

OPEN GOVERNMENT DATA AND VALUE CREATION

Exploring the Roles Canadian Data Intermediaries Play in the Value Creation Process

Salah Merhi

Thesis submitted to the University of Ottawa

in partial fulfillment of the requirements for the degree of Doctor of Philosophy

in

Digital Transformation and Innovation

Tri-Faculty Graduate Program

Faculties of Engineering, Business and Arts

University of Ottawa

© Salah Merhi, Ottawa, Canada, 2023

Abstract

Open government data, concerned with the opening and publishing of government data in a free, accessible, and machine-readable format, aims to encourage public participation in government affairs, increasing government transparency and accountability. It is also posited that open government data will inspire businesses, the public and government agencies to use it and contribute to economic growth and value creation. The Canadian federal, provincial, and local governments have been actively opening and releasing open datasets about multiple subjects of interest to the public. However, evidence of the benefits of using open government datasets by Canadian businesses is scant, with no empirical research undertaken in Canada to understand how the data are used and what value is being created.

Based on a qualitative approach, this thesis focuses on the works and experiences of 17 professed open data intermediary firms in Canada. It aims to discover patterns and themes that provide insights into how open government data were used, the challenges facing open data intermediaries, the state of open government data, and the economic value created. The data collection is based on semi-structured interviews conducted virtually with the founder or company's executives.

In addition, the findings highlight the key similarities and differences in the activities open data intermediaries performed and the importance of resources and capabilities in developing products/services that contribute to economic value creation. Finally, five critical challenges impacting the use of open government data are identified: awareness, quality of open government data, competencies of users, data standards, and value creation.

Acknowledgment

My journey to complete this dissertation was difficult and full of challenges. However, support from family and friends and the program's co-director, Dr. Daniel Amyot, made this achievement possible.

I want to thank my present supervisor, Dr. Rocci Luppicini, and my past supervisor, Dr. Heather Morrison, for agreeing to supervise my dissertation. My thanks also go to my Ph.D. committee members, Dr. Ajax Persaud for his support and guidance, and to Dr. Herna Viktor and Dr. Seamus Ross for agreeing to participate in the examination committee, and for their insights.

This dissertation research would not have been possible without the participation of the 17 Canadian organizations who shared their knowledge, experience, and results generated from working with and using open government data. I am very grateful to all of them.

I am also grateful for the University of Ottawa's scholarship support and Dr. Amyot's NSERC grant that enabled me to complete this dissertation.

Last, I want to thank Dr. Jihad Abdalnour, Dr. Lynda Khalaf, Dr. Paul Makdissi, Dr. Myra Yazbek, and all my friends in Ottawa and abroad for their guidance, support, and motivation. I am grateful to all of you.

Table of Contents

Abstract.....	ii
Acknowledgment.....	iii
Table of Contents	iv
List of Tables	viii
List of Figures.....	ix
List of Terms and Abbreviations.....	x
Chapter 1. Introduction.....	1
1.1 Terminology.....	4
1.2 Background.....	7
1.3 Canada Open Government Strategy.....	11
1.4 Open Data Intermediary (ODI).....	15
1.5 Purpose of the Study	16
1.6 Significance of the Study	18
1.7 Structure of the Thesis	20
Chapter 2. Literature Review	22
2.1 Open Government, Open Data and Open Government Data	22
2.1.1 Open Government.....	23
2.1.2 Open Data and Open Government Data.....	25
2.1.2.1 <i>The Open Data Concept.</i>	25
2.1.2.2 <i>The Open Government Data Concept.</i>	27
2.2 Data Intermediary (DI) and Open Data Intermediary (ODI)	32
2.2.1 <i>Data Intermediaries</i>	32
2.2.2 <i>Open Data Intermediary (ODI)</i>	33
2.3 The Concept of Value	35
2.4 OGD Value Creation.....	37
2.5 Business Model (BM)	38
2.6 Theoretical Frameworks and Methods.....	41
2.6.1 <i>Resource-Based View (RBV).</i>	42
2.6.2 <i>Dynamic Capabilities (DC)</i>	44
Chapter 3. Methodology	48
3.1 Research Design.....	49
3.2 Case Study Approach.....	51

