaluation employed in this study is limited to the information available on the top-level, central data portals.

The criteria used for each category of evaluation are presented in Appendix 2. These criteria have been defined in such a way as to eliminate as much ambiguity as possible in the study's findings and to facilitate reproducibility of the research results. Frequencies were calculated for the exclusive categories of each criteria, at each level of government. Cross-tabulation was used to determine how much of the data in each category was also found relevant to the Region of Waterloo.

3.4 Key Informant Interviews

A total of thirteen (13) key informant interviews were conducted for this study. Sampling was purposeful, with the total sample stratified into different perspectives in the open data community of interest. Four (4) interviews were conducted with public servants working with government open data portals. These four included one (1) interview at each federal, provincial levels of government and two at the upper-tier municipal (regional) level of government. A further two (2) interviews were conducted with public servants that work or have worked at one the lower-tier municipal governments in the Region of Waterloo. These participants were selected for their familiarity with those municipal governments and their likeliness to be using public data in their day to day work.

The other strata in the sample was initially composed of the open data community of interest selected from the *Open Data Region of Waterloo* Google Group (N=70). Potential participants were selected based on their activity in that group and in other related events such as open data conferences and hackathons. Five (5) interviews were conducted with open data advocates, a combination of civic activists and computer programmers. Finally, two (2) data

librarians were interviewed for their expertise in information use and management. New participants were sought until saturation was achieved based on repetition within the interviewees' responses.

Interviewees were sent an informational letter about the study prior to agreeing to participate, and invited to ask questions. Written consent was obtained prior to beginning the interviews, with oral consent given to have their responses recorded using a digital recorder. All participants were reminded of their right to skip questions they were not comfortable answering, or to exit the interview at any time.

Interviews were conducted over March and April of 2014, and took between 45 minutes and an hour and a half each. In some cases, follow-up questions were sent via email when clarification on specific comments were needed. The majority of interviews were conducted in-person, while two were conducted via Voice over Internet Protocol software (Skype). In one case, following multiple interview requests and a lengthy approval process, questions were emailed and written responses were supplied.

All interview data was alphanumerically labelled by source and transcribed. All recording are stored on an encrypted, password protected hard drive stored in a secure location. Transcripts were imported into the NVIVO software package for qualitative analysis. Open coding was applied to the data to search for common themes among and between publisher and user respondents relevant to the study's research objectives. Data analysis files are also stored on an encrypted, password protected hard drive stored in a secure location. The final coding framework is presented in Appendix 5, while a summary of the results presented in the findings section of this document.

3.5 Systematic literature review

Systematic literature reviews should use a clearly formulated question to identify relevant studies, appraise their quality and summarize the collective evidence using an explicit methodology (Keele, 2007). The well-defined, explicit methodology reduces the risk of a biased selection of literature, while allowing for the results to be reproduced. With respect to the research question, if the literature yield consistent findings, the review suggests these to be robust and transferable. If the findings are inconsistent, the source of variation should be further explored (Keele, 2007; Khan, Kunz, Kleijnen, and Antes, 2003).

The systematic review of grey literature employed in this study is unusual in that it focuses entirely on grey literature rather than peer-reviewed articles. This is an intentional choice made to survey the open data discourse amongst governments and advocacy groups in Canada. Grey literature, in this study, includes documents (web pages, memos, conference proceedings, white papers, reports and working papers) from both government bodies and research organizations. Parameters for inclusion in the study include that the document:

1) was produced in the last ten years;
2) was produced by or directly address Canadian open data initiatives,
3) explicitly discuss the use of open data in or with benefit to rural areas or small municipalities.

Systematic literature reviews in the social sciences, according to Papaioannou et al. (2010), are especially challenging due to an ambiguity of terms across multiple databases. Search terms included different combinations of the terms "open data", "rural", "effective use", "uptake".

Wildcards and truncation were employed where appropriate. To fit within the scope and scale of this research project, records reviewed for the criteria listed above was limited to the first two hundred (200) items returned on online searches.

Three general areas were targeted in the search for literature: a general scan of the world wide web, Canadian content databases offered through the University of Guelph, and government websites at the three levels being examined in this study. The Google search engine was first queried for these terms in general, and then in a narrower search of domains specific to government resources in Canada (.gc.ca, .on.ca). The parliamentary libraries of both the federal government and the province of Ontario were searched, as were the websites of the agencies responsible for their open data agencies (the Treasury Board of Canada and the Ministry of Government Services, at the federal and provincial levels, respectively). The website of the Region of Waterloo was searched using the site's own query tool, as were the sites of the cities of Waterloo, Kitchener, and Cambridge. The websites of each township in the Region (the townships of North Dumfries, Wellesley, Wilmot and Woolwich) were small enough to search manually. Finally, two databases covering Canadian policy, trade and research publications offered through the University of Guelph, the Canadian Research Index, the Canadian Business and Current Affairs Complete database and Ebrary. The final number of results included from each source are included in a table in Appendix 5.

Each result was evaluated for inclusion using the following criteria 1) the document must be published in or after 2004; 2) the document must directly address Canadian open data initiatives. Finally, the reference lists included in any of the reports found using these criteria were checked for further literature meeting the parameters for this study. A final list of

documents included in the study is included in Appendix 6, referenced in APA style with the addition of a brief summary of each resource.

Open coding has been applied to each document individually seeking to identify themes related to the use of open data in small communities and rural areas. The coding framework was the same used for key informant interviews, and is presented in Appendix 4. The body of literature reviewed and the results of the open coding process are presented in the findings section of this document.

3.6 Limitations

Four main limitations affected this study, two related to conducting key informant interviews: the sample of interview participants; a lack of cooperation at the federal level of government; limited resources for stakeholder engagement; and, personal bias. These were overcome through the use of inter-method triangulation and careful treatment of the study's findings.

Key informant sample

While the Region of Waterloo does have an active open data community of interest, this community primarily resides within urban areas of the region. Very few of the key informants interviewed for this study reside in the Region's townships, and those that did stated they were unaware of any other open data advocates in their respective area. While most respondents still felt they could speak to the subject generally, a few limited their answers so as not to speak for a population they felt they could not represent. This limitation will be accounted for in coding and

analysis, while a systematic review of grey literature will provide another source of data for the research questions the key informant interviews were meant to address.

Censored responses

Entering interviews with public servants, it was expected that participants would be responding as representatives of the government. An attempt to interview a public servant from the federal level of government, however, surpassed expectation. The request was forwarded to a communications department, which requested the interview questions for screening. Despite regular contact by the researcher, approval from the communications department was held for six weeks. Furthermore, rather than an interview with a live representative of the open data agency, the communications agency responded with a set of written responses, limiting any opportunity for follow up questions. The study thus recognizes and notes that these responses were carefully controlled and are weighted as such in analysis. This data is also balanced by documents on the federal open data initiative in the systematic review of grey literature, prepared by authors with greater independence.

Limited resources

As noted in the review of literature and methodology chapters of this thesis, an assessment of adaptive capacity must be undertaken by stakeholders in a community. They alone are able to define the composition of the five categories of resources for adaptive capacity, as described by Wall and Marzall (2006) that suits the context of their own community. The resources available for this research were not sufficient to undertake such a prolonged process with the communities

that make up the case study for this research. Instead, the findings of this thesis regarding the use of open data to assess adaptive capacity has been limited to describing the availability and useability of open data that is relevant to the Region of Waterloo. This study does not claim to assess adaptive capacity in the Region of Waterloo, nor to prescribe the correct combination of resources that need to be secured in order to bolster that capacity.

This study also does not attempt to make any conclusions on how different data management practices or other characteristics related to the open data initiatives at each level of government affect the attributes of datasets measured below. It is limited to describing the attributes thought to contribute to effective use among the data at each portal without seeking any cause or correlation between those attributes and characteristics of the agencies that supply them. While this is outside the scope of this study, future research may attempt to do this with a more in depth examination of the policies and practices of each open data portal.

Personal bias

Over the course of this research, the researcher has been involved in organizing open data events and participating in 'hackathons' using open data, albeit as part of the Toronto community of interest. While care must be taken not to let personal experience influence the findings of this study, an understanding of knowledge systems associated with the open data community of interest may also have benefited the ability of the researcher to comfortably engage with participants in key informant interviews to uncover findings that may not otherwise have been accessible.

Chapter 4 – Findings

The following chapter presents findings from the study's qualitative and quantitative data collection. An analysis of open coding from key informant interviews and a systematic review of grey literature on open data first address the barriers faced by small and rural communities interested in using open data. These sources are also used to add insight to a series of quantitative descriptions of open data currently available to the Region of Waterloo as the study seeks to assess that data for qualities thought to contribute its potential for effective use.

The qualitative portion of these findings include interviews with public servants from the agencies responsible for the three open data portals examined in this study. One interview was conducted each with public servants from the federal and provincial government and two with the upper-tier regional government. It should be noted that the request to interview a public servant working directly on open data was redirected to Treasury Board of Canada's communications department. Responses to interview questions were provided to the researcher in writing via email. Two public servants with experience working at lower-tier municipalities within the Region of Waterloo were also interviewed. These key informants can be viewed as representing both the supply and prospective user side of open data, with interviewees providing incite on the potential challenges of implementing open data initiatives in small municipalities while also responding to suggestions that public servants were among those who may benefit from open data. The remainder of the study's 13 key informant interviews can be viewed as coming from the user side of the Region of Waterloo's open data community of interest, and include developers, civic activists, data librarians. However, most of these informants, and

especially the data librarians, felt they could provide insight to questions regarding the supply of open data.

A systematic review of grey literature, described in the methodology chapter of this study, has been used to test the validity of responses from key informants through triangulation. The selection of grey literature was limited to material directly relevant to Canadian open data initiatives and community of interest. A more detailed summary of parameters used in the systematic review of grey literature can be found in Appendix 6. It should be noted that many influential studies on open data have been written in the Europe, the UK or the US and thus fall outside of the parameters of this systematic review of grey literature.

A number of the documents included in the systematic review of grey literature represent findings from and submissions to government consultations with open data stakeholders, while others are records of evidence presented by open data experts to two federal committees studying open data in the last three years: the Standing Committee on Access to Information, Privacy and Ethics and the Standing Committee on Government Operations and Estimates. Other material includes relevant reports from government and non-government organizations, government press releases and presentations, minutes from the legislative floor and news, magazine and blog articles written by members of the open data community of interest.

Quantitative findings presented below have been segmented to describe the Regional, Provincial, and Federal open data portals separately. This has been done for two reasons. First, such segmentation may better reveal the practices applied by open data stewards at each level of government. As such, key informant interviews with public servants at those initiatives have been examined for further insight when juxtaposing the characteristics of each portal.

Second, the federal open data portal has much larger catalogue of datasets than the other two portals examined. As such, analysis of the federal open data portal used a stratified random sample from the overall catalogue of datasets, whereas the Regional and Provincial catalogues were evaluated in their entirety. Findings from the evaluation of datasets from the federal portal are therefore treated separately in order to report their margins of error.

While the quantitative assessment describes the open data catalogues at each level of government, cross-tabulation has also been used to isolate the characteristics of the datasets found to be geographically relevant to the case study. Although it may be possible to classify some of the results as ordinal data, the majority is nominal. As such, all cross tabulation has been limited to statistical tools available to analyze nominal data.

4.1 Challenges to using open data faced by small and rural municipalities

While a wide range of issues related to the use of open data were discussed during key informant interviews, informants saw many of the problems faced by rural communities as being the same as those faced by urban communities. General challenges to the use of open data by small and rural communities uncovered in the qualitative findings are discussed below, but the point of focus remains those felt more acutely by small and rural communities.

When seeking to identify the challenges small and rural communities face in using open data, it should first be noted that many of the informants interviewed on both the user and supply side of the open data community of interest questioned whether the general population would even be aware of the concept of open data or its suggested potential. One informant felt

that a lack of understanding of open data was one of the biggest challenges to the success of open data in smaller communities. Another suggested that while awareness of and vocal support for open data exists in small pockets, broader support was needed. Awareness was seen as an important factor for the sustainability of open data initiatives, as it would both increase demand for data among the public and provide incentive within the public sector to comply with open data demands.

The onus for raising awareness was viewed by interview respondents as a responsibility shared across the community. Key informants representing government open data initiatives spoke of their agency's efforts to raise awareness online, through public consultations, hackathons and conferences. A number of respondents from the user side of the community of interest also spoke of their roles in organizing open data conferences and in making presentations on the topic to raise awareness amongst colleagues.

Similar concerns around awareness were expressed and addressed in materials examined in the systematic review of grey literature. Findings from recent consultations initiated by the Ontario government suggested the government needed to further promote the economic benefits that may be derived from the use of open data, and recommended more public engagement efforts and the creation of permanent office devoted to open government as ways open data may be promoted (Government of Ontario, 2014b).

At the federal Standing Committee on Access to Information, Privacy and Ethics, Members of Parliament expressed concern that the public in their constituencies had never heard of open data (SCAIPE, 2011d). Presenters from municipal open data initiatives stressed that awareness raising, through the media, consultations, contests and partnerships with universities

and colleges, had been a major focus of their efforts (SCAIPE, 2010e). While some experts suggested for awareness campaigns using advertising, others argued it would be more effective to reach out to strategically placed figures or bodies to raise awareness among prospective users (SCGOE, 2014b, SCGOE, 2014d, SCGOE, 2014e). Experts from the federal open data initiative suggested that the efforts of federal Treasury Board President Tony Clement had done much to raise awareness of open data around the country (SCGOE, 2013). Indeed, the systematic review of grey literature includes many press releases from events the Treasury Board initiated over the last year with the President meeting leaders from industry and academia.

A wide range of prospective users were discussed in this study's systematic review of grey literature, broadly spanning businesses in all manner of industry; software developers; public servants; academics; non-profit organizations; transparency advocates; journalists; and, civic-minded citizens. Stakeholders at a national consultation on open data undertaken by the federal government recognized that not all prospective users would have the same demands, but urged the government to focus on meeting the needs of a select set of users rather than provide a mediocre service for all (Government of Canada, 2014e).

> Transparency advocates sought data about government accountability, academics cared about provenance, metadata and context, software developers and business sectors cared about data that would provide business intelligence or front line service information while local and provincial counterparts were interested in exploring how data from different jurisdictions could be rolled up and shared more effectively. And this is to say nothing of numerous other stakeholder groups, sub-groups or even the many Canadians who will come to the portal with little sense of purpose or potentially experience in working with data. (Government of Canada, 2014e)

Key informants admitted that the Region of Waterloo, with two universities and a vibrant tech sector, was an anomaly among smaller communities in terms of the number of developers local to the community. They suggested possible users outside of the Region's urban centres could include local journalists, community organizations, agriculturists, small business, researchers and public servants.

Public servants from open data initiatives interviewed for this study suggested they were not always aware of what the data their organization supplied was being used for, relying on social media and relationships with the community for such feedback. They emphasized that open data use is often unpredictable. Each of these supply-side key informants emphasized that there was no expectation or preferred user in mind when data is released.

When considering users of open data, a data librarian suggested that the people working with open data today are the 'keeners', and not the users that open data managers and proponents should be worrying about reaching. Most of those interviewed from the open data community of interest recognized that some people and communities would be in a better poised to use open data than others. The same concerns were found in the systematic review of grey literature.

While the majority of reviewed materials spoke to issues of accessibility and inclusive use in terms of web standards for accessibility and user-friendliness, others addressed the deeper issues of literacy, language, skills, resources and access. Some sources warned of a 'data divide', parallel to the digital divide and similarly resulting from differences in income, education and literacy (Bates, 2012; Gurstein, 2011; Government of Canada, 2014c). Issues identified as barriers to the effective use of open data during this study's qualitative analysis, on both the supply and user side of the open data community of interest, have been characterized in the

remainder of this analysis as capacity issues. These are presented in three categories below: technology / infrastructure, know-how, financial / human resources.

4.2 Challenges: Infrastructure/Technology

User access to technology and internet connectivity were not seen identified by key informants as particularly large barriers. It was acknowledged that internet connectivity can be an issue in some rural communities, but that was not seen as the case within the Region of Waterloo. Much of the data available on open data portals was not seen as particularly bandwidth intensive, and thus even slow connections were seen as sufficient. Informants felt that most users wanting to access open data or products of open data would likely have access to a home computer or to a smartphone to do so.

Sources examined in the systematic review of grey literature presented a more mixed view. Some experts presenting evidence to the Standing Committee on Access to Information, Privacy and Ethics suggested Canada on the whole was doing fairly well in terms of connectivity (SCAIPE, 2010a), while new, affordable technology also meant that what was once only possible for well funded organizations could now be accomplished by everyday citizens (SCAIPE, 2011f). Others admitted internet connectivity and device ownership to be a real issue in some rural areas (SCAIPE, 2011k; EWB, 2013). The federal government's recently released Digital Canada 150 action plan commits the federal government to ensuring that 98% of Canadians will have access to broadband internet at a minimum of 5 megabits per second by 2017. The plan hails the success of a $225 million investment in rural broadband already made by the federal government through the 'Broadband Canada: Connecting Rural Canadians'

program. It is beyond the scope of this study to comment on the success of that program or the goals of the new federal plan, but it can be stated that while the federal government recognizes rural connectivity as an important issue, it also seems optimistic that adequate measures are being taken to address it (Clement, Gurstein and Longford, 2012).

Key informants did, however, see technological infrastructure on the supply side of open data as a greater challenge in small communities. They suggested the cost of developing or licensing software and of purchasing the hardware needed for an open data initiative may be prohibitive to many small communities. A number of informants suggested these investments should instead be integrated into existing information technology (IT) renewal plans. These informants, from the user side of the community of interest no less, argued that the move to 'open by default' needed to be seen as a long-term process rather than something that could happen instantaneously.

A key informant working on the Region's open data initiative stated opening data was currently a very manual process for their organization, requiring input from a number of people and departments. Another informant with experience working at the Region's lower-tier municipalities stated that some of the townships in the region do not have their own IT departments, making any change to technology a larger undertaking.

Experts presenting evidence to the federal Standing Committee on Access to Information, Privacy and Ethics also recognized previously scheduled processes of 'evergreening' older technology as the opportune moment for public IT departments to integrate new infrastructure for open data (SCAIPE, 2010d). One expert suggested that switching to open formats may also provide greater leverage when negotiating with IT vendors, as the government

would no longer be dependent on systems that use specific proprietary formats (SCAIPE, 2010d). The Ontario government's recent report summarizing province-wide consultations on open data recommends that all procurement policies be amended to require that all new information technology purchases support open data while a strategy is developed to transition all IT systems to comply with this standard by 2017 (Government of Ontario, 2014b).

Other experts presenting evidence to the federal Standing Committee on Access to Information, Privacy and Ethics reminded the committee that policies must also change to reflect the updated technology used for open data (SCAIPE, 2011b; SCAIPE, 2011h). As one expert pointed out, the concept of 'Freedom of Information' was created in the era of hard copy documents (SCAIPE, 2011e). An expert presenting to the Standing Committee on Government Operations and Estimates noted that open data is currently seen as something done after the data has been created or collected, making compliance a major problem. In their opinion, this would cease to be the case if open data were integrated into the infrastructure as part of the overall process (SCGOE, 2014c).

4.3 Challenges: Know-how

While local key informants acknowledged that the Region of Waterloo was likely to have an exceptional level of capacity to work with open data, it was suggested by a number of those interviewed that the capacity to do so was likely less among residents of the Region's smaller settlements and rural areas. A submission by the Media Awareness Network to the federal Digital Canada 150 initiative argues the two main barriers to digital literacy in rural Canada are a lack of infrastructure and attitude, with the latter referring to culture and educational levels

(Government of Canada, 2014d). A successful strategy for digital literacy, the submission argues, must address barriers such as attitude, age, socioeconomic status, language and regional availability of resources, targeting specific populations and situations where needed (Government of Canada, 2014d).

A lack of capacity to use the data, however, was not seen by key informants as a valid argument against releasing it. One informant used the analogy of libraries to suggest that making data available would actually lead to greater data literacy, as making data available provides raw material with which one can learn. For another, providing data related to a user's interests also provides motivation to learn to use the data. Experts presenting evidence to the federal Standing Committee on Access to Information, Privacy and Ethics were also of the mind that having open data available could lead to greater data literacy (SCAIPE, 2010b; SCAIPE, 2010d; SCAIPE, 2010e). The experts argued data should be made available even if it did not meet all criteria for accessibility, and suggested the public has a greater ability to deal with complex information than is widely assumed (SCAIPE, 2010e).

While it was acknowledged that it would be a long-term solution, key informants agreed that data literacy should be incorporated into the educational curriculum. The issue of data literacy was seen as generational, with so-called 'digital natives' already more apt to make use of data even without formal training. A librarian amongst the respondents suggested a typical user's knowledge need not be particularly advanced, but enough to critically read and interpret data that can at times be messy and inaccurate. The same respondent summarized data literacy as “a matter of use, interpretation and critical thinking”.

Addressing data literacy was not solely seen as the responsibility of government but also that of the community advocating for open data. While one community member noted they work with a community group on programming to get people interested in open data, the informant stated this was merely a matter of planting the seed. It was suggested by a number of key informants that enough training to perform simple operations with data can motivate people to learn independently using other resources. Key informants believed that although rural areas may not have immediate access to the educational institutions found in urban areas, adequate learning resources were freely available to any user with an internet connection. Some informants suggested there can be cultural or generational barriers when asking people to learn new things, but that these are not limited to small towns or rural areas.

Educational programming on data literacy is called for in reports commissioned by provincial agencies (IPCO, 2011) and by stakeholders consulted in the federal 'Open Data Roundtable' consultations completed in 2013. The findings of the latter also states many of the stakeholders felt the federal open data portal would itself be an ideal space for hosting educational resources on open data. A representative of the federal open data program presenting to the Standing Committee on Government Operations and Estimates (SCGOE, 2014b) point out there are already some basic resources on the federal portal, as did the spokesperson for the portal interviewed, via email, for this study.

While arguing digital literacy should be supported by all levels of government, a report by the Stratford Institute for Digital Media suggests it can be easier to assess needs and deliver such programs at the local level (Stratford Institute, 2012). The report quotes the Canadian Research Alliance for Community Innovation and Networking as suggesting local community

networks can provide the needed social infrastructure for information and communication technology (ICT) access and adoption. Other sources reviewed identified initiatives such as the Community Access Program, Canada Job Grant and Computers for Schools Program, public libraries and the private sector as all possibly playing a role in developing public digital literacy skills (EWB, 2013; Government of Canada, 2014b; Stratford Institute, 2012; Government of Canada, 2014d).

4.4 Infomediaries

Many of the key informants interviewed for this study did not feel open data would only benefit a minority due to a disparity in the public's capacity to directly take advantage of it. One key informant stated the importance of accounting for the overall data life-cycle when considering the benefits derived from open data. Many key informants from the user community envisioned primary users processing raw open data to produce applications or value-added information that would ultimately be consumed by the public. As such, the need for wider data literacy would be eased through consumable products created by developers, research organizations or data journalism, often described as 'infomediaries'. An informant for this study pointed out that infomediaries could also overcome more general technological limitations, as reports can be printed or integrated into more traditional broadcast programming. The informant suggested that data journalism in particular could lead to more public awareness, informed discussion, and better decision making.

Experts presenting to the federal standing committee on Access to Information, Privacy and Ethics also saw a need for infomediaries. Experts saw business, non-profits and academia as

infomediaries that may use data to create applications which Canadians may access through their own preferred channels (Routine Proceedings, 2012). One expert noted recent innovation in user friendly data visualization and story telling tools, suggesting these could be hosted on open data portals . Other reports examined in the systematic review of grey literature mentioned traditional media, maps and custom applications as possible mediums to deliver information derived from open data, but warned that such tools would have be tailored to diverse audiences and user groups (Geist, 2009; EWB, 2013).

A number of open data users interviewed suggested the community of interest could itself play a role in connecting people with technical expertise with others that have substantive interests in specific types of data. Academics, community activist groups and non-profit organizations were mentioned as possibly taking on the role of infomediary. While some interviewed saw existing users of open data as community-minded, others felt it was more realistic for people or interest groups without data skills to contract the services of infomediaries through financial remuneration. A developer among the key informants pointed out that while it is certainly possible for open data products to be developed in an open source manner, this depends on developers taking a personal interest in an issue or collection of data. The informant suggested that organizations wanting to develop a product or conduct analysis using open data will still see savings as access to the data is free.

Another informant suggested government could create an incentive mechanism to attract startups to address social policy questions, and dove-tail such efforts with open data. A short-term example of this may be the open data hackathon or appathon, events or competitions that bring together various combinations of coders, user experience experts, designers, and intended

consumers to create products from open data (FED-R-03, Baker, 2011; 27 Shift, 2013; SCAIPE, 2010c). Such events are lauded within some of the grey literature surveyed for this study as opportunities to spur creative solutions to public services, economic development and innovation (SCAIPE, 2010c). However, key informants and some of the other grey literature sources also expressed a certain amount of scepticism regarding the benefits of hackathons and the sustainability of their products . While key informants suggested they had taken part in such events in the past, some felt that open data providers or special interest groups relying on them to generate innovative and meaningful outcomes may be sorely disappointed.

Experts presenting to both the Standing Committee on Access to Information and Privacy and Standing Committee on Government Operations and Estimates recognized the need for infomediaries to maximize the benefit from open data. But while community or special interest groups were mentioned as possibly acting as infomediaries (SCAIPE, 2010c), the discussion just as often centred on how the need to fill these roles can create economic opportunities (SCAIPE, 2010c; SCGOE, 2014d). It was argued that there would still be benefit to the general public if a business was to add value to raw data and redistribute that product for profit (SCAIPE, 2010c). Others noted the success stories of the digital economy are made up of companies skilled in organizing and leveraging data (SCAIPE, 2010d), and that companies with existing and compatible software (ie. Google maps) will be well-poised to integrate government information into their programs for more efficient transmission to the public (SCAIPE, 2011e). The economic opportunities presented by open data were promoted at all levels of government examined in this case study, but especially by the federal open data initiative. In March of 2014, the agency sponsored a national appathon entitled the Canadian Open Data Experience, or

CODE. This was promoted in a series of high profile meetings between Tony Clement, the president on the Treasury Board, and leaders from the technology industry and academics emphasizing the event's entrepreneurial spirit.

While all key informants recognized the need for infomediary services and were aware of existing applications using open data to provide services for the broader public, there was little consensus on the extent to which infomediaries could supplant the need for data literacy programming. One librarian interviewed for this study felt that with the provision of visualization tools was a more immediate solution for less data literate users wanting to take advantage of open data. Another felt that while infomediaries were useful, it did not supplant the need to address wider data literacy. One informant even suggested an education in civics and service design were ultimately more important for users of open data than technical skills, as technology would eventually provide user-friendly solutions to data manipulation and/or computer programming.

While the Regional open data initiative suggested open data would allow developers to create services for others in the community in a report to the regional council (Region of Waterloo, 2013a), key informants recognized that some small communities would have fewer resources to draw on as infomediaries than their urban counterparts. Even within the Region of Waterloo, one informant suggested it was too early to say if the local tech community would take it upon itself to use open data to create broader social good. Conversely, another informant living in a rural part of the region spoke of the difficulty of finding other community members interested in using open data to address local concerns despite his own technological skillset.

It was, however, argued by one informant that developers in small communities may have more incentive to work with local open data due to an interest in local issues and lower expectations for remuneration, where larger companies may not foresee enough return on such an investment. Another key informant suggested data standards could eventually allow technological solutions developed for one community to be used by others, providing access without having to develop a local version of the same application. A key informant from the Region's open data portal stated the initiative was considering options to make data more accessible to people on the portal, particularly through visualization. Because resources were limited, the initiative was interested in adopting existing platforms, including the open source CKAN and solutions developed by other open data initiatives. Finally, a number of informants spoke of the continued importance of local journalism as a traditional form of infomediary and suggested resource-strained local newspapers could benefit from reporting using free and readily available government data.

4.6 Know-how on the supply side of open data

Know-how within the agencies meant to supply open data was also identified as a challenge by informants from the public sector. In addition to the data management skills needed to maintain an open data initiative, informants suggested public servants needed a level of data literacy sufficient to field public requests for data from both basic and advanced users; to increase the awareness of the value of the data being opened to the public; and, to use the data themselves to draft policy and design programs based on evidence-based decision making and business intelligence.

Many sources reviewed in the systematic review of grey literature also emphasized the need to build the capacity of public servants to collect, manage and use data. Experts presenting evidence at both the federal Standing Committee on Access to Information, Privacy and Ethics stated the importance of building the capacity of all public servants to use data, not just those working directly on making open data available to the public. A consultation by the provincial government on Open Data in Ontario also found there was a need to develop educational programs and tools to promote data literacy within the government.

A key informant from the provincial open data initiative indicated capacity building efforts were already underway with internal workshops and training courses taking place and more planned for the coming year. It was also suggested that as the province's open data initiative became more mature, ministries would establish and reinforce their own practices. At the time of this study's data collection, no capacity building efforts were underway within the Regional government. However, representatives from the region met regularly with public servants from other municipalities with open data initiatives along with representatives from the province to share experiences and learn from one another (Region of Waterloo, 2013b). It was suggested by a number of informants interviewed for this study, with direct reference made to townships in the Region of Waterloo, that the skillset necessary to maintain an open data initiative may not be available to smaller organizations hoping to implement an open data program. The ability to develop this capacity, however, was largely seen as limited by a lack of resources.

4.7 Challenges: Resources

In both interviews with the open data community of interest and the systematic review of grey literature, resource constraints are seen as some of the biggest challenges to organizations providing open data. As noted by an expert from Treasury Board presenting to the Standing Committee on Government Operations and Estimates, "a balance must be struck between the portal's curatorial needs (maintaining and updating the data, responding to questions) and respect for the organization's resource constraints" (SCGOE, 2014b).

The overall costs that stem from an open data initiative can be difficult to measure, as they largely rely on existing assets of government agencies (SCGOE, 2014b; Region of Waterloo, 2013a). The discussion of infrastructure capacity, above, notes obvious investments needed to acquire hardware and developing an online portal. But while financial costs were seen as an initial barrier to implementing open data in small communities, key informants and grey literature sources suggest the human resources aspect of implementing an open data initiative may be a greater obstacle.

Key informants seemed to agree that the challenge of resources was likely more pronounced in small municipalities. Those familiar with the lower-tier townships viewed them as already operating at maximum-capacity. One informant felt small municipalities may not even have the resources to learn about the value an open data initiative may offer them or their constituents. Another stated the long-term sustainability of open data initiatives were dependent on a proper allotment of resources, and claimed to have seen open data initiatives start enthusiastically only to fall away for this very reason. The informant suggested small

municipalities need not devote a full time person to open data, but that someone must be held responsible for maintaining the initiative rather than asking public servants to "do it from the side of their desk". A librarian suggested open data proponents in small communities could possibly step in to fill the gap in resources.

A number of tasks were identified by key informants as consuming resources in addition to basic data management. Public servants and data librarians interviewed suggested addressing privacy concerns through data aggregation and anonymization required extra resources. The same issue was noted in reviewed grey literature, while the creation of metadata was also viewed as particularly labour intensive by those consulted for the federal Open Data Roundtable. Stakeholders suggested user-generated metadata be included along with a disclaimer warning against any inaccuracies (Government of Canada, 2014e).The resources required to translate data into both official languages was a particularly contentious issue at the federal Standing Committee on Access to Information, Privacy and Ethics. Some experts before the committee argued that waiting until every dataset was translated would ensure that none of it would never be available. Instead, they suggested making as much data as open as possible, in particular datasets that were mostly numeric would not require very much translation, and trying to improve the offerings later (SCAIPE, 2010c; SCAIPE, 2011c).

4.8 Resources at upper and lower tier municipalities in the Region of Waterloo

Despite any concern for resources, a report to a meeting of the Region of Waterloo municipal council during the early stages of the local open data initiative suggested the initiative could be implemented without additional staff or resources (Region of Waterloo, 2013a). According to

the Region's Open Governance Model, the introduction of an open data initiative was not intended to replace existing practices, but enhance those that were already in place. It states the initiative “will make use of existing committees/working groups and approval structures so as not to create a burden on the organization and to take advantage of processes already in place” (REG-I-01). Interviews with public servants at the upper-tier municipality did not reveal the addition of any new staff, but suggested those working directly on the open data initiative balance those duties with previously existing job descriptions. This is not to suggest, however, that resource constraints are not recognized. When opening a dataset, part of the assessment undertaken by staff evaluates the level of resources needed to publish and maintain the data.

It is at the region's small, lower-tier municipalities that key informants suggested resource constraints are felt particularly hard. Public servants within the region felt the townships did not have the resources needed to compile and release data. One informant described the townships as 'piggybacking' off the upper-tier municipality, with most of the data they use internally coming from the upper-tier Region or provincial governments. While the lower-tier municipalities often contribute financially to data collection activities such as aerial photography, this must be planned well in advanced so that provision can be made in the townships' limited budgets. With limited staffing, public servants at the township level are already often expected to 'wear multiple hats', taking on responsibilities that may be divided among multiple employees in larger organizations. Informants felt assigning staff to implement an open data initiative would be a major resource challenge at the township level, as would finding the budget to pay for a dedicated position. One public servant with experience working

at the Region's lower tier municipalities questioned whether there was enough demand and benefit to justify a small municipality devoting its scarce resources to an open data program.

4.9 Savings

It should also be noted that while key informants interviewed agreed that an open data initiative would require the use of organizational resources, many also felt that open data would bring different forms of resource savings in the long run. Key informants and grey literature sources both indicated open data could lead to fewer requests for information as the public began to 'self-serve' via the open data portal, thus reducing costs associated with manual access to information requests (SCAIPE, 2010c; SCAIPE, 2011d; Government of Canada, 2014a; Mowat Centre, 2011). It was also thought open data would create more efficient internal data sharing while also reducing redundancies in data collection and storage (SCAIPE, 2011a; SCGOE, 2014c; Government of Canada, 2014c; Reitano, 2013). Key informants and multiple sources in the reviewed grey literature felt access to more data would lead to greater efficiencies in government through improved analysis and data-driven decision making (SCAIPE, 2010d; SCGOE, 2014c; Government of Canada, 2014c; KPMG, 2013; Eaves, 2012; Government of Ontario, 2014b), while analysis by the public may also reveal opportunities for savings (SCAIPE, 2010d; SCGOE, 2014c; IPCO, 2011; Mowat Centre, 2011). If nothing else, it was suggested in one report that greater transparency would lead some agencies to be more careful in their spending (Mowat Centre, 2011). Solutions delivered by developers, the private sector and non-profits using open data were also seen as potentially relieving some of the burden for public services from governments (Government of Canada, 2014a; SCAIPE, 2010c; SCAIPE,

2011d; Mowat Centre, 2011). Instead of having to spend their budgets to access data, community organizations, researchers and students would be able to invest more of their resources into delivering benefits for the public.

4.10 The usefulness of open data

Prioritizing the release of data seems a difficult proposition for those public servants working on open data initiatives. As limited resources meant governments could only release a finite amount of data in a given time period, public servants working on the provincial and regional portals interviewed for this study indicated that in the ideal situation, priority for releasing data was given to datasets with the most public value. However, the key informants were also quick to admit that it was difficult to predict the value of data or know how data would be used.

Key informants at the regional open data initiative indicated public requests for data could be made through email, social networks or over the phone. At the provincial level, the public servant interviewed also indicated demand for data was gauged through requests made through social media and email, as well as by reaching out to the open data community of interest at conferences and events.

While requests to the region's initiative were stated to be infrequent, one public servant suggested that this was when the process for releasing data worked best. The informant suggested a request for a specific dataset provided added justification to open the data when approaching data managers. Another key informant working on the regional portal suggested data was often selected for release by reviewing what was available at other municipal open data catalogues and looking for comparable data at the Region of Waterloo.

Key informants on the user side of the community of interest stated it was clear that they could request data, but that the larger challenge was not knowing what was available. This need appears to have been recognized at the provinical level where a province-wide inventory of data collected by various ministries was being created at the time of data collection. The inventory has since been posted online, inviting the public to vote on which datasets they wish to be released. While there is no clear indication of the follow up process given on the voting tool itself, a related provincial web page (Government of Ontario, 2013b) suggests the process for releasing a dataset can take from several months to a year or more, depending on the volume and complexity of the data.

Additional concerns raised by key informants around the demand for data were related to the opaqueness of the request process and timeliness of responses. One informant pointed out many people may not always know which level of government to approach for the data they sought. Another suggested that in addition to making requests for raw data the public should be able to suggest ideas for open data products that they may not be able to implement themselves.