3.3 Case Selection.....	52
3.4 Data Collection	55
3.5 Identification of Participants.....	56
3.6 Theoretical Sampling	59
3.7 Ethical Considerations	60
3.8 Selection and Recruitment of Participants	61
3.9 Pilot Case	62
3.10 Interview Process	62
3.11 Data Coding and Analysis	64
3.11.1 Purposive Sampling	65
3.11.2 Open Coding	68
3.11.3 Axial Coding	69
3.11.4 Selective Coding.....	70
3.11.5 Constant Comparison	71
3.12 Trustworthiness.....	71
3.13 Chapter Summary	72
Chapter 4. Findings	73
4.1 Open Data Intermediaries Overview.....	73
4.2 Open Data Intermediary Profiles	76
4.3 Open Data Intermediaries and Open Government Data	99
4.4 How Are Open Government Data Used?.....	107
4.4.1 Custom-Made Solutions Builder	108
4.4.2 Data/Information Service Provider.....	109
4.4.3 Independent Software Developer	114
4.5 What Are the Resources and Capabilities Needed to Create Value from OGD?.....	116
4.5.1 Data Resources	117
4.5.2 Human Resources (HR)	122
4.5.3 Financial Resources.....	125
4.6 Open Government Data and Value Creation	132
4.6.1 Open Data Intermediaries and Value Creation	133
4.7 Challenges and Barriers	146
4.7.1 Organization Development Challenges	147
4.7.2 Financial Challenges	148
4.7.3 Data Challenges.....	150
4.8 Chapter Summary	153

Chapter 5. Discussion	154
5.1 Open Government Data	155
5.2 ODIs' Use of Open Government Data	155
5.2.1 Data Portal Creation	159
5.2.2 Insights and Knowledge Extraction Services	160
5.2.3 Custom-Made Solutions/Services	161
5.2.4 General Application Software Solution	161
5.3 ODIs' Resources and Capabilities	163
5.3.1 Human Resources	163
5.3.2 Financial Resources	165
5.3.2.1. Revenue	166
5.3.2.2 Government Grants and Funds	166
5.3.2.3 Business Incubators and Accelerators (BI and BA)	166
5.3.2.4 Angel Investors and Venture Capital (VC)	167
5.4 Challenges and Barriers	168
5.4.1 Organizational Development	168
5.4.2 Fundraising	169
5.4.3 Product Development	170
5.4.4 Data Acquisition	170
5.5 Data Quality and Metadata Issues and Challenges	173
5.5.1 Data Quality	173
5.5.2 Metadata	176
5.6 OGD, ODIs, and Value Creation	178
5.6.1 Direct Benefits	179
5.6.2 Indirect Benefits	179
5.6.3 Value Creation Process	180
Chapter 6. Conclusions, Implications, Limitations, and Future Work	186
6.1 Conclusions and Insights	186
6.2 Implications	193
6.3 Limitations	195
6.4 Future Research	196
References	197
Appendix A. Consent Form	228
Appendix B. Invitation Letter to Participate in a Doctoral Research Project	231
Appendix C. Certificate of Ethics Approval	234
Appendix D. Interview Protocol	235

Appendix E. Issues with Open Government Data	237
Appendix F. List of Open Data Intermediaries found on the Web	239
Appendix G. NVivo-based Coding Process.....	251

List of Tables

Table 1: <i>Open Data Definition</i>	9
Table 2: <i>Canada's Directive on Open Government Requirements</i>	14
Table 3: <i>List of Companies Used for Verifying Open Data Intermediaries' Legal Status</i>	58
Table 4: <i>List of Participating Open Data Intermediaries</i>	68
Table 5: <i>Overview of ODI Organizations</i>	74
Table 6: <i>ODIs' Locations, Motivations, and Backgrounds</i>	76
Table 7: <i>Target Markets, Products/Services, and Business Models</i>	104
Table 8: <i>List of Open Data Challenges Reported by Open Data Intermediaries</i>	152
Table 9: <i>European vs. Canadian ODI Archetypes</i>	158
Table 10: <i>ODIs' Data Services and Target Customers</i>	160
Table 11: <i>Data Types and Sources</i>	162
Table 12: <i>Factors Affecting ODIs' Success</i>	172