4.11 General findings on the usefulness of open data

When asked if they felt that the data currently available on the three portals under consideration in this study would be useful to residents of the Region of Waterloo, opinions were very mixed. Many from the pool of key informants suggested data related to local services and local geospatial data could be of interest to the users in the Region of Waterloo. A public servant with experience working at the region's lower-tier municipalities did not see much value at all in

currently available datasets and felt some of the most generally useful local data collected by government could not be released due to issues of confidentiality.

Overall, the majority of informants suggested the usefulness of data depends on the interests of individual users, with many stating they could only speculate on whether or not users from the region's smaller settlements and rural areas would find any of the available data to be useful. Many felt that releases should not be limited to data with obvious use, because one cannot accurately predict all of the possible uses of open data.

Finally, when discussing the usefulness of or demand for data, informants were often as concerned with attributes of the data as the substance of its contents. Many of these attributes overlap issues raised in this study's review of literature and conceptual framework related to the concept of effective use. As such, the perspectives of key informants interviewed for this study and relevant perspectives from the systematic review of grey literature have been included at relevant points in this study's evaluation of data offerings available to the Region of Waterloo, found below.

4.12 Relevance to the Region of Waterloo

One of the more basic questions that needs to be addressed when assessing the possible use of open data to assess adaptive capacity in the Region of Waterloo is whether or not the data available is relevant to the local community. The quantitative approach in this study operationalized the concept of a dataset's relevance in terms of the geographical coverage of its contents to the Region of Waterloo, as well as the granularity of that content to the Region.

A dataset has been counted as covering the region geographically if it contains observations that cover or originate from within the Region of Waterloo. As such, although the figures are not specific to the Region of Waterloo, in this study statistical data for the province of Ontario is still taken as relevant to the Region. For geospatial data, administrative boundaries for provincial bodies would also be seen as relevant. Geospatial point data, however, is not viewed as relevant unless those points fall within the geographic boundaries of the Region. It is arguable that a lack of features in the Region (for example, no occurrence of forest disease) is still relevant, and this has been noted in the study's limitations.

Table 1. Relevance of Data on Provincial and Federal Open Data Portals to the Region of Waterloo

	Provincial (N=212) % of total catalogue	Federal (n = 385, CL=95%) % of total catalogue
Not relevant	15.6	94.3 (±2.3)
Relevant, not granular	37.3	5.2 (±2.2)
Relevant, granular	47.2	0.5 (±0.7)
Total	100.0	100.0 (±0.0)

Unsurprisingly, all data on the Region of Waterloo's open data portal was found to be relevant and of a granularity at the Regional level or finer. Results of the evaluation at the provincial level were more mixed, with almost half of all datasets on the portal found to be relevant and granular at the Regional level or finer. A further 79 of the 212 datasets on the portal were also found relevant, albeit at a coarser level of granularity than the Regional level.

The evaluation of the federal portal in this regard yielded quite different results. 94.3 per cent (95% confidence interval [CI]: 92.0% - 96.6%) of the datasets in the sample were found to be irrelevant to the Region of Waterloo. This may be due to the way the stewards of the federal open data portal have chosen to list its offerings, which at the time of data collection totalled 206,278 datasets.

As noted below, the sample suggests that most of the federal catalogue is made up of geospatial data and satellite imagery. An inspection of this data on the open data portal reveals that rather than presenting all selection data of this kind originating from a single database under a single record in the catalogue, each file is given its own record. As such, many of the 385 records selected in this study's sample actually originate from only a handful of sources.

Thus, the sample of datasets from the federal open data portal used in this study represents the overall proportion of relevant datasets in the catalogue based on the inflated manner that the portal stewards have chosen to present their data. While the original data sources may have included imagery covering the Region of Waterloo, none of the base layer data randomly selected for the sample used in this study was found to be relevant to the case study.

It should also be noted here that it would be extremely difficult to find relevant offerings from those sources without knowledge of how each individual filename is coded. For example, datasets in the Canadian Digital Elevation Data include NTS grid coordinates in their title and are thus only searchable for those familiar with this system. No link is provided within the metadata to an index or legend for this system. Others, such as the 'Landsat 7 Orthorectified Imagery Over Canada' or the 'Ground Control Database' series, with thousands of datasets each,

are even more cryptically coded. For those unfamiliar with the codes, the only indication of the area an image covers is a polygon overlying a map. While these representations are clear on sight, they are not searchable.

4.13 Open data's usefulness for assessing community adaptive capacity

It was noted earlier in this chapter that the usefulness of open data is dependent on the needs of individual users and applications. For this study a general application for open data has been proposed by asking if open data could be used to assess a community's adaptive capacity. The data available on the three open data portals examined in this study will thus be evaluated for their usefulness when attempting to use data towards these ends.

The study employs a framework, developed by Wall and Marzall (2006), which suggests five categories of resources that contribute to adaptive capacity (economic, human, institutional, natural and social resources). These categories are discussed at length in this study's review of literature. It is likely that many of the datasets evaluated in this study could be used for a multitude of purposes, possibly combining or spanning multiple resource categories. As such, and in the interest of creating replicable results, operationalizing the concept of resources for adaptive capacity for quantitative analysis required clear parameters to distinguish datasets as being in one category or another.

This should be identified as a major limitation of this analysis. When discussing the potential of open data to assess adaptive capacity, the potential use of a dataset is just as or more important than the data's original use. For proponents of open data, both in the literature and

among those interviewed in this study, this flexibility of potential use is a key quality of open data and the source of its lauded benefits. But therein lies the problem: just as it would be difficult to predict all of the ways in which open data may be used, it is probably even more so for two researchers to make the same set of predictions.

In the face of this limitation, this study uses the data's original purpose to categorize each dataset into one of Wall and Marzall's (2006) five resource categories. This original purpose was first sought in the metadata of the dataset provided by the data stewards. If uncertainty remained, the general mission of the agency that originally collected or created the data was used to decide. A small number of datasets at the provincial level, geospatial projections, were categorized as 'Other'.

Table 2. Resource Category Frequencies for Open Data Relevant to the Region of Waterloo

	Regional (N=32)	Provincial (N=212)		Federal (n = 385, level of confidence = 95%)	
	% of total	% of total	Relevant data as % of total	% of total	Relevant data as % of total
Economic	12.5	17.9	14.1	1.8 (± 1.3)	1.6 (±1.3)
Human	0.0	9.9	9.9	1.0 (± 1.0)	0.8 (±0.9)
Institutional	75.0	32.1	27.8	0.0	0.0
Social	0.0	3.3	3.3	1.3 (± 1.1)	1.3 (± 1.1)
Natural	12.5	34.4	26.9	95.8 (± 2.0)	2.1 (±1.4)
Other	0.0	2.4	2.4	0.0	0.0
Total	100	100	84.5	100	5.7 (± 2.3)

Over seventy percent of the data offered on the Region's open data portal has been categorized as institutional. This is data related to the Regional government's assets, operations and programming, including schedules for waste pickup and council meetings, Region-owned infrastructure and administrative boundaries, and transit and cycling routes. Data related to human and social resources are notably absent at the Regional level.

At the provincial level, economic, institutional and natural resource data also dominate the catalogue, but the portal does contain data related to human and social resources. While these make up a smaller proportion of the catalogue, all of the datasets that fall within those categories have been judged to be geographically relevant to the Region of Waterloo. For the most part human resource related datasets are concerned with the operations, programming, and student populations of Ontario's schools, colleges and universities. Social resource data includes but is not limited to records of marriages, births and deaths.

At the federal level, the vast majority of data making up the catalogue falls within the natural resource category. As mentioned elsewhere, the vast majority of this is satellite imagery. While Natural Resources Canada had supplied 198,022 of the 206,279 available datasets on the federal portal at the time of data collection, Environment Canada and the Department of Fisheries and Oceans have also contributed data that likely fall into this category. Apart from data related to institutional resources, of which there is no data on the federal open data portal, limiting analysis to data relevant to the Region of Waterloo show the proportion of datasets for each resource category in a statistical tie.

As noted in Wall and Marzall (2006), the importance of each category of resource for adaptive capacity, and thus of each category of data and the individual datasets within them, is

largely determined by the idiosyncrasies of each community. As such, this study cannot attempt to assign any value to each category of resource without meaningful consultations with residents of the Region of Waterloo. Such a feat is beyond the scope and resources of this study. Instead, the conclusions of this study are limited to stating that open data is available for each category, while relying on findings from the evaluation of other attributes to infer whether that data meets the conditions for effective use.

4.14 Data content and Format - Granularity

The level of granularity in this study has been determined by the unit of observation used in a dataset. As such, a dataset broken down by municipality would be seen as having a finer level of granularity than those with data recorded at the provincial level. Given that this is a case study of the Region of Waterloo, data with no geographical coverage of the Region have been noted as 'not relevant', with observations of the granularity discarded.

Granular data was viewed by key informants and grey literature sources to be of greater use than aggregated data. A librarian interviewed for this study believed there would be great research benefits if rural-specific data were made available. However, concerns over privacy have often led to the aggregation of data in areas with low population density, making much of the data on rural areas less useful to work with. When presenting evidence to the federal Standing Committee on Government Operations and Estimates, a leader from the small business community suggested aggregated data is less relevant to small firms trying to understand their town or neighbourhood as a marketplace. An open data advocate presenting to the Standing Committee on Access to Information, Privacy and Ethics felt that data aggregation should occur

at geographies commonly recognized by communities, be they neighbourhoods, city wards or administrative districts. The lack of a standard unit for reporting data at finer levels of granularity was also identified in reviewed grey literature sources as making the comparison and combination of different datasets a difficult proposition (SCAIPE, 2011k; Government of Canada, 2014e).

Table 3. Relevance of Data on Provincial and Federal Open Data Portals to the Region of Waterloo

	Provincial (N=212) % of total catalogue	Federal (n = 385, CL=95%) % of total catalogue
Not relevant	15.6	94.3 (±2.3)
Relevant, not granular	37.3	5.2 (±2.2)
Relevant, granular	47.2	0.5 (±0.7)
Total	100.0	100.0 (±0.0)

Each of the 629 datasets evaluated in this study were classified as a) not relevant to the Region of Waterloo, b) relevant, but with a level of granularity coarser than upper-tier regional municipalities such as the Region of Waterloo, or c) relevant with granularity at the level of upper-tier regional municipalities or finer. All datasets from the Region of Waterloo's open data portal were found to have a level of granularity at the level of upper-tier regional municipality or finer.

4.15 Data Format

The primary concern of much of the grey literature reviewed in this study was that data be, at a minimum, machine-readable. (Tychon, 2014; Eaves, 2012; Government of Ontario, 2014b; Region of Waterloo, 2013a). Structured, machine readable data is conducive to exploration and analysis (SCAIPE, 2010b; 04, FED-ME-023, Region of Waterloo, 2013a; Geist, 2009). One key informant felt the issue of machine-readable data was particularly poignant with financial data, which they saw as often being made available publicly in PDF. Key informants also emphasized that raw data should always be made available over data that had been 'massaged or manipulated'. Raw data was essential for users to do their own analysis.

Key informants interviewed for the study were also concerned with specific file formats. Most recognized the value of non-proprietary formats, but felt common proprietary formats compatible with common software were no reason to hold back the release of data. One librarian interviewed for the study felt there was a need to explain to users which formats were compatible with what software. Another emphasized the need to ensure the ability to preserve data once a format became outdated. Other informants debated the value of static data dumps to data accessible via an application programming interface (API). While the latter were seen by some to be inaccessible to many users, it was suggested developers prefer them over having to push updates to their audiences when data used by their applications became outdated. This aspect of the discussion of formats again depended on the prospective user, with the ideal solution being to offer both data dumps and an API.

Experts presenting the federal Standing Committee on Access to Information, Privacy and Ethics also seemed to advocate for making data available in multiple formats (SCAIPE,

2010d; SCAIPE, 2010e), while the Open Data Charter commits member countries to releasing as much data in as many formats as possible (TBSC, 2014a). The federal Open Data Roundtable summary reported stakeholders were also mixed in their demand for an API, with developers and some researchers among those in favour. It was also suggested that applications regularly accessing an API would provide open data portals with better analytics on how data was being used (Government of Canada, 2014e).

None of the three portals examined in this study offers access to open data available through an API. The federal portal does have an API, but this access is limited to the metadata of datasets in the portal's catalogue. This study uses three initial categories to describe the format of datasets available on the three open data portal evaluated: structured data, geospatial data, and non-machine readable, fixed-layout documents such as images and PDFs. The first two categories can be described as machine readable, while the third is not. While PDFs may contain any combination of images, text or tables, they are not generally seen as structured data and thus are placed in the same category as images.

Table 4. Data formats of available open data

	Regional (N=32)	Provincial (N=212)		Federal (n=384, CL=95%)	
	% of Total	% of Total	Relevant % of Total	% of Total	Relevant % of Total
Tabular	34.4	48.6	42.9	6.0 (±2.4)	3.7 (±1.9)
Geospatial	53.1	50.5	40.6	67.5 (±4.7)	1.6 (±1.3)
Tabular + Geospatial	12.5	0.9	0.9	0.3 (±0.5)	0.0 (±0.0)
Image / PDF	0.0	0	0.0	25.7 (±4.4)	0.5 (±0.7)
Total	100	100	84.5	100 (±0.0)	5.8 (±2.3)

Geospatial data makes up approximately more than half of the regional catalogue and includes basemap features such as contours, building footprints and rivers, in addition to transit routes, library locations and administrative boundaries.

The provincial open data portal is nearly evenly split between tabular and geospatial data, with the overall proportion relevant to the Region of Waterloo falling along a similar ratio. The majority of geospatial datasets originate from the Ministry of Natural Resources, but also the Ministry of Environment and Ministry of Transportation. A very small proportion of the Federal catalogue is made up of tabular data, but it is in this category that most of the federal datasets relevant to the Region of Waterloo fall. Most of this data originates from Statistics Canada and Agriculture and Agri-food Canada.

All geospatial datasets in the sample taken from the Federal catalogue are supplied by Natural Resources Canada. Satellite imagery has been included in this category if it is in a format that allows for the inclusion of additional data, such as GeoTIFF. The federal portal is the only one to include PDF and imagery in its catalog as open data. Some of these records are satellite imagery with possible geo-spatial applications, but because they do not contain any additional, machine-readable data, they have not been classified as such. Strict accordance to the open definition discussed in this study's review of literature would not qualify these offerings as open data at all, although the five star open data scheme discussed in the next section still does as long as the files are covered by an open license.

4.16 Openness

As a general measure of the 'openness' of data made available through the three open data portals examined in this case study, the analysis uses Sir Tim Berners-Lee's Five Star Open Data deployment scheme. This scheme has been outlined in the review of literature and is operationalized for analysis in accordance with the five star open data schema (Hausenblas, 2012). The scheme has been adopted by many in the open data community of interest as a way to gauge the openness of the data offered as open data. The scheme has become so well accepted that the federal government uses the system to rate the offering on its own open data portal. These of course were not taken at face value, as each dataset was independently assessed by the researcher for this study. Where catalogue entries showed more than one available format available for a dataset (not including additional metadata) the evaluation used the most 'open' format for this ranking.

Table 5. Five Star open data ratings of Available open data

	Regional (N=32)	Provincial (N=212)		Federal (n=384, CL=95%)	
	% of Total	% of Total	Relevant % of Total	% of Total	Relevant % of Total
☆	0.0	0.0	0.0	28.1 (±4.5)	1.6 (±1.3)
☆☆	40.6	58.0	53.3	15.6 (±3.6)	0.3 (±0.5)
☆☆☆	59.4	42.0	31.1	56.4 (±5.0)	3.9 (±1.9)
Total	100	100	84.4	100	5.7 (±2.3)

None of the portals denote their datasets using uniform resource identifiers or link the available data to other data for context, the requirements of four and five star data ratings, respectively. Because all data on each portal is covered by an Open Government License, no dataset in this evaluation has less than one star. A two star rating requires data be machine-readable in addition to being covered by an open license, while three stars are awarded if the file format is non-proprietary.

Only the federal portal contains data with only one star, which means the data is not machine-readable. While the five star scheme still calls this open data, such data does qualify for a strict interpretaion of the open definition in which data must be machine-readable (OKF, 2013). This is in part because a large proportion of the portal is made up of imagery files. This evaluation has, however, differentiated between basic image files (jpeg, tiff) and those capable of recording metadata of use in geospatial software (GeoTiff).

These latter formats are also non-proprietary, contributing to the 56.4 percent (95% CI: 51.4% - 61.4%) of the federal catalogue that achieves a three star rating. The regional open data portal fares best in this regard out of the three examined, with 59.4 percent of its data in non-proprietary formats. Like the Region's portal, all of the provincial portal's datasets are machine-readable, although over half remain in proprietary formats. This proportion is mostly a mix of Microsoft Excel spreadsheets and geospatial Shapefiles, or .SHP format. This latter format is very popular and actually compatible with most geographic information system (GIS) programs. However, Esri, the company behind the ubiquitous ArcGIS software, still holds the rights to the format and as such it is treated as proprietary in this study.

It should also be noted that 75 of the 212 datasets catalogued on the provincial open data site are in fact hosted on the Ministry of Natural Resources Land Information Organization database, and that in order to access these datasets users must supply contact information in order to access them. The link to download data is provided in an email to the address input by the user. There does not appear to be any restriction to access based on the contact details submitted, but some may balk at the belaboured process. The requirement to supply contact details could be interpreted as counter the spirit of the Open Definition which states that open data must be provided without technological obstacles to the data's access, and that a license should not discriminate against persons, groups or fields of endeavour (OKFN, 2012) (eg. someone without an email address or uncomfortable with the requirement to supply one). Stated more explicitly in the provincial Information and Privacy Commissioner's 2011 annual report, non-discriminatory access means registration for access is not required (IPCO, 2011).

4.17 Longitudinal data

While some of the key informants interviewed for this study would prefer to see version histories and archives of all updates to the open data, none of the three data portals currently offer this. Instead, the analysis here data that could be deployed in time series analysis for the availability of the data at regular intervals, such as those indicated in the update frequency provided in the metadata.

Time-series data in this analysis is treated as any dataset in which observations or measurements are recorded at regular intervals, at the same unit of observation, or set of units. The open data portals may provide these as separate files, sometimes with separate records in the open data catalogue. For tabular data, they may also be presented as separate sheets or columns within the same file. If a geospatial dataset is one made up of recurring observational data (health of forests, water or air quality), it will be deemed as time series. If the file consists of administrative boundaries or static point data, it will not be – even if these files are subject to update with new data points added.

Table 6. Availability of times series open data

	Regional (N=32)	Provincial (N=212)		Federal (n=384)	
	% of Total	% of Total	Relevant % of Total	% of Total	Relevant % of Total
Not Time Series	75.0	56.1	46.3	95.1 (±2.2)	1.3 (±1.1)
Available	12.5	32.1	26.4	4.2 (±2.0)	3.6 (±1.9)
Unavailable	12.5	11.8	11.8	0.8 (±0.9)	0.8 (±0.9)
Total	100.0	100.0	84.5	100.0	5.7 (±2.3)

The majority of datasets on all three portals are not classified as time-series by this study's definition. Exactly half of time-series datasets available at the Region's portal and the majority of those available on the provincial and federal portals have data available at multiple intervals of time. These majorities remain true, although decreasing slightly, when only considering the time-series data relevant to the Region of Waterloo on the provincial and federal portals.

4.18 Metadata

As noted in the requirements for effective use of open data presented in this study's conceptual framework, the information made accessible to a prospective user about the data is key to their ability to reuse it.

Key informants and sources from the review of grey literature were all in agreement that high quality metadata is key to the effective use of open data (SCAIPE, 2011b; SCGOE, 2014b).

Metadata was viewed by key informants as being important to finding data, to accurately interpret the data (Government of Canada, 2014e), and to know where to ask questions about the data if needed. While recognizing the importance of metadata, one key informant warned that it should not be prescriptive about the interpretation of data in a sort of gatekeeper function. A key informant from the provincial open data portal agreed with this sentiment but, along with other key informants suggested metadata can ensure the accurate interpretation of data by including information about collection methodology and important technical considerations or imperfections in the data.

Key informants on the user side of the community of interest unanimously felt the metadata available at open data portals was lacking in both quantity and quality. However, the metadata desired varied among responses. Definitions of variables were the most often cited elements, along with contacts for data stewards (SCGOE, 2014b). Some key informants questioned the value of including too much metadata if users would not use it and that a basic minimum should be aimed for. While some felt the use of a standard could be used to ensure consistency in metadata, others felt existing standards would not be adhered to as they were too demanding. Findings from the federal Open Data Roundtable surveyed in this study's systematic review of grey literature also reported a range of metadata needs based on the stakeholder being consulted. Academics, policy wonks and librarians wanted specific metadata standards to be adhered to, while developers and others in the community wanted information accessible to the layperson (Government of Canada, 2014e).

Many key informants recognized the process of creating metadata as a labor intensive one. While a public servant working with the provincial open data portal suggested some

provincial ministries have yet to even establish metadata practices, key informants from the user side suggested internal use of metadata would lead to its improved quality. The idea that users of open data could contribute to crowd-sourced metadata came up both in interviews with key informants and in reviewed grey literature (Government of Canada, 2014e). Other suggestions included metadata being generated through automated programs, and sharing the responses to questions about the data previously asked by other users.

This study assessed open data for the availability of information about the data available through open data portals, and how easily a potential user could find more information if needed. The quantitative evaluation examined three facets: the level of description provided for the content of each dataset, the information provided on the process and methods of data collection or creation, and the availability of contact information for the original data providers if the user chooses to seek further information. The evaluation is limited to data available on the portal page, and does not look at metadata provided separately. In some cases, particularly with the federal portal, there is more metadata once you have downloaded the dataset. As such, findings from this evaluation are only applicable at the point of discovery.

4.19 Metadata: Level of description

While it has been left to respondents of key informant interviews to provide deeper commentary on the general quality of metadata, this quantitative assessment simply asked if a description of each dataset was provided, and if so, if it went to the extent of some form of data dictionary for terms used in the dataset's fields or attributes.

Table 7. Level of description for available open data

	Regional (N=32)	Provincial (N=212)		Federal (n=384, CL=95%)	
	% of Total	% of Total	Relevant % of Total	% of Total	Relevant % of Total
Not described	0.0	0.0	0.0	10.9 (±3.1)	0.5 (±0.7)
General	3.1	37.3	31.6	89.1 (±3.1)	5.3 (±2.2)
Field level	96.9	62.7	52.9	0.0	0.0
Total	100	100	84.5	100	5.8 (±2.3)

Of the three portals, the federal is the only one that omits descriptions of its datasets, albeit a minority of them. But while the vast majority of data in the catalogue and almost all of the data relevant to the Region of Waterloo is described in a general sense, no dataset from this portal is provided with an attribute-level description. The provincial and regional portals do a much better job at describing their data, with the Region's portal containing attribute level descriptions for all but one of its datasets.

4.20 Metadata: Data Collection / Creation Methods

Table 8. Description of collection / creation methods for available open data

	Regional (N=32)	Provincial (N=212)		Federal (n=384)	
	% of Total	% of Total	Relevant % of Total	% of Total	Relevant % of Total
Not described	90.6	53.3	43.4	51.7 (±5.0)	3.9 (±1.9)
Described	9.4	46.2	41.1	48.3 (±5.0)	1.9 (±1.4)
Total	100	100	84.5	100	5.8 (±2.3)

The availability of information on the collection and/or creation of data was assessed at an even more basic level, with the evaluation simply identifying whether such a process had been described at any level or was omitted. The federal and provincial catalogues are both nearly evenly split in regards to this question, although this is only true for approximately one-third of the data relevant to the Region of Waterloo found on the federal portal. Only a small proportion of datasets on the Region's portal include information on their collection or creation.

4.21 Metadata: Contacts for data stewards

Table 9. Supplied contacts for available open data

	Regional (N=32)	Provincial (N=212)		Federal (n=384)	
	% of Total	% of Total	Relevant % of Total	% of Total	Relevant % of Total
Agency	12.5	100.0	84.5	100.0	5.7 (±2.3)
Data Steward	87.5	0.0	0.0	0.0	0.0
Total	100.0	100.0	84.5	100.0	5.7 (±2.3)

The study's quantitative assessment also sought to characterize the ease for a prospective user to contact someone responsible for the creation or collection of the data. The results were stark, with all metadata on the federal and provincial portals only providing the agency from which the data originated. While the federal portal provides a link to that agency's website, the province does not even go that far. Almost all datasets on the Region's portal, however, contained the name and contact information of an individual data steward.

4.22 Metadata: Time Coverage / Date of collection

The date or time period of data collection is imperative to knowing the currency of the data one is working with. As such, metadata available on the portal for each dataset was evaluated for the provision of a date of collection or period of coverage. The findings do not attest to the accuracy of such information, but only whether or not the data stewards have provided it.

Table 10. Provision of time coverage / date of collection

	Regional (N=32)	Provincial (N=212)		Federal (n=384)	
	% of Total	% of Total	Relevant % of Total	% of Total	Relevant % of Total
Not supplied	43.8	0.0	0.0	0.0	0.0
Supplied	56.3	100.0	84.5	100.0	5.7 (±2.3)
Total	100.0	100.0	84.5	100.0	5.7 (±2.3)

All records examined in the sample of the federal portal provided dates/time periods for all datasets, as did the entire population of the provincial portal. The Region's portal, however, failed to provide dates for 43.8 percent of the data available on its portal.

4.23 Metadata: Update Frequency

As discussed in the review of literature, the sustainability of a project, application or product developed to use open data can often be, at least in part, dependent on the frequency at which that data is updated. The overall sustainability of open data initiatives and products built using the data they supply is viewed by this study as being reflected in continued availability of and frequency of updates to datasets. Multiple sources reviewed in the systematic review of grey literature reflected open data users' need to be reassured that any product they create using open data will be able to continue accessing that data in the future (SCAIPE, 2011a; SCGOE, 2014c; SCGOE, 2014e; Eaves, 2012). While in one sense this requires frequent updates to data, a

librarian interviewed for this study pointed out that it also means data from previous intervals should be preserved and left accessible for the long-term. Findings from the federal Open Data Roundtable suggest desire for longitudinal data is particularly strong in the non-profit, government, academic and policy sectors (Government of Canada, 2014e).

The recent provincial study on open data recommends that this could be achieved if government were becomes a major user of its own open data infrastructure (Government of Ontario, 2014b). Stakeholders contributing to the federal Open Data Roundtable consultation revealed frustration at not being kept up to date about updates to data (Government of Canada, 2014e). It is for this reason that some key informants interviewed for this study expressed a preference to access data through APIs rather than static data dumps.

This study assessed data available on the three open data portals under review for listing their update frequency within the metadata available on the dataset's portal web page. This is not an assessment of the accuracy of the update frequency stated in each dataset's metadata, but of the portal providing notice of the frequency of updates at all. Although all three portals make it clear that the terms of use can change, these stated frequencies may be enough commitment to give users the confidence to invest in using the data. As such, each dataset was assessed as providing the update frequency or not. If the frequency is listed as 'Never', this is still taken as a valid response. In certain cases a static, unchanging dataset is still useful, while in others this will be enough warning for users not to plan a project dependent on updates to the dataset.

In some cases, data stewards have declared the update frequency as 'As Needed'. These datasets have been judged to contain an ambiguous response. While understandable for an agency's geospatial datasets that may not change regularly enough to warrant regular updates,

the ambiguity around updates may be more frustrating than a 'Never' for users for any user wishing to ascertain the sustainability of any product they may develop based on the data. As such, it may be just as or more useful for users of the data to know how often the currency of the data is reviewed as it is the frequency at which it is updated.

Table 11. Update frequencies of available open data

	Regional (N=32)	Provincial (N=212)		Federal (n=384)	
	% of Total	% of Total	Relevant % of Total	% of Total	Relevant % of Total
Included	43.8	13.7	10.9	1.3 (±1.1)	1.3 (±1.1)
Not Included	43.8	44.3	39.1	5.2 (±2.2)	1.0 (±1.0)
Ambiguous	12.5	22.6	34.4	93.5 (±2.5)	3.4 (±1.8)
Total	100.0	100.0	84.4	100.0	5.7 (±2.3)

An overwhelming 93.5 percent (95% CI: ±2.5%) of the federal open data catalogue list ambiguous update frequencies. The province, on the other hand, fails to provide any update frequency for 44.3 percent of its datasets. Only approximate one-fifth of the data relevant to the Region of Waterloo available on both the provincial and federal portal include a definitive update frequency. Even the Region's portal, providing specific update frequencies for less than half of the datasets it offers, could be more reassuring to its users.

Summary

Awareness of open data was identified as an initial barrier to its success in small and rural communities, while local developers, journalists, community organizations, small business, researchers and public servants were most commonly indicated as prospective users. Most interviewed from the open data community of interest admitted that some people and communities would be in a better position to take advantage of open data than others.

Challenges for those already aware of and interested in using open data in small and rural communities, both on the user and supply sides of the community of interest fell into three general themes: infrastructure, resources and know-how. Small municipalities were seen as resource constrained and lacking the necessary technology and data management practices needed to implement an open data initiative. A lack of data literacy was identified as a challenge on both the supply and user side of open data. It was suggested this be addressed through educational programming, while free resources were also thought to be available for those inclined to learn independently. A more immediate solution was thought to be the infomediary, data literate knowledge brokers that could use open data on behalf of those interested but lacking the skills to do so.

An inventory of available open data found that much of the data was indeed relevant to the Region of Waterloo. While data was also available for each of the five categories of resource found in Wall and Marzall's (2006) framework for adaptive capacity, some categories fared proportionally better than others across the three levels of government data examined. The format of and metadata included with each dataset was also examined, with findings revealing a

number of areas open data providers could improve on in order to better meet the conditions necessary for effective use. These are discussed in greater detail in the next chapter.

Chapter 5 – Discussion

The purpose of this study was two-fold: to identify challenges to the effective use of open data in small and rural communities and to explore the potential use of open data to inform assessments of the adaptive capacity of rural communities. The previous chapter addressed the first of these tasks through informants from the open data community of interest relevant to the Region of Waterloo, with a systematic review of grey literature used test the validity of those responses.

Awareness of open data was identified as an initial barrier to its success in small and rural communities, while local developers, journalists, community organizations, small business, researchers and public servants were most commonly indicated as prospective users. Most interviewed from the open data community of interest admitted that some people and communities would be in a better position to take advantage of open data than others.

A number of barriers were identified for those already aware of and interested in using open data in small and rural communities, both on the user and supply sides of the community of interest. These challenges were organized into three general themes: infrastructure, resources and know-how.

5.1 Infrastructure

While rural broadband connectivity remains an issue in much of Canada, key informants did not see this as the case in the Region of Waterloo. Key informants also felt that most people in the

Region's rural areas would have access to a device that would allow them to benefit from open data, be it through direct access or via a product developed using open data such as a smartphone app. Admittedly, this is not surprising as the Region was selected to be a case study with the knowledge that it contained one of Canada's major centres of technology. It must be stated that the same is not necessarily true for other rural communities in the province and country (Government of Canada, 2014c).

Infrastructure was seen, however, to be an issue on the supply side of the open data community of interest. It was suggested that the cost of both the software and hardware needed to implement a successful open data initiative could be prohibitive for small municipalities. Key informants suggested the process should be thought of in the long-term, rather than making data open overnight, with the necessary upgrades integrated into existing plans for IT renewal. But even this may be a feat in the region's lower-tier municipalities that do not have dedicated internal IT staff.

Infrastructure issues identified on the supply side of the open data community of interest overlapped with those placed in another theme, that of resources. The purchase of new technology requires financial resources that small rural municipalities may find difficult to meet, but human resources were more often cited as a challenge by key informants. While the upper-tier Region of Waterloo was able to implement an open data program without adding to its staff, key informants familiar with the lower-tier municipalities were already at maximum capacity, with staff performing multiple roles.

Given the likelihood that other small municipal governments face similar issues, the open data model employed in the Region of Waterloo may be of interest to Ontario's other

regional and county two-tier municipal governments. At the time of writing, Niagara, Peel and York Regions are the only other upper-tier municipal governments in Ontario to offer open data. One obvious shortcoming of this model is that the regional portal only hosts data relevant to the whole of the region, and will not post a dataset unless it covers all lower-tier municipalities. As such, demands for data specific to a single municipality will be left unaddressed unless similar data is available for all other municipalities in the region.

5.2 Resources

While resources were indicated as a major barrier to the viability of open data initiatives run by small municipalities, key informants also indicated that, in the long run, many internal efficiencies could be gained from open data. It was thought resources may see savings from fewer information requests from the public, reduction of redundant data collection, more efficient data sharing between departments, more efficient operations and programming through improved analysis and data-driven decision making. It was also thought that third parties using open data to offer services to the public could relieve demand on the government.

Whether or not these new efficiencies would surpass the new demands on organizational resources resulting from running an open data initiative needs further exploration. It should be noted that the lower-tier municipalities in the Region of Waterloo already receive data from the upper-tier municipal and provincial governments. Furthermore, as one informant pointed out, data sharing between departments in small organizations may be as simple as speaking to someone who works in the same office and with whom a relationship already exists. Without the siloed organizational structure found in larger governments and the supposed data 'gatekeeping'

that can accompany this, it seems that some of the efficiencies to be gained from data sharing may be smaller than first perceived.

5.3 Know-how

In addition to the data management skills needed to maintain an open data initiative, key informants suggested public servants needed a level of data literacy sufficient to field public requests for data from both basic and advanced users; to increase the awareness of the value of the data being opened to the public; and, to use the data themselves to draft policy and design programs based on evidence-based decision making and business intelligence. This view was reiterated in the study's systematic review of grey literature, particularly in the final report of consultations conducted by the provincial government (Government of Ontario, 2014b).

On the user side of the community, the capacity for the general public to use the data was identified as the largest barrier to the use of open data. Key informants suspected that data literacy may be lower outside of urban centres, but this could not be tested or proved within the scope of this study. A lack of capacity to use the data, however, was not seen by key informants as a valid argument against releasing it. On the contrary, most thought the availability of open data could lead to greater data literacy, as this data would be relevant to the public and thus provide material and motivation to learn.

Most key informants saw a need for data literacy programming in school curricula, but many also believed that learning materials were available online if aspiring users were properly motivated. A number of informants felt the open data community of interest could also play a

role in increasing data literacy, through knowledge-sharing and by advocating for relevant programming from the government. This study's systematic review of grey literature shows some of the same agencies working on issues of digital literacy through community informatics programming have now begun to advocate for data literacy.

However, the use of knowledge, or 'infomediaries', was seen as the most immediate solution to issues of data literacy. Local developers, journalists, and community organizations were viewed by they study's key informants as having an important role to play in processing raw open data to produce applications or value-added information that would ultimately be consumed by the public. The same sentiment can be found in open data literature (Halonen, 2013; Davies, 2013), although some authors warn against an over-reliance on infomediaries. Janssen (2012) suggests those translating the data for wider consumption would be placed in a position of power that risks being misused to advance their own agendas.

One key informant suggested it was too early to tell if the developer community within the Region of Waterloo would provide infomediary services in a noncommercial way. Many suggested it was more realistic for special interest groups to directly contract the services of developers or data analysts to create value added products for them, an option made more affordable by the provision of free data. Traditional infomediaries like local news outlets were seen as well-placed to use open data for low-cost, in-depth journalism.

One key informant suggested that while local markets would be too small for larger companies to be interested in producing locally focused applications, there would be enough incentive for local developers to want to do so. Others felt the optimum scenario would be for the standardization of municipal data, allowing small communities to use applications developed

elsewhere. Key informants on the supply side of the open data community of interest reported standards had been discussed in municipal open data working groups, but no meaningful progress had been made to such ends.

5.4 Evaluating available data

The three themes discussed above have some obvious areas of overlap. A lack of financial resources can prevent an organization from investing in the infrastructure needed for an open data initiative, while a larger staff may allow for specialization in data management. On the user side, broadband connectivity can provide learning materials to work with data.

These three themes mostly fall within two of the three categories of elements presented in this study's adapted framework for effective use of open data (see Chapter 2 – Review of Literature). These are issues related to user capacity and access, and issues related to government policy priorities and financing. The third category in the adapted framework, issues related to data quality and format, is examined in greater depth in the study's analysis of data made available by the federal, provincial and regional open data portals.

While data content and quality were not included in the discussion above, this should not be taken as an indication that these are not issues in small and rural communities. Key informants had much to say about data quality, formats and metadata, but many pointed out that such issues were as applicable to the data available from the whole of the federal and provincial portals as they were to data from those portals relevant to the Region of Waterloo or the data on the region's own portal. As such, these issues were viewed as applicable to open data in general.

Nor should the elements for effective use presented in the study's adapted framework be seen as independent of one another. It is easy to infer how a lack of human resources for an open data initiative can lead to quality issues in the data it releases, or how the amount of metadata provided may influence the level of 'know-how' required for a potential user in order to make use of it.

5.5 Available data: Data relevant to the Region of Waterloo

The evaluation of data at three open data portals, federal, provincial and regional made it clear that data relevant to the region is available. While the ratio of relevant data as part of total data offered decreased at higher levels of government decreased, the gross amount of data available from larger governments meant they had more to offer in terms of the sheer number of datasets. For example, 100% of datasets available at the regional portal were deemed relevant, only 5.7% (CI 95%: 3.4-8%) of the data from the federal portal was classified as such. However, when characterizing this in terms of actual numbers of relevant datasets available, this means 32 relevant datasets are available from the regional portal compared to between roughly seven and sixteen thousand datasets from the federal portal (N=206,279). Meanwhile, 84.5% or 179 of 212 datasets from the provincial portal were found to be relevant.

Sheer volume, however, means very little with regard to the usefulness of the data, especially with regard to a specific use such as the one originally suggested by this study, assessing adaptive capacity. Halonen (2013) points out that the overall number of data users may even be irrelevant, with the more important indicators found among those actually making

use of open data with relatively high impact. For Halonen (2013), the accessibility and usability of data is as or more important than the actual amount of data available.

5.6 Available data: Datasets by category of resource determining adaptive capacity

This study makes use of Wall and Marzall's framework for adaptive capacity, which posits adaptive capacity as access to five categories of resource: economic, social, institutional, human and natural. With assessments of adaptive capacity acting as a hypothetical use case in this thesis, the quantitative analysis sought to organize available, geographically relevant data into each of these five categories. Table 2 presents the datasets from within the five categories found on each portal.

The regional portal is dominated by institutional data, which in this study includes data on the upper-tier's assets and infrastructure. The province has the largest proportion of human resource data relevant to the region. When excluding data not found to cover the region, the small proportion of relevant data on the federal portal was fairly evenly spread across four of the five categories of resource. The exception was a complete lack of data related to institutional resources.

The relevant data available from each level of government may reflect the responsibilities that government has to the community. For example, education is primarily the jurisdiction of the province and thus results in a sizable proportion of the data available on the provincial open data portal dealing with schools, enrolment and educational programming. As noted in the review of literature, the determinants of adaptive capacity occur at a variety of

scales (Smit & Wandel, 2006). Some are local while others reflect larger scale socio-economic and political systems. In Wall and Marzall's (2006) framework, the importance of each category of resource determining adaptive capacity will vary by community. However, one can foresee potential issues arising if data essential for assessing adaptive capacity falls within the jurisdiction of an under-resourced municipal government.

It is also possible that the priorities of governments are reflected in what data is released or not released. While interviews with representatives of open data initiatives have shown agencies begin by releasing the data which requires the least amount of effort, some have expressed concerns that these agencies will not go beyond providing the 'low-hanging fruit' (Dawes, 2012). While providing evidence to the federal Standing Committee on Access to Information, Privacy and Ethics, an expert working on the American federal open data initiative noted a tendency for 'mission-driven disclosure', in which agencies provide data that advances their own mission and goals (SCAIPE, 2011b). Many have been quick to point out the Canadian government's repeatedly stated commitment to open data and open government while imposing strict controls on federal science programs.

Again, the number of datasets in each category indicates little other than the potential of finding data needed for the community to assess its adaptive capacity. As noted in this study's literature review (Brooks & Adger, 2004; Wall & Marzall, 2006), a process of stakeholder engagement is needed within communities to determine the importance of each resource category and related indicators. Without the required process of stakeholder engagement, far beyond the capacity of this study, little can be said with regard to the appropriateness of available data for such an assessment.

It should also be noted that this study's findings are only a snapshot of what was available at the time of data collection. These catalogues have grown since data was collected for this study, and will continue to expand as the open data movement matures. These findings are thus useful as a guide for recommending areas in which the open data portals may want to concentrate their efforts as they continue to make more data available.

The open data managers interviewed for this study indicated that they were regularly making more data available. However, their priority was to meet user demand rather than attempting to encourage specific uses. While all portals surveyed in this study have some mechanism for demanding data, the province had taken the extra step of putting a list of all releasable data to a public vote. This approach addresses an issue raised by many of the informants interviewed for this study, who said they felt most people would not know what data the government has and what they may request.

As such, any conclusions on the ability to use open data to assess adaptive capacity should not be limited to what is currently available. Furthermore, it should be reiterated here that this study's methodology sought the dataset's original use when classifying them by resource category. Open data can be reused for purposes different from its original use, and could thus possibly contribute to assessing access to resources in another of the five categories identified by Wall and Marzall (2006) than the one indicated in this study.

Finally, if 'data by default' becomes a reality, one would expect to be spoiled for choice in each category of resource. However, one outstanding question is whether or not government has a complete collection of the data needed to assess adaptive capacity. Saying nothing of the potential for original research, one might expect to find valuable secondary data among crown

corporations that do not fall under the government open data mandate, non-profits and charities, various community, trade and service organizations, and academic research that profile the community or their online social networks.

The third category of this study's adapted framework for the effective use of open data also refers to how data is preprocessed and presented for discovery and reuse. This includes formatting, tagging and metadata, as well as matters of presentation such as choice of interval in time series data, or units of observation. All of these attributes will affect user ability to interpret data, although their impact will vary between users and their intended use (Davies, 2013; Gurstein, 2011). Findings related to these issues are discussed below into two loosely defined sections: format and metadata.

5.7 Format

Purely geospatial data makes up half of the region's open data catalogue. The provincial open data portal is close to evenly split between tabular and geospatial data, with the proportion of data relevant to the Region of Waterloo falling along a similar ratio. A very small proportion of the federal catalogue is made up of tabular data, but it is in this category that most of the datasets deemed relevant to the Region of Waterloo fall.

When discussing the format of open data, there is nothing to say that geospatial or tabular data would be more useful when assessing adaptive capacity. As with so many other factors, this depends on which indicators the community find appropriate for their local assessment. It is worth noting that maps produced from geospatial data can be useful for

community engagement, although some tabular data can be mapped or visualized in charts and graphs. Such a consideration would sit well with Juhola and Kruse (2013), who argue the translation of the results of adaptive capacity assessments requires careful consideration of how that information is communicated to the target audience, especially those stakeholders who will influence what actions are taken.

A more important consideration when discussing data format may be the availability of non-proprietary formats, although this too ultimately depends on the user. Well-funded organization hired to consult on adaptive capacity assessments are more likely to be able to access the software needed to use proprietary formats than grassroots community groups or even academics. Using Berners-Lee's five star scale of open data ("Openness Rating", 2013), the study showed the majority of data available at the Region of Waterloo's portal and the majority of data relevant to the region available at the federal portal to be in non-proprietary formats. Over half of the data relevant to the region available through the provincial portal was in proprietary formats. The federal open data portal was the only one of the three to include data that was not machine-readable, one of the most basic requirements of open data.

While standards were not assessed quantitatively, interviews with key informants suggested little data apart from transit schedules followed any standards beyond those implemented internally in the collection agency. A number of key informants suggested standard data formats would allow for open data applications and tools to be used with data across jurisdictions. It's unclear how much value this would have for assessments of adaptive capacity, but standardization would allow for comparison between communities.

From the data deemed relevant to the Region of Waterloo, this study differentiates between data which is granular at the level of upper-tier municipalities or finer and that with a more coarse level of granularity. Well over half of provincial data found to be relevant to the Region of Waterloo was categorized as granular at a level finer than the region itself. While a miniscule proportion of the federal portal was found to have such a fine level of granularity, with over 200,000 records in the catalogue this still represents hundreds of available datasets.

This level of granularity was selected because it is at this level that plans for climate change adaptation are currently being considered within the region (Region of Waterloo, 2013). Admittedly, lower-tier townships conducting their own assessment of adaptive capacity may want data with a finer level of granularity to provide measurements for some of their selected indicators.

That said, it should be remembered that the adaptive capacity of a community is affected by factors at multiple levels. Smit and Wandel (2006) view the determinants of adaptive capacity as reflecting the interactions of conditions set by socio-economic and political systems at a number of scales. As such, data that uses upper-tier municipalities or provinces as a unit of observation is still relevant to a lower-tier municipality, especially if the indicator measured applies equally to the whole of that unit's contents (eg. Geospatial data showing regional administrative boundaries for provincial agencies).

More than two thirds of time series data on the provincial portal has data available for previous increments, while only half of the time-series data available on the regional portal does. A large majority of the time-series data available on the federal portal has time-series data available. Engle (2011) argues anticipatory adaptation measures taken to lessen the impacts of

future events depend on a community's ability to learn from success and failure in past experiences, as well as being having some idea of what the future holds – something that may be helped with the ability to look for long-term trends using up to date data.

5.8 Metadata

Dawes (2012) argues understanding the processes of data collection, management, access and dissemination is necessary in order to fully understand data and it potential use. The literature suggests poor documentation, a lack of commonly agreed standards and ambiguous data definitions commonly leave the metadata provided with open data often insufficient for its effective reuse (Dawes, 2012; Peled 2013; Zuiderwijk et al., 2012).

Five attributes were evaluated in the metadata provided at the three portals examined in this study: the level of description of the data, information on the method employed to collect or create the data, contacts for data stewards, period of coverage or collection date and update frequency. This is a selection of attributes based on considerations raised in this study's review of literature. The requirements of most metadata standards are much more demanding. The study's key informant interviews indicated a general dissatisfaction with the quality of metadata provided at open data portals, but also that the average did not require an overwhelming amount of information either.

Almost all of the data on the regional portal is described at the field level, while the same is true for just over half of the data on the provincial portal. None of the datasets sampled from the federal open data portal had field-level descriptions. Creation/collection methods were

described for less than 10% of the data offered on the regional portal, and almost half of all data relevant to the region on the provincial portal. The majority of relevant data on the federal portal did not include a description of collection or creation methods. Obviously, the failure to provide this information increases the possibility for the misinterpretation of data and, in the context of assessing adaptive capacity, an uncertainty of whether the data truly measures selected indicators.

Neither the provincial or federal portals provided any contact information for the data stewards responsible for datasets, posting only the agency that originally produced it. Contacts for data stewards were supplied for the vast majority of datasets at the regional open data portal. At the region, this may allow for more efficient consultation with data stewards on questions specific to the data itself, about its suitability to assess a particular aspect of adaptive capacity, and what other relevant data may be available. At the very least, one would expect questions sent to general contact addresses for agencies supplying data to have a much slower response time. Given only the name of the agency, as is the case on the provinical and federal portals, users may feel inclined to contact the portal managers first and thus place greater strain on their resources. Open data literature (Peled, 2013; Helbig et al.,2012) notes the potential for the public to collaborate with open data stewards to improve data quality, but this may prove difficult without contacts for those public servants responsible for supplying the data.

Davies (2013) argues open data users are much more likely to invest time and resources in projects using data when they know it will continue to be available and can expect frequent updates and timely access. The date of collection or period of time covered by the data was supplied for all offerings at the provincial portal and within the sample from the federal

catalogue, but just over half of the data supplied from the regional portal included this information. Meanwhile, more than half of the relevant data at the provincial portal and of all data at the regional open data portal were either ambiguous about or did not include an indication of the frequency of updates to datasets. This is true for an even larger proportion of relevant data from the federal open data portal. While knowing the date of data collection or time period data covers is essential for meaningful data analysis, recent and frequently updated data is key to assessing current and future resource access and adaptive capacity.

The findings above show each portal as having a number of strengths and weaknesses in its metadata contributing to the overall potential for data in those catalogues to be used effectively. Ultimately the most important attribute may be the contact information needed for users to ask questions and fill the gaps in the metadata included with a dataset. Given that such contact information is a major shortcoming for both provincial and federal open data portals, users of data from those catalogues must hope they data they need is well-documented and/or simple enough to understand without thorough metadata, or be willing to undertake the investigation needed to find the appropriate data stewards.

5.9 For further research

The scope of this study's quantitative assessment was too limited to be able to test for a number of other data attributes that contribute to the effective use of open data. These include, but are not limited to data accuracy and discoverability.

While data accuracy is somewhat tied to the frequency of updates for a dataset, there are many more variables specific to each dataset that need to be considered. Unfortunately, it is impossible to examine each one within the scope of this study. It should be noted that both key informants and sources in the systematic review of grey literature recognized that data released to the public would not be perfect. However, many saw access to imperfect data as better than no data at all, as long as those imperfections were noted in the provided metadata. Some, including the public servants providing data, saw an opportunity for open data users to collaborate with the government to improve the quality of data. Others suggested the accuracy of data would improve if open data were used internally by government. Given the evident optimism around the opportunity provided by open data to improve the overall quality of government data, a study of whether or not this occurs as the open data movement matures could prove to be valuable research.

The quality of metadata contributes to open the discoverability of data, but other factors related to the structure and capabilities of the open data portals also play a role. Sources reviewed in the systematic review of grey literature suggested the ability to use categorical filters, to sort by popularity, to preview data, to share data through social networks, to comment on or tag datasets would all increase their discoverability. A librarian interviewed for this study suggested effectively addressing the discoverability of data required an examination of user needs. The informant cited libraries as examples when exploring how different interfaces for discovery of the same set of information can be used to address the needs of advanced or basic users. These issues are complex unto themselves and merit a more thorough examination in another study.

It was noted in the discussion above that government may not have all of the data needed for the assessment of adaptive capacity, as is likely true for many objectives. Further study is needed on the possibility of combining data from other sources with that available from government open data portals. Furthermore, if open data is to be used in any community-led process, more study is required on how a balance can be struck between the 'expert knowledge' that results from the analysis of that data and local knowledge offered by the community.

Finally, over the course of this study a new question has emerged in the mind of the researcher. If the hailed benefits of open data are realized, will they contribute to any of the five resource categories of adaptive capacity? The benefits are said to include economic development through better-informed business decisions and the provision of data-related applications and services (Janssen et al., 2012; Davies, 2012). Open data is supposed to lead to greater transparency and accountability, developing institutional resources (Janssen et al., 2012). Open data could also contribute to better delivery of services provided by community organizations and to greater civic participation by a well-informed public (Noveck, 2010; Evans-Cowley & Hollander, 2010). It was even suggested in this study's findings that open data will lead to greater data literacy, thus contributing to the development of human resources. Each of these possibilities must be examined on their own, in depth and over an adequate period of time.

Chapter 6 – Final Summary, Conclusion, Further Research and Recommendations

6.1 Final Summary

This study opened with a review of literature in two major parts. First, the concept of adaptive capacity was reviewed with an emphasis on the concept's role in the vulnerability framework. The five categories of resource argued by Wall and Marzall (2006) to be determinants of adaptive capacity were then explored in greater detail. The second half of the review of literature examined open data, with the origins and perceived benefits of open data discussed prior to moving into a review of the conditions identified in the literature as necessary for open data to be used effectively.

Chapter 3 described this study's mixed methods approach, including key informant interviews, a systematic review of grey literature, and an inventory of available open data relevant to the Region of Waterloo, evaluated for its potential to assess local adaptive capacity and properties supporting effective use. Chapter 4 reported the findings of the research conducted using those methods. Results from key informant interviews and the systematic review of grey literature related to challenges to open data use in small and rural communities are presented within three themes: technology and infrastructure, resources, and know-how. Findings from the evaluation of available data were presented categorically, with tables reporting frequencies. Chapter 5 connected the three themes of challenges findings identified in the study's qualitative analysis with those of the quantitative evaluation of available open data, and discussed each within the context of the literature reviewed in Chapter 2.

6.2 Conclusions

The goal of this research was to explore the use of open data to contribute to assessments of adaptive capacity in rural Ontario communities. Wall and Marzall (2006) create a framework for adaptive capacity based on access to five categories of resource: economic, institutional, human, social and natural resources. This study sought to explore whether open data could be a valuable source of information for communities seeking to assess their access to locally-available resources in those five categories.

The study recognized that the composition and importance of each of these categories in a community would be dependent on the local context (Brooks & Adger, 2004; Wall & Marzall, 2006), and was thus limited in what it could conclude with regard to the content of available open data beyond evaluating how much data was available within each category at open data portals relevant to the Region of Waterloo. The study does, however, examine issues around the ability to use open data within the context of small and rural communities that may preclude its application to assess adaptive capacity. Special consideration is given to the general needs for indicators used to assess adaptive capacity laid out by authors employing the vulnerability framework, as noted in the review of literature.

It can also be concluded that key informant interviews identified a lack of awareness about open data in small and rural communities, as well as a number of issues that stood in the way of open data's goals in those places. These findings were, for the most part, corroborated by those from a systematic review of open data grey literature. On the supply side of the community of interest, a lack of IT infrastructure needed to implement an open data initiative and of financial and human resources needed to maintain such a program were seen as

particularly difficult issues to overcome in small municipalities. A general lack of data literacy was identified among prospective users of open data, but it was suggested that this could eventually be addressed by educational programming in schools and community organizations. Comprehensive learning materials were said to be available online for those communities without access to such resources, but more immediately through the use of knowledge brokers, coined infomediaries, conducting research and developing applications using open data on behalf of the community.

Data related to each of the five resource categories in Wall and Marzall's (2006) framework was available, but it cannot be stated with certainty if any of this data would contribute to meaningful indicators of the Region of Waterloo's adaptive capacity. Data related to natural and economic resources figured strongly at each of the three portals, while data related to institutional resources was also made up a sizeable proportion of the regional and provincial portals. The proportion of open data catalogues related to human resources was insignificant except at the provincial level, while social data was largely absent from all but the federal portal. The reasons behind these gaps need to be explored with further research.

Finally it is concluded, that the data available from the federal, provincial and regional open data portals each had strengths and weaknesses when evaluated through the lens of the study's adapted framework for effective use. While all three portals had some work to convert data to non-proprietary formats, the federal portal needs to ensure all of its data is in fact machine-readable. The quality of metadata needs to improve at all three portals in different respects. Time-series data need to go back further in time, while users need to know when they can expects updates. Without more elaborate descriptions of datasets or the ability to contact

data stewards, the risk of misinterpreting some data remains increases.

6.3 Recommendations

In addition to its currently mitigation-focused strategies to address climate change (Region of Waterloo, 2013), this study recommends the Region of Waterloo should consider applying the vulnerability framework to its climate change adaptation plans by identifying the area's levels of exposure, sensitivity and adaptive capacity. Stakeholder engagement should be undertaken to identify the resources needed to obtain a suitable level of adaptive capacity and decide on appropriate indicators. Once this process is completed, data for those indicators can be sought internally and through contacts with the provincial and federal governments. Any data found to be relevant should be made open if it is not already, so that the public may be able to comment on and contribute to any government assessment of adaptive capacity.

In order to support the ability to find data related to resources for adaptive capacity, or for any other use, open data initiatives would do well to follow the provincial government in releasing an inventory of all non-sensitive datasets so that users may know what data governments have and make the appropriate request for its release. The release of data under the jurisdictions of arms-length associations, boards and crown corporations should be encouraged by governments, while the wider open data community should reach out to community organizations and charities to do the same.

Organizations should take a long-term view when implementing open data. Information technology and human resource requirements should be integrated into the organization's long-

term vision in order to accommodate their costs and to ensure the initiative's sustainability. Where smaller municipalities within two-tier systems cannot afford the resources needed to open their data, upper-tier municipalities should explore how they provide assistance, perhaps by hosting and maintaining the data on their own portals. The Region of Waterloo's policy of not opening data unless it covers the entirety of the region should be revisited, with study made as to how the program can still provide equal levels of service to each of the lower-tier municipalities.

Improved metadata should be made a priority as government agencies integrate the eventual release of data into their workflow. Time-series data should be released as far back as possible, and all data should be released at a level of granularity that allow for meaningful analysis at the local level. The federal open data portal should reconsider how it organizes geospatial data in its catalogue. New information technology systems should support non-proprietary formats while legacy data in proprietary formats should be converted where possible.

In addition to efforts to increase awareness of open data and its potential benefits, advocates within the open data community of interest should also advocate for and contribute to the development of data literacy programming. Furthermore, mechanisms should also be developed to aid connections between infomediaries and those needing their services. An incentive program for infomediaries to provide services to organizations with minimal resources could be explored as a way to promote open data and benefit the public.

References

Adger, N. W., Arnell, N. W., & Tompkins, E. L. (2005). Successful adaptation to climate change across scales. *Global environmental change*, *15*(2), 77-86

Ahmed, A.K. (2006). Concepts and Practices of "Resilience": A Compilation From Various Secondary Sources. Thailand: United States Agency for International Development.

Bates, J. (2012). “This is what modern deregulation looks like”: co-optation and contestation in the shaping of the UK’s Open Government Data Initiative. *The Journal of Community Informatics*, *8*(2).

Bizer, C., Heath, T., & Berners-Lee, T. (2009). Linked data-the story so far. *International journal on semantic web and information systems*, *5*(3), 1-22.

Brooks, N. and Adger, N.W. (2004). “Assessing and Enhancing Adaptive Capacity” in *Adaptation Policy Frameworks for Climate Change: Developing Strategies, Policies and Measures* Eds. Bo Lim, B., Spanger-Siegfried, E. Cambridge University Press.

Bryant, C. R., Smit, B., Brklacich, M., Johnston, T. R., Smithers, J., Chjotti, Q., & Singh, B. (2000). Adaptation in Canadian agriculture to climatic variability and change. *Climatic Change*, *45*(1), 181-201.

C-CIARN. (2002, March 1). *Agricultural Adaptation in a Changing Climate.* Workshop conducted at the Cutten Club, Guelph, ON.

Clement, A. H., Gurstein, M., & Longford, G. (Eds.). (2012). *Connecting Canadians: investigations in community informatics*. Athabasca University Press.

Climate Action WR. (2013). *A Climate Action Plan for Waterloo Region: Living Smarter in 2020*. Waterloo, ON.

Cohen, S. J. (2011). Overview: Climate Change Adaptation in Rural and Resource-Dependent Communities. In *Climate Change Adaptation in Developed Nations* (pp. 401-412). Springer Netherlands.

Davidson, D. J., Williamson, T., & Parkins, J. R. (2003). Understanding climate change risk and vulnerability in northern forest-based communities. *Canadian Journal of Forest Research*, *33*(11), 2252-2261.

Davies, A., & Lithwick, D. (2010). Government 2.0 and Access to Information: 1. Recent Developments in Proactive Disclosure and Open Data in Canada. Parliamentary

Information and Research Service; Canadian Electronic Library.

Davies, T. (2010). *Open data, democracy and public sector reform: A look at open government data use from data. gov.uk*. Practical Participation.

Davies, T. (2012). Supporting open data use through active engagement. *W3C Using Open Data Workshop, June 2012*. Brussels.

Davies, T. (2013) *Open Data Barometer: 2013 Global Report*. London: Open Data Institute and World Wide Web Foundation.

Davies, T., Perini, F., & Alonso, J.M. (2013). *Researching the emerging impacts of open data: ODDC conceptual framework*. London: World Wide Web Foundation and ODDC.

Dawes, S. S. (2010). Stewardship and usefulness: Policy principles for information-based transparency. *Government Information Quarterly*, *27*(4), 377-383.

Dawes, S. S. (2012). A realistic look at open data. *Center for Technology in Government, University at Albany/SUNY Available at http://www. w3. org/2012/06/pmod/pmod2012_submission_38. pdf*.

Desouza, K. C., & Bhagwatwar, A. (2012). Citizen Apps to Solve Complex Urban Problems. *Journal of Urban Technology*, *19*(3), 107-136.

Dickinson, T., & Burton, I. (2011). Adaptation to climate change in Canada: A multi-level mosaic. In *Climate Change Adaptation in Developed Nations* (pp. 103-117). Springer Netherlands.

Expert Panel on Climate Change Adaptation (EPPCA) Report to The Minister of The Environment. (2009). *Adapting to Climate Change in Ontario: Towards the Design and Implementation of a Strategy and Action Plan*. Toronto: Queen's Printer for Ontario.

Engle, N. L. (2011). Adaptive capacity and its assessment. *Global Environmental Change*, *21*(2), 647-656.

Evans-Cowley, J., & Hollander, J. (2010). The new generation of public participation: Internet-based participation tools. *Planning, Practice & Research*, *25*(3), 397-408.

Eversole, R. (2011). Community agency and community engagement: re-theorising participation in governance. *Journal of Public Policy*, *31*(01), 51-71.

Eversole, R. (2012). Remaking participation: challenges for community development practice. *Community development journal*, *47*(1), 29-41.

Fischer, G. (2001, August). Communities of interest: Learning through the interaction of multiple knowledge systems. In *Proceedings of the 24th IRIS Conference* (pp. 1-14). Department of Information Science, Bergen.

Foley, P., Alfonso, X., & Wiseman, I. (2007). *Local Information Systems: A Review of Their Role, Characteristics and Benefits; Summary*. Department for Communities and Local Government.

Folke, C., Hahn, T., Olsson, P., & Norberg, J. (2005). Adaptive governance of social-ecological systems. *Annual Review Environmental Resources*, *30*, 441-473.

Frazier, T. G., Thompson, C. M., & Dezzani, R. J. (2014). A framework for the development of the SERV model: A Spatially Explicit Resilience-Vulnerability model. *Applied Geography*, *51*, 158-172.

Government of Canada. 2012. Canada's action plan on open government. Canadian Electronic Library (Firm) Ottawa, Ont

Gupta, J., Termeer, K., Klostermann, J., Meijerink, S., van den Brink, M., Jong, P., Nooteboomf S., & Bergsma, E. (2010). The adaptive capacity wheel: a method to assess the inherent characteristics of institutions to enable the adaptive capacity of society. *Environmental Science & Policy*, *13*(6), 459-471.

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

Halonen, A. (2012). *Being Open About Data: Analysis of the UK open data policies and applicability of open data*. London: The Finnish Institute in London.

Hausenblas, M. (2012). 5 ★ Open Data [website]. Retrieved from http://5stardata.info/

Hazell, R. (1989). Freedom of Information in Australia, Canada and New Zealand. *Public Administration*, *67*(2), 189-210.

Helbig, N., Cresswell, A. M., Burke, G. B., & Luna-Reyes, L. (2012). *The Dynamics of Opening Government Data: A White Paper*. Albany: Center for Technology in Government.

Hogge, B. (2010). *Open data study: a report commissioned by the Transparency and Accountability Initiative*. Retrieved from http://www.opensocietyfoundations.org/reports/open-data-study

Holsten, A., & Kropp, J. P. (2012). An integrated and transferable climate change vulnerability assessment for regional application. *Natural hazards*, *64*(3), 1977-1999.

Huijboom, N., & Van den Broek, T. (2011). Open data: an international comparison of strategies. *European journal of ePractice*, *12*(1), 1-13.

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

Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, adoption barriers and myths of open data and open government. *Information Systems Management*, *29*(4), 258-268.

Jetzek, T., Avital, M., & Bjørn-Andersen, N. (2012). The Value of Open Government Data: A Strategic Analysis Framework. In *2012 Pre-ICIS Workshop*.

Juhola, S., & Kruse, S. (2013). A framework for analysing regional adaptive capacity assessments: challenges for methodology and policy making. *Mitigation and Adaptation Strategies for Global Change*, 1-22.

Khan, K. S., Kunz, R., Kleijnen, J., & Antes, G. (2003). Five steps to conducting a systematic review. *Journal of the Royal Society of Medicine*, *96*(3), 118-121.

Keele, S. (2007). *Guidelines for performing systematic literature reviews in software engineering* (pp. 1-57). Technical report, EBSE Technical Report EBSE-2007-01.

Kulig, J. C., Edge, D. S., & Joyce, B. (2008). Understanding Community Resiliency in Rural Communities Through Multimethod Research. *Journal of Rural and Community Development*, *3*(3), 77-94.

Lakomaa, E., & Kallberg, J. (2013). Open Data as a Foundation for Innovation-The Enabling Effect of Free Public Sector Information for Entrepreneurs. *Access, IEE.* (1), 558-563.

Lemmen, D.S., Warren, F.J., Lacroix, J., and Bush, E. (Eds.). (2008). *From Impacts to Adaptation: Canada in a Changing Climate 2007.* Ottawa, ON: Government of Canada.

MacLean, D., & St. Arnaud, B. (2008). In Don MacLean B. S. A. (Ed.), *ICTs, innovation and the challenge of climate change* Winnipeg, MB: International Institute for Sustainable Development, 2008.

Maier-Rabler, U., & Huber, S. (2011). " Open": the changing relation between citizens, public administration, and political authority. *eJournal of eDemocracy & Open Government*, *3*(2).

McCarthy, J. J. (Ed.). (2001). *Climate change 2001: impacts, adaptation, and vulnerability: contribution of Working Group II to the third assessment report of the Intergovernmental Panel on Climate Change*. Cambridge University Press.

McCarthy, J. J. (Ed.). (2001). *Climate change 2001: impacts, adaptation, and vulnerability: contribution of Working Group II to the third assessment report of the Intergovernmental Panel on Climate Change*. Cambridge University Press.

McLeman, R. A., Brklacich, M., Woodrow, M., Vodden, K., Gallaugher, P., & Sander-Regier, R. (2011). Opportunities and Barriers for Adaptation and Local Adaptation Planning in Canadian Rural and Resource-Based Communities. In *Climate Change Adaptation in Developed Nations* (pp. 449-459). Springer Netherlands.

Meier, K., Brudney, J., & Bohte, J. (2011). *Applied statistics for public and nonprofit administration*. Cengage Learning.

Meijer, A. J. (2007). Publishing public performance results on the Internet: Do stakeholders use the Internet to hold Dutch public service organizations to account?. *Government Information Quarterly*, *24*(1), 165-185.

Meijer, A. (2009). Understanding modern transparency. *International Review of Administrative Sciences*, *75*(2), 255-269.

Meijer, A. J., Curtin, D., & Hillebrandt, M. (2012). Open government: connecting vision and voice. *International Review of Administrative Sciences*, *78*(1), 10-29.

Noveck, B. (2010). The Single Point of Failure. in Eds. Lathrop, D., & Ruma, L.. *Open government: Collaboration, transparency, and participation in practice* (pp. 51-72). O'Reilly Media, Inc.

NEU O'Neill, S., & Boykoff, M. (2011). The role of new media in engaging the public with climate change. In *Engaging the Public with Climate Change: Behaviour Change and Communication* (pp. 233-251). Routledge.

Obama, B. (2009). Transparency and open government. *Memorandum for the heads of executive departments and agencies*.

Open Knowledge Foundation (OKFN). (2012). *Open Data Handbook Documentation, Release 1.0.0.* Retrieved from: http://opendatahandbook.org/pdf/OpenDataHandbook.pdf

Pant, L. P. & Heeks, R. (2012). ICT-Enabled Development of Capacity for Climate Change Adaptation, pp. 116-132. In R. Heeks & A. V. Ospina (Eds.). ICTs, Climate Change and Development: Themes and Strategic Actions. University of Manchester, UK, and International Development Research Centre (IDRC) of Canada.

Papaioannou, D., Sutton, A., Carroll, C., Booth, A., & Wong, R. (2010). Literature searching for social science systematic reviews: consideration of a range of search techniques. *Health*

Information & Libraries Journal, *27*(2), 114-122.

NEU Pasquier, M., & Villeneuve, J. P. (2007). Organizational barriers to transparency a typology and analysis of organizational behaviour tending to prevent or restrict access to information. *International Review of Administrative Sciences*, *73*(1), 147-162.

Peled, A. (2011). When transparency and collaboration collide: The USA open data program. *Journal of the American society for information science and technology*, *62*(11), 2085-2094.

Peled, A. (2013). Re-Designing Open Data 2.0. *eJournal of eDemocracy & Open Government*, *5*(2).

Peristeras, V., Mentzas, G., Tarabanis, K. A., & Abecker, A. (2009). Transforming E-government and E-participation through IT. *Intelligent Systems, IEEE*, *24*(5), 14-19.

Piotrowski, S. J., & Van Ryzin, G. G. (2007). Citizen attitudes toward transparency in local government. *The American Review of Public Administration*, *37*(3), 306-323.

Reilly, J. M. (2011). Overview: Climate change adaptation in the agricultural sector. In *Climate Change Adaptation in Developed Nations* (pp. 347-357). Springer Netherlands.

Robinson, D., Yu, H., Zeller, W. P., & Felten, E. W. (2008). Government data and the invisible hand. *Yale JL & Tech.*, *11*, 159.

Sawicki, D. S., & Craig, W. J. (1996). The democratization of data: Bridging the gap for community groups. *Journal of the American Planning Association*, *62*(4), 512-523.

Sayer, A. (2000). *Realism and social science*. Sage.

Scassa, T. (2013, June 18). *Canada's Open Government Licence V2.0 Is Released* Retrieved from: http://www.teresascassa.ca/index.php?option=com_k2&view=item&id=131:canada%E2%80%99s-open-government-licence-v20-is-released&Itemid=81

Schneiderbauer, S., Pedoth, L., Zhang, D., & Zebisch, M. (2013). Assessing adaptive capacity within regional climate change vulnerability studies—an Alpine example. *Natural hazards*, *67*(3), 1059-1073.

Shakespeare, S. (2013). Shakespeare Review: An indedpendent review of public sector information. Department for Business, Innovation and Skills. UK

Simpson, L. E. (2005). Community informatics and sustainability: why social capital matters. *The Journal of Community Informatics*, *1*(2).

Smit, B., & Pilifosova, O. (2003). Adaptation to climate change in the context of sustainable development and equity. *Sustainable Development*, *8*(9), 878-912.

Smit, B., & Skinner, M. W. (2002). Adaptation options in agriculture to climate change: a typology. *Mitigation and adaptation strategies for global change*, *7*(1), 85-114.

Smit, B., & Wandel, J. (2006). Adaptation, adaptive capacity and vulnerability. *Global environmental change*, *16*(3), 282-292.

Treasury Board of Canada (2013). *Openness Rating.* Retrieved from: http://data.gc.ca/eng/openness-rating

Ubaldi, B. (2013). *Open government data: Towards empirical analysis of open government data initiatives* (No. 22). OECD Publishing.

Uhlir, P. F., & Schröder, P. (2007). Open data for global science. *Data Science Journal*, *6*, OD36-OD53.

Urwin, K., & Jordan, A. (2008). Does public policy support or undermine climate change adaptation? Exploring policy interplay across different scales of governance. *Global environmental change*, *18*(1), 180-191.

NEU Van Aalst, M. K., Cannon, T., & Burton, I. (2008). Community level adaptation to climate change: the potential role of participatory community risk assessment. *Global Environmental Change*, *18*(1), 165-179.

Wall, E., & Marzall, K. (2006). Adaptive capacity for climate change in Canadian rural communities. *Local Environment*, *11*(4), 373-397.

Wenger, E. (2000). Communities of practice and social learning systems. *Organization*, *7*(2), 225-246.

Wimsatt, W. C. (1994). The ontology of complex systems: levels of organization, perspectives, and causal thickets. *Canadian Journal of Philosophy*, *24*(sup1), 207-274.

Wolf, J. (2011). Climate change adaptation as a social process. In *Climate change adaptation in developed nations* (pp. 21-32). Springer Netherlands.

Worthy, B. (2014). Making Transparency Stick: The Complex Dynamics of Open Data. Birbeck College, London.

Yohe, G., & Tol, R. S. (2002). Indicators for social and economic coping capacity—moving toward a working definition of adaptive capacity. *Global Environmental Change*, *12*(1), 25-40.

Yu, H., & Robinson, D. (2012). The New Ambiguity of "Open Government". *UCLA Law Review Discourse, 59,* 178-230.

Zuiderwijk, A., Janssen, M., Choenni, S., Meijer, R., & Sheikh Alibaks, R. (2012). Socio-technical impediments of open data. *Electronic Journal of e-Government*, *10*(2), 156-172.

Appendix 1 – Sampling

Because of the large number of datasets available from the federal open data portal (N=193,105, in June, 2014), this study opted to evaluate a sample of the catalogue rather than the entire population, as it did for the provincial and regional open data portals.

A value of 0.5 was used to estimate of the population's unknown standard deviation, while a 95% confidence level was selected following common practice in the social sciences (Meier, Brudney and Bohte, 2009). 5% was accepted as a tolerable amount of error, with findings related to the federal open data portal limited to getting a general sense of the catalogue rather than any precise measurement of its attributes. Thus the sample was calculated as:

$$n = (z \times \sigma / E)^2$$
$$n = (1.96 \times 0.5 / 0.05)^2$$
$$n=384.16$$
$$n=384$$

Stratified sampling was employed primarily to ensure the sample included a proportional number of datasets from agencies other than Natural Resources Canada, which makes up more than 95% of the catalogue. The researcher felt this was important in order to reflect the diversity of data types and management practices used by different agencies, and because the records listed as the ten most downloaded datasets by the portal's own reporting all originate from agencies other than Natural Resources Canada.

The federal open data portal allows users to filter datasets by the mutually exclusive category of federal agency, supplying a count for the datasets that fall within each. These categories were used to stratify the sample using proportionate allocation.

Agency	**Datasets**	**Proportion**	**Sample Strata**
Natural Resources Canada:	185,068	95.99%	368.02 (368)
Statistics Canada:	5,583	2.86%	11.10 (11)
Agriculture and Agri-food Canada	1,623	0.84%	3.23 (3)
Other	741	0.39%	1.47 (1)

It was unclear in which order datasets would be displayed when listed by agency. As such, systematic sampling was employed within each strata in order to take a cross-section of each list. The collection from each agency was by the number of datasets to be sampled, yielding the interval at which datasets would be selected. An online random number generator was used to generate a whole number between one and the interval. The random number was used as the initial point for sampling in the catalogue. In the case of the combined categories of 'Other', all

datasets from the remaining agencies were listed together as they appeared on the site, with the random number between one and 741 used to select a dataset.

Appendix 2 – Criteria Used in the Evaluation of Available Open Data

Five resource categories
0 - Other
1 - Economic
2 - Human
3 - Institutional
4 - Natural
5 - Social

It is possible that datasets could be used for other purposes. However, this survey will use the original purpose of the data to determine whether or not the data relates to one of five resource categories. Description was examined first. If still unclear the source of the data and the mission of that organization were used to decide.

Dataset Description
0 - No description provided.
1 - A general description of the dataset is provided.
2 - The data is described at the field level (eg. a definition is provided for each column of tabular data or for each feature class in geospatial data).

Description of Collection/Creation method
0 - Not Described
1 - Described

Contact provided for Data Stewards
0 - No source/contact provided.
1 - The agency providing data is named.
2 - An individual data steward is named.

Description of Update frequency
0 - Described
1 - Not Described
2 - Ambiguous

In some cases 'As Needed' is used. While potentially frustrating for those looking to use the data in a sustainable manner, this makes sense for an agency's geospatial datasets that may not

change regularly enough to warrant regular updates. In such cases, it would be better for users of the data to know how often the currency of the data is reviewed.

Description of Time coverage / Date of collection
0 - Date of collection or range of time coverage is not described
1 - Date of collection or range of time coverage is described

Time series
0 - It is not a time-series dataset
1 - It is a time-series dataset and previous increments of the data are not available.
2 - It is a time-series dataset and previous increments of the data are available.

Time-series data, in this analysis, is any data with which a comparative analysis of the same unit of observation at different points in time is possible. As such, a contact list is not seen as time-series data, nor is a a geospatial dataset that would be updated with new coordinates in the same file.

Format
0 - The data is tabular
1 - The data is geospatial Geospatial
2 - The data is offered in both geospatial and tabular formats (eg. list with addresses and a shape file)
3 - The data is in an image or PDF file

Five Star Openness Rating
0 – Not open
1 – The dataset is covered by an open licence – JPEG, PDF, TIFF
2 – The data is machine-readable – SHP, XLS(X), CAP, SAR/CTT,
3 – The data is machine-readable and in a non-proprietary format – KML, GML, GeoTIFF, GeoPDF, EDI, CDED ASCII (DEM), DIMAP
4 – The data is machine-readable, in a non-proprietary format and uses RDF standards to facilitate data interchange on the Web.
5 – The data is machine-readable, in a non-proprietary format and is already linked to other data using RDF standards.

See http://5stardata.info/ for more detail.

Geographical Relevance and Granularity

0 = not relevant

1 = relevant at a level of granularity higher than the region

2 = relevant with a level of granularity specific to the region or one of its component municipalities.

If a dataset covers the whole of Ontario or the whole of Canada, it has been deemed relevant. If a dataset is specific to one or multiple areas, it has only deemed relevant if one of those areas falls within the Region of Waterloo (eg. A dataset that uses lower-tier municipalities as a unit of observation). The level of granularity in this study has been determined based on the unit of observation used in a dataset. As such, a dataset broken down by municipality would be seen as having a finer level of granularity than those with data recorded at the provincial level. Given that this is a case study of the Region of Waterloo, data with no geographical coverage of the Region have been noted as 'not relevant', with observations of the granularity discarded.

Shapes and outlines such as provincial administrative boundaries that encompass the Region of Waterloo have been deemed geographically relevant, but not seen as granular to the level of the region. Point data, if recorded within the region, has been deemed as both granular and relevant to the region. It is arguable, however, that even with a lack of features in that area, it is still relevant to know - for example if a file tracking forest disease in the province shows that there is no record of forest disease in the Region of Waterloo.

Appendix 3 – Qualitative Coding Framework

Benefits of Open Data
-------- adaptive capacity
-------- data as new natural resource
-------- economic
-------- innovation
-------- measures of impact
-------- participation
-------- prestige
-------- savings
-------- services
-------- transparency + accountability

Community of Interest
-------- awareness
-------- collaboration
-------- consultation on OD & OG
-------- demand for data
-------- hackathons
-------- prospective users
-------- -------- developers
-------- -------- researchers and librarians
-------- relationships

Releasing Data
-------- institutional change
-------- -------- culture change
-------- -------- leadership
-------- -------- ownership
-------- -------- policy + priorities
-------- -------- sustainability
-------- risk
-------- -------- privacy
-------- -------- risk averse
-------- -------- security

Working with Data
-------- accessibility
-------- -------- inclusive use
-------- -------- language
-------- capacity
-------- -------- capacity (infrastructure)
-------- -------- capacity (know-how)
-------- -------- capacity (resources)
-------- data accuracy
-------- data format
-------- data management
-------- data usefulness
-------- discovery
-------- granularity
-------- indirect use
-------- -------- infomediary
-------- -------- tradiitional sources
-------- licensing
-------- metadata
-------- standards
-------- timeliness

Sources
-------- federal
-------- local to RW
-------- municipal
-------- provincial
-------- regional

Appendix 4 – Results of Systematic Review of Literature Search

Canadian Business and Current Affairs Database: 15 relevant results
Press release: 12
Opinion Piece / Blog: 3
http://search.proquest.com.subzero.lib.uoguelph.ca/cbcacomplete/

Google (limited to Canadian results): 13 relevant results
Opinion Piece / Blog: 6
Report: 7
https://www.google.ca/webhp?cr=countryCA

Google Custom Search of Government Domains: 9 relevant results
Informational Website:2
Report: 5
Presentation: 2
https://www.google.com/cse/home?cx=007843865286850066037%3A3ajwn2jlweq

Ontario Council of University Libraries Scholars Portal: 9 relevant results
Report: 9
http://journals2.scholarsportal.info/

Ontario Legislative Assembly (House Hansard Search): 3 relevant results
Parliamentary/Legislature Transcript: 3
http://hansardindex.ontla.on.ca/

Ontario government website: 4 relevant results
Informational website: 2
Press Release: 2
http://ontario.ca

Parliament of Canada: 27 relevant results
Parliamentary/Legislature/Council Meeting Notes/Transcripts: 24
Report: 3
http://www.parl.gc.ca

Treasury Board of Canada Secretariat: 29 relevant results
Press release: 21
Report: 1
Speech: 7
http://www.tbs-sct.gc.ca

Region of Waterloo: 2 relevant results
Parliamentary/Legislature/Council Meeting Notes/Transcripts: 2
http://www.regionofwaterloo.ca

Appendix 5 – Systematic Review of Literature References

27 Shift. (2013). *Beyond Giving and Volunteering: How and Why Individuals Are Exploring New Ways to Advance Social Good*. Vancouver, BC: Isakson, T. Retrieved from: http://www.27shift.com/wp-content/uploads/Beyond-giving-and-volunteering-full-report.pdf

Baker, M.B. (2011). Canada's Growing Open Data Movement. *Feliciter, 57(3), 96-98*. Retrieved from: http://www.cla.ca/AM/Template.cfm?Section=Vol_57_No_3

Benzie, R. (2013, October 21). Wynne seeks new era of 'openness'. *Toronto Star.* Retrieved from: http://torontostar.newspaperdirect.com/epaper/viewer.aspx

Clement, T. (2012a). *Speech by Tony Clement, President of the Treasury Board of Canada, to Regina and District Chamber of Commerce.* Speech presented at Regina and District Chamber of Commerce, Regina, SK. Retrieved from: http://www.tbs-sct.gc.ca/media/ps-dp/2012/0411-eng.asp

Clement, T. (2012b). *Speech by Tony Clement, President of the Treasury Board of Canada, at the Ignition Awards.* Speech presented at 2012 Ignition Awards, Ottawa, ON. Retrieved from: http://www.tbs-sct.gc.ca/media/ps-dp/2012/1019-eng.asp

Clement, T. (2012c). *Speaking Notes for the Honourable Tony Clement, President of the Treasury Board of Canada - "Using Technology to Challenge the Status Quo in Government Operations" – GTEC.* Speech presented at GTEC government technology conference, Ottawa, ON. Retrieved from: http://www.tbs-sct.gc.ca/media/ps-dp/2012/1106a-eng.asp

Clement, T. (2012d). *Speaking Notes for the Honourable Tony Clement, President of the Treasury Board of Canada, at PSENGAGE.* Speech presented at PSEngage conference, Ottawa, ON. Retrieved from: http://www.tbs-sct.gc.ca/media/ps-dp/2012/1127-eng.asp

Clement, T. (2013). *Speaking Notes for the Honourable Tony Clement, President of the Treasury Board of Canada, GTEC.* Speech presented at GTEC Government Technology Conference, Ottawa, ON. Retrieved from: http://www.tbs-sct.gc.ca/media/ps-dp/2013/1008-eng.asp

Canada. Parliament. House of Commons. Standing Committee on Access to Information, Privacy and Ethics (SCAIPE). (2010a). *Minutes of Proceedings.* 40th Parliament, 3rd session, meeting no. 031. Retrieved from: http://www.parl.gc.ca/HousePublications/Publication.aspx?DocId=4783697&Language=E&Mode=2&Parl=40&Ses=3

Canada. Parliament. House of Commons. Standing Committee on Access to Information, Privacy and Ethics (SCAIPE). (2010b). *Minutes of Proceedings.* 40th Parliament, 3rd session, meeting no. 037. Retrieved from: http://www.parl.gc.ca/HousePublications/Publication.aspx?DocId=4859333&Language=E&Mode=1

Canada. Parliament. House of Commons. Standing Committee on Access to Information, Privacy and Ethics (SCAIPE). (2010c). *Minutes of Proceedings.* 40th Parliament, 3rd session, meeting no. 038. Retrieved from: http://www.parl.gc.ca/HousePublications/Publication.aspx?DocId=4875736&Language=E

Canada. Parliament. House of Commons. Standing Committee on Access to Information, Privacy and Ethics (SCAIPE). (2010d). *Minutes of Proceedings.* 40th Parliament, 3rd session, meeting no. 040. Retrieved from: http://www.parl.gc.ca/HousePublications/Publication.aspx?DocId=4910281&Language=E

Canada. Parliament. House of Commons. Standing Committee on Access to Information, Privacy and Ethics (SCAIPE). (2010e). *Minutes of Proceedings.* 40th Parliament, 3rd session, meeting no. 042. Retrieved from: http://www.parl.gc.ca/HousePublications/Publication.aspx?DocId=4934696&Language=E

Canada. Parliament. House of Commons. Standing Committee on Access to Information, Privacy and Ethics (SCAIPE). (2011a). *Minutes of Proceedings.* 40th Parliament, 3rd session, meeting no. 043. Retrieved from: http://www.parl.gc.ca/HousePublications/Publication.aspx?DocId=4944847&Language=E&Mode=1

Canada. Parliament. House of Commons. Standing Committee on Access to Information, Privacy and Ethics (SCAIPE). (2011b). *Minutes of Proceedings.* 40th Parliament, 3rd session, meeting no. 045. Retrieved from: http://www.parl.gc.ca/HousePublications/Publication.aspx?DocId=4969132&Language=E&Mode=1

Canada. Parliament. House of Commons. Standing Committee on Access to Information, Privacy and Ethics (SCAIPE). (2011c). *Minutes of Proceedings.* 40th Parliament, 3rd session, meeting no. 046. Retrieved from: http://www.parl.gc.ca/HousePublications/Publication.aspx?DocId=4991691&Language=E&Mode=1

Canada. Parliament. House of Commons. Standing Committee on Access to Information, Privacy and Ethics (SCAIPE). (2011d). *Minutes of Proceedings.* 40th Parliament, 3rd session, meeting no. 047. Retrieved from: http://www.parl.gc.ca/HousePublications/Publication.aspx?DocId=5001510&Language=E&Mode=1

Canada. Parliament. House of Commons. Standing Committee on Access to Information, Privacy and Ethics (SCAIPE). (2011e). *Minutes of Proceedings. 4*0th Parliament, 3rd session, meeting no. 048. Retrieved from: http://www.parl.gc.ca/HousePublications/Publication.aspx?DocId=5015178&Language=E&Mode=1

Canada. Parliament. House of Commons. Standing Committee on Access to Information, Privacy and Ethics (SCAIPE). (2011f). *Minutes of Proceedings.* 40th Parliament, 3rd session, meeting no. 049. Retrieved from: http://www.parl.gc.ca/HousePublications/Publication.aspx?DocId=5036242&Language=E&Mode=1

Canada. Parliament. House of Commons. Standing Committee on Access to Information, Privacy and Ethics (SCAIPE). (2011g). *Minutes of Proceedings.* 40th Parliament, 3rd session, meeting no. 050. Retrieved from: http://www.parl.gc.ca/HousePublications/Publication.aspx?DocId=5057741&Language=E&Mode=1

Canada. Parliament. House of Commons. Standing Committee on Access to Information, Privacy and Ethics (SCAIPE). (2011h). *Minutes of Proceedings.* 40th Parliament, 3rd session, meeting no. 052. Retrieved from: http://www.parl.gc.ca/HousePublications/Publication.aspx?DocId=5067532&Language=E&Mode=1

Canada. Parliament. House of Commons. Standing Committee on Access to Information, Privacy and Ethics (SCAIPE). (2011i). *Minutes of Proceedings.* 41st Parliament, 1st session, meeting no. 008. Retrieved from: http://www.parl.gc.ca/HousePublications/Publication.aspx?DocId=5156870&Language=E&Mode=1

Canada. Parliament. House of Commons. Standing Committee on Access to Information, Privacy and Ethics (SCAIPE). (2011j). *Minutes of Proceedings.* 40th Parliament, 3rd session, meeting no. 041. Retrieved from: http://www.parl.gc.ca/HousePublications/Publication.aspx?Language=E&DocId=4923340

Canada. Parliament. House of Commons. Standing Committee on Access to Information, Privacy and Ethics (SCAIPE). (2011k). *Minutes of Proceedings.* 40th Parliament, 3rd session, meeting no. 044. Retrieved from: http://www.parl.gc.ca/HousePublications/Publication.aspx?Language=E&DocId=4959359

Canada. Parliament. House of Commons. Standing Committee on Government Operations and Estimates (SCGOE). (2012). *Minutes of Proceedings.* 41st Parliament, 1st session, meeting no. 045. Retrieved from: http://www.parl.gc.ca/HousePublications/Publication.aspx?DocId=5582949&Language=E&Mode=1

Canada. Parliament. House of Commons. Standing Committee on Government Operations and Estimates (SCGOE). (2013). *Minutes of Proceedings.* 41st Parliament, 2nd session, meeting no. 005. Retrieved from: http://www.parl.gc.ca/HousePublications/Publication.aspx?DocId=6321775&Language=E&Mode=1

Canada. Parliament. House of Commons. Standing Committee on Government Operations and Estimates (SCGOE). (2014a). *Minutes of Proceedings.* 41st Parliament, 2nd session, meeting no. 013. Retrieved from: http://www.parl.gc.ca/HousePublications/Publication.aspx?DocId=6426510&Language=E&Mode=1

Canada. Parliament. House of Commons. Standing Committee on Government Operations and Estimates (SCGOE). (2014b). *Minutes of Proceedings.* 41st Parliament, 2nd session, meeting no. 013. Retrieved from: http://www.parl.gc.ca/HousePublications/Publication.aspx?DocId=6458624&Language=E&Mode=1

Canada. Parliament. House of Commons. Standing Committee on Government Operations and Estimates (SCGOE). (2014c). *Minutes of Proceedings.* 41st Parliament, 2nd session, meeting no. 018. Retrieved from: http://www.parl.gc.ca/HousePublications/Publication.aspx?DocId=6513254&Language=E&Mode=1

Canada. Parliament. House of Commons. Standing Committee on Government Operations and Estimates (SCGOE). (2014d). *Minutes of Proceedings.* 41st Parliament, 2nd session, meeting no. 019. Retrieved from: http://www.parl.gc.ca/HousePublications/Publication.aspx?DocId=6524084&Language=E&Mode=1

Canada. Parliament. House of Commons. Standing Committee on Government Operations and Estimates (SCGOE). (2014e). *Minutes of Proceedings.* 41st Parliament, 2nd session, meeting no. 020. Retrieved from: http://www.parl.gc.ca/HousePublications/Publication.aspx?DocId=6534497&Language=E&Mode=1

Canadian Geomatics Community Roundtable. (2013). *Data Resources Dimension Community Discussion Document.* Retrieved from: http://cgcrt.ca/wp-content/uploads/2013/12/Data-Sources-V-3.0.pdf

Canadian Internet Policy and Public Interest Clinic (CIPPIC). (2011). *Analysis of Share-Alike Obligations in Municipal Open Data Licenses*. Ottawa, ON: Fewer, D. & Mewhort, K. Retrieved from: https://cippic.ca/en/publications/analysis_of_share_alike_obligations

Canadian Internet Policy and Public Interest Clinic. (2012). *Creative Commons Licenses: Options for Canadian Open Data Providers*. Ottawa, ON: Mewhort, K. Retrieved from: https://cippic.ca/sites/default/files/Creative%20Commons%20Licenses%20-%20Options%20for%20Canadian%20Open%20Data%20Providers.pdf

Day, S. (2011). *Speaking Points for the President of the Treasury Board and Minister for the Asia-Pacific Gateway for the Open Government Announcement.* Statement to the media, delivered in Edmonton, AB. Retrieved from: http://www.tbs-sct.gc.ca/media/ps-dp/2011/0318-eng.asp

Day, S. (2011). *Speaking Points for the President of the Treasury Board and Minister for the Asia–Pacific Gateway for the Open Data Announcement*. Statement to the media, delivered in Vancouver, BC. Retrieved from: http://www.tbs-sct.gc.ca/media/ps-dp/2011/0317-eng.asp

Davies, A., & Lithwick, D. (2010). *Government 2.0 and Access to Information: 1. Recent Developments in Proactive Disclosure and Open Data in Canada*. Retrieved from the Library of Parliament Research Publications website: http://www.parl.gc.ca/Content/LOP/ResearchPublications/2010-14-e.htm

Douglas, K., Hurtubise-Loranger, E., & Lithwick, D. (2012). *The Access to Information Act and Proposals for Reform Report* (Publication No.2005-55-E). Legal and Legislative Affairs Division, Parliamentary Information and Research Service. Retrieved from: http://www.parl.gc.ca/Content/LOP/ResearchPublications/2005-55-e.pdf

Eaves, D. (2012). *My Canadian Open Government Consultation Submission.* Eaves.ca [blog]. Retrieved from: http://eaves.ca/2012/01/19/my-canadian-open-government-consultation-submission/

Engineers Without Borders (EWB). (2013). *2013 EWB Policy Forum on Global Development: Lessons Learned & Paths Forward*. Toronto, ON. Retrieved from: http://my2.ewb.ca/site_media/static/library/files/878/ewb-policy-forum-2013-lessons-learned-paths-forward.pdf

Federation for the Humanities and Social Sciences. (2014). Highlights Federal Budget 2014. Ottawa, ON. Retrieved from:
http://www.ideas-idees.ca/sites/default/files/budget_2014_-_highlights_e.pdf

FIMS UGC Research Team. (2010). *Mobilizing User-generated Content for Canada's Digital Advantage.* Faculty of Information and Media Studies at the University of Western Ontario: Trosow, S.E. Retrieved from: http://ir.lib.uwo.ca/cgi/viewcontent.cgi?article=1022&context=fimspub

Geist, M. (2008, August 3). 'Crowdsourcing' puts many extra hands to work. *Toronto Star.* Retrieved from: http://www.thestar.com/business/2009/08/03/crowdsourcing_puts_many_extra_hands_to_work.html

Geist, M. (2009, December 14). Canada dragging its feet on open data initiatives. *Toronto Star.* Retrieved from: http://www.thestar.com/business/2009/12/14/geist_canada_dragging_its_feet_on_open_data_initiatives.html

Government of Canada. (2014a). *Canada's Action Plan on Open Government*. Retrieved from: http://data.gc.ca/eng/canadas-action-plan-open-government

Government of Canada. (2014b). *Digital Canada 150 Final Report.* Ottawa, ON. Retrieved from: http://www.ic.gc.ca/eic/site/028.nsf/eng/00576.html

Government of Canada. (2014c). *Digital Canada 150 – Consensus Submission*. Ottawa, ON. Retrieved from: https://www.ic.gc.ca/eic/site/028.nsf/eng/00527.html

Government of Canada. (2014d). *Digital Literacy in Canada: From Inclusion to Transformation - Submissions*. Retrieved from: https://www.ic.gc.ca/eic/site/028.nsf/eng/00537.html

Government of Canada. (2014e). *Open Data Roundtables Summary Report*. Retrieved from: http://data.gc.ca/eng/open-data-roundtables-summary-report

Government of Canada. (2014f). *Open Data Institute*. Retrieved from: http://actionplan.gc.ca/en/initiative/open-data-institute

Government of Ontario. (2013a). *Backgrounder: Ontario's Open Government Initiative*. Retrieved from: http://news.ontario.ca/mgs/en/2013/10/ontarios-open-government-initiative.html

Government of Ontario. (2013b). *Backgrounder: Opening Ontario's Data*. Retrieved from: http://news.ontario.ca/mgs/en/2014/04/opening-ontarios-data.html

Government of Ontario. (2013c). News Release: Ontario Launches Open Government [press release]. Retrieved from: http://news.ontario.ca/opo/en/2013/10/ontario-launches-open-government.html

Government of Ontario. (2014a). News Release: Engaging the Public on Open Data [press release]. Retrieved from: http://news.ontario.ca/mgs/en/2014/04/engaging-the-public-on-open-data.html

Government of Ontario. (2014b). *Open by Default: A New Way Forward for Ontario*. Retrieved from: https://www.ontario.ca/government/open-default-new-way-forward-ontario

Gurstein, M. (2010, September 2). Open Data: Empowering the Empowered or Effective Data Use for Everyone? *Gurstein's Community Informatics* [Blog]. Retrieved from: http://gurstein.wordpress.com/2010/09/02/open-data-empowering-the-empowered-or-effective-data-use-for-everyone/

Gurstein, M. (2011, July 11). A Data Divide? Data "Haves" and "Have Nots" and Open (Government) Data. Gurstein's Community Informatics Blog [blog]. Retrieved from: http://gurstein.wordpress.com/2011/07/11/a-data-divide-data-%E2%80%9Chaves%E2%80%9D-and-%E2%80%9Chave-nots%E2%80%9D-and-open-government-data/

Information and Privacy Commissioner, Ontario (IPCO). (2011). *IPC Ontario 2011 Annual Report*. Retrieved from: http://www.ipc.on.ca/english/Resources/Annual-Reports/Annual-Reports-Summary/?id=1193

Information and Privacy Commissioner, Ontario (IPCO). (2012). *Open Data, Big Data, Yes ... Personal Data, No!* [presentation slides]. *Retrieved from:* http://www.ipc.on.ca/english/Resources/Presentations-and-Speeches/Presentations-and-Speeches-Summary/?id=1219

Information Commissioner of Canada (ICC). (2010). *Address by the Information Commissioner of Canada Standing Committee on Access to Information, Privacy and Ethics* [Presentation slides]. Retrieved from: http://www.oic-ci.gc.ca/eng/pa-ap-appearance-comparution-2010_6.aspx

KPMG. (2013). *Public Service Transformed: Harnessing the Power of Behavioural Insights*. Retrieved from: http://www.kpmg.com/Ca/en/IssuesAndInsights/ArticlesPublications/Documents/Public-Service-Transformed-FINAL2.pdf

Lithwick, D. & Thibodeau, M.O. (2012). *Canada and the International Open Government Partnership Report*. Retrieved from the Library of Parliament Research Publications website: http://www.parl.gc.ca/Content/LOP/ResearchPublications/2012-15-e.htm

Newspapers Canada. (2011). *National Freedom of Information Audit 2011*. Retrieved from: http://www.newspaperscanada.ca/public-affairs/FOI2011

Mowat Centre. (2011). Fiscal Sustainability and the Future of Public Services: A Shifting Gears Report. Toronto, ON: Gold, J., Hjartarson, J. & Mendelsohn, M. Retrieved from: http://mowatcentre.ca/fiscal-sustainability-and-the-future-of-public-spending/

Ontario Legislative Assembly. (2013). *Hansard Issue L078*. Retrieved from: http://hansardindex.ontla.on.ca/hansardeissue/40-2/l078.htm

Ontario Legislative Assembly. (2014). *Hansard Issue L126*. Retrieved from: http://hansardindex.ontla.on.ca/hansardeissue/40-2/l126.htm

Ontario Legislative Assembly. (2014). *Hansard Issue L129*. Retrieved from: http://hansardindex.ontla.on.ca/hansardeissue/40-2/l129.htm

Ontario Ministry of Rural Affairs (OMRA). (2014). *Rural Roadmap: The Path Forward for Ontario*. Retrieved from: http://www.omafra.gov.on.ca/english/rural/roadmap-2014.pdf

Policy Horizons Canada. (2012). *Driving Policy on a Shifting Terrain: Understanding the Changing Policy Environment Amid 21st-Century Complexity*. Ottawa, ON. Retrieved from: http://www.horizons.gc.ca/eng/content/driving-policy-shifting-terrain-21st-century-complexity

Public Safety Canada. (2012). Reducing the Impact of Disasters and Promoting Resilience [press release]. Retrieved from: http://www.publicsafety.gc.ca/cnt/nws/nws-rlss/2012/20121024-eng.aspx

Region of Waterloo. (2013a). *Regional Municipality of Waterloo Administration and Finance Committee Agenda, September 27, 2011*. Retrieved from: http://www.regionofwaterloo.ca/en/regionalGovernment/resources/FA2011-0927.pdf

Region of Waterloo. (2013b). *Regional Municipality of Waterloo Administration and Finance Committee Agenda, April 30, 2013*. Retrieved from: http://www.regionofwaterloo.ca/en/regionalGovernment/resources/FA2013-0430.pdf

Reitano, S. (2013, August 17). Research Findings – Literature Review. The Benefits of Open Data [website]. Retrieved from: http://beautifuldata.ca/opendata/?p=373

Research Data Strategy Working Group. (2011). *Mapping the Data Landscape: Report of the 2011 Canadian Research Data Summit*. Retrieved from: http://longrangeplan.ca/wp-content/uploads/2010/10/Data_Summit_Report.pdf

Routine Proceedings. (2012, November 5). Canada. Parliament. House of Commons. *Edited Hansard 146(176).* Retrieved from: http://www.parl.gc.ca/HousePublications/Publication.aspx?DocId=5820476

Routine Proceedings. (2013, February 7). Canada. Parliament. House of Commons. *Edited Hansard 146(206)*. Retrieved from: http://www.parl.gc.ca/HousePublications/Publication.aspx?DocId=5966209

Smith, C. (2011, October 28). *Profiting from open data.* Ottawa Business Journal. Retrieved from: http://www.obj.ca/Opinion/2011-10-28/article-2787950/Profiting-from-open-data/1

Stratford Institute for Digital Media. (2012). *Becoming a Digital Nation - Provincial and Territorial Open Government Initiatives Report*. Waterloo, ON. Retrieved from: http://brainmaven.com/static/corp/OpenGovernment.pdf

Treasury Board of Canada Secretariat (TBCS). (2009). *OIC 2010-2011 Estimates - Report on Plans and Priorities*. Retrieved from: http://www.tbs-sct.gc.ca/rpp/2010-2011/inst/nd5/nd5-eng.pdf

Treasury Board of Canada Secretariat (TBCS). (2011a). Minister Day Launches Open Data Portal [press release]. Retrieved from: http://www.tbs-sct.gc.ca/media/nr-cp/2011/0317a-eng.asp

Treasury Board of Canada Secretariat (TBCS). (2011b). Expansion of Open Government [press release]. Retrieved from: http://www.tbs-sct.gc.ca/media/nr-cp/2011/0318a-eng.asp

Treasury Board of Canada Secretariat (TBCS). (2011c). Minister Clement Announces Progress on Key Open Government Initiatives [press release]. Retrieved from: http://www.tbs-sct.gc.ca/media/nr-cp/2011/1116-eng.asp

Treasury Board of Canada Secretariat (TBCS). (2011d). Expansion of Open Data Portal [press release]. Retrieved from: http://www.tbs-sct.gc.ca/media/nr-cp/2011/1202-eng.asp

Treasury Board of Canada Secretariat (TBCS). (2012a). Minister Clement Announces New Open Data Portal [press release]. Retrieved from: http://www.tbs-sct.gc.ca/media/nr-cp/2012/1024-eng.asp

Treasury Board of Canada Secretariat (TBCS). (2012b). Minister Clement Highlights Proposed New Licence for Open Data Portal [press release]. Retrieved from: http://www.tbs-sct.gc.ca/media/nr-cp/2012/1106-eng.asp

Treasury Board of Canada Secretariat (TBCS). (2012c). Minister Clement Releases Open Government Action Plan [press release]. Retrieved from: http://www.tbs-sct.gc.ca/media/nr-cp/2012/0412-eng.asp

Treasury Board of Canada Secretariat (TBCS). (2012d). Minister Clement Announces Expansion of Open Data Portal [press release]. Retrieved from: http://www.tbs-sct.gc.ca/media/nr-cp/2012/0628-eng.asp

Treasury Board of Canada Secretariat (TBCS). (2012e). Minister Clement Highlights Government's Innovation Efforts in Open Data [press release]. Retrieved from: http://www.tbs-sct.gc.ca/media/nr-cp/2012/0712-eng.asp

Treasury Board of Canada Secretariat (TBCS). (2012f). Minister Clement Highlights Government's Innovation Efforts in Open Data [press release]. Retrieved from: http://www.tbs-sct.gc.ca/media/nr-cp/2012/0717-eng.asp

Treasury Board of Canada Secretariat (TBCS). (2012g). Minister Clement Champions Government's Innovation Efforts on Open Data and Open Government to International Experts [press release]. Retrieved from: http://www.tbs-sct.gc.ca/media/nr-cp/2012/1015-eng.asp

Treasury Board of Canada Secretariat (TBCS). (2012h). Minister Clement Participates in Tweet Chat on Proposed Open Government Licence [press release]. Retrieved from: http://www.tbs-sct.gc.ca/media/nr-cp/2012/1211a-eng.asp

Treasury Board of Canada Secretariat (TBCS). (2013a). Minister Clement Speed Dates for International Open Data Day [press release]. Retrieved from: http://www.tbs-sct.gc.ca/media/nr-cp/2013/0223-eng.asp

Treasury Board of Canada Secretariat (TBCS). (2013b). Minister Clement Taps Toronto Open Data Crowd for Relaunch of data.gc.ca [press release]. Retrieved from: http://www.tbs-sct.gc.ca/media/nr-cp/2013/0311-eng.asp

Treasury Board of Canada Secretariat (TBCS). (2013c). Minister Clement Taps Edmonton Expertise at Open Data Roundtable [press release]. Retrieved from: http://www.tbs-sct.gc.ca/media/nr-cp/2013/0403-eng.asp

Treasury Board of Canada Secretariat (TBCS). (2013d). Minister Clement Connects With Ottawa Open Data Expertise [press release]. Retrieved from: http://www.tbs-sct.gc.ca/media/nr-cp/2013/0509-eng.asp

Treasury Board of Canada Secretariat (TBCS). (2013e). Minister Clement Launches Next Generation Open Data Portal [press release]. Retrieved from: http://www.tbs-sct.gc.ca/media/nr-cp/2013/0618-eng.asp

Treasury Board of Canada Secretariat (TBCS). (2013f). Minister Clement Talks Open Data with OpenText CEO Mark Barrenechea and Waterloo Tech Entrepreneurs [press release]. Retrieved from: http://www.tbs-sct.gc.ca/media/nr-cp/2013/0905-eng.asp

Treasury Board of Canada Secretariat (TBCS). (2013g). Minister Clement Promotes Open Data with Toronto Innovators and Entrepreneurs [press release]. Retrieved from: http://www.tbs-sct.gc.ca/media/nr-cp/2013/0919-eng.asp

Treasury Board of Canada Secretariat (TBCS). (2013h). Minister Clement and Canadian Innovators Explore Potential of Open Data with International Partners [press release]. Retrieved from: http://www.tbs-sct.gc.ca/media/nr-cp/2013/1011a-eng.asp

Treasury Board of Canada Secretariat (TBCS). (2013i). Canada Leads First International Working Group on Open Data [press release]. Retrieved from: http://www.tbs-sct.gc.ca/media/nr-cp/2013/1031a-eng.asp

Treasury Board of Canada Secretariat (TBCS). (2013j). Minister Clement Promotes the Canadian Open Data Experience (CODE) with Tech Students in the Waterloo Area [press release]. Retrieved from: http://www.tbs-sct.gc.ca/media/nr-cp/2013/1122-eng.asp

Treasury Board of Canada Secretariat (TBCS). (2013k). Tony Clement, President of the Treasury Board, announcing the Launch of the Next Generation Open Data Portal, data.gc.ca, and the new Open Government Licence [press release]. Retrieved from: http://www.tbs-sct.gc.ca/media/pg-gp/2013/0618-eng.asp

Treasury Board of Canada Secretariat (TBCS). (2013l). Minister Clement Highlights the Value of Open Data and Announces Plans to Launch a National Open Data Appathon [press release]. Retrieved from: http://www.tbs-sct.gc.ca/media/nr-cp/2013/1011-eng.asp

Treasury Board of Canada Secretariat (TBCS). (2013m). Minister Clement Hosts Canadian Government's First "Google Hangout" [press release]. https://www.tbs-sct.gc.ca/media/nr-cp/2013/0301a-eng.asp

Treasury Board of Canada Secretariat (TBCS). (2013n). Minister Clement Talks Open Data with OpenText CEO Mark Barrenechea and Waterloo Tech Entrepreneurs [press release]. Retrieved from: http://www.tbs-sct.gc.ca/media/nr-cp/2013/0905-eng.asp

Treasury Board of Canada Secretariat (TBCS). (2013o). Minister Clement Underlines the Value of Open Data with Geographic Information Systems Experts [press release]. Retrieved from: https://www.tbs-sct.gc.ca/media/nr-cp/2013/1022-eng.asp

Treasury Board of Canada Secretariat (TBCS). (2013p). Leads First International Working Group on Open Data [press release]. Retrieved from: https://www.tbs-sct.gc.ca/media/nr-cp/2013/1031a-eng.asp

Treasury Board of Canada Secretariat (TBCS). (2014a). Canada Publishes its G8 Open Data Action Plan for International Open Data Day [press release]. Retrieved from: http://news.gc.ca/web/article-en.do?nid=818109

Treasury Board of Canada Secretariat (TBCS). (2014b). CODE Appathon to Unlock Value of Federal Open Data [press release]. Retrieved from: http://news.gc.ca/web/article-en.do?nid=814609

Treasury Board of Canada Secretariat. (2014). Minister Clement Celebrates the Launch of CODE [press release]. Retrieved from: http://news.gc.ca/web/article-en.do?nid=820209

Tychon, G.G. (2014). Open Data Principles. *DataToGo: Open Data, Public Data and Data Visualization [website]*. Retrieved from: http://datatogo.xspatial.com/?q=content/open-data-principles