List of Figures

Figure 1: <i>Canada's First National Action Plan Commitments</i>	13
Figure 2: <i>The Data Environment</i>	27
Figure 3: <i>Elements of a Business Model</i>	41
Figure 4: <i>Data Collection and Analysis Process</i>	64
Figure 5: <i>ODI Distribution by Category</i>	107
Figure 6: <i>Open Government Data-based Products/Services</i>	107
Figure 7: <i>Critical ODI Resources</i>	116
Figure 8: <i>Benefits of Open Government Data Use</i>	133
Figure 9: <i>Jobs Created by ODIs</i>	145
Figure 10: <i>ODIs Challenges and Barriers</i>	146
Figure 11: <i>Canadian ODI Archetypes</i>	156
Figure 12: <i>ODI-created Products/Services</i>	159
Figure 13: <i>The Process of Value Creation from Data</i>	180
Figure 14: <i>Expanded Value Creation Process</i>	185
Figure 15: <i>Coding Files (NVivo Screenshot)</i>	251
Figure 16: <i>Final Code Categories (Top) and Sub-categories of One Higher-level Category</i> <i>(Bottom)</i>	252
Figure 17: <i>Example Sub-categories for Data Startup (Top) and for Investment, Funds and Grants</i> <i>(Bottom)</i>	253

List of Terms and Abbreviations

AEC	Architecture, Engineering, Construction
AI	Artificial Intelligence
CAD	Computer-Aided Design
DI	Data Intermediary
EC	European Commission
EU	European Union
FOIA	Freedom of Information Act (U.S. law on the freedom to access governments' data)
GIS	Geographic Information System
GT	Grounded Theory
ICT	Information and Communication Technology
LIDAR	Light Detection and Ranging (technology that uses eye-safe laser beams to create a 3D representation of the surveyed environment)
ML	Machine Learning
NGO	Non-Government Organization
NFP	Not-For-Profit
OD	Open Data as defined by the Open Knowledge Foundation (OKF)
OGD	Open Government Data is also referred to as Open Data (OD)
ODI	Open Data Intermediary (data intermediary who uses open government data)
OGP	Open Government Partnership organization
OKF	Open Knowledge Foundation
Pre-seed funding	Refers to the founders, friends, supporters, and family investment
PSI	Public Sector Information
REB	Research Ethics Board
Seed funding	Refers to funds that private investors provide, usually in exchange for an equity
Series A funding	A level of investment in a startup that follows initial seed capital
SaaS	Software as a Service

SP01-17	Code assigned to participating data intermediary startups to protect their identity
VC	Venture Capital
Web 2.0	Refers to the second generation of the World Wide Web

Chapter 1. Introduction

The open government movement reached a significant milestone in 2009 when US President Obama signed a Memorandum on Transparency and Open Government on his first day in office, committing his administration to “establish a system of transparency, public participation, and collaboration” (Obama, 2009). The US move towards an open government was followed in 2011 by the formation of the Open Government Partnership (OGP) due to efforts by the US administration. OGP, formed by eight member states (Brazil, Indonesia, Mexico, Norway, Philippines, South Africa, United Kingdom, and the United States), called on all governments worldwide to join the partnership and “commit to upholding the principles of open and transparent government by endorsing the Open Government Declaration¹.”

The Open Government Declaration commits member states to embrace the principles of transparency and open government. It also commits them to “uphold the value of openness in their engagement with citizens to improve services, manage public resources, and create a safer community².” The Open Government Declaration consists of four commitments:

1. Increase the availability of information about governmental activities.
2. Support civic participation.
3. Implement the highest standards of professional integrity throughout our administrations.
4. Increase access to new technologies for openness and accountability.

¹ <https://www.opengovpartnership.org/process/joining-ogp/open-government-declaration/>

² *ibid*

The declaration has been implemented by 76 countries and 106 local jurisdictions worldwide (OGP, 2023). The first commitment of the declaration called for increased access to information held by governments and to “proactively provide high-value information, including raw data, in a timely manner, in formats that the public can easily locate, understand and use, and in formats that facilitate reuse³.”

Janssen et al. (2012) defined open government data as “non-privacy-restricted and non-confidential data which is produced with public money and is made available without any restrictions on its usage or distribution” (p. 258). The Organization for Economic Co-operation and Development (OECD) describes open government data as “a philosophy - and increasingly a set of policies – that promotes transparency, accountability and value creation by making government data available to all.” The organization’s webpage⁴ describes the benefits of making datasets available as follows: “Public institutions become more transparent and accountable to citizens, and by encouraging the use, reuse and free distribution of datasets, they promote business creation and innovative, citizen-centric services.” The Open Knowledge Foundation (OKF) is a not-for-profit organization founded in the UK by Rufus Pollok to lobby for making all non-personal information open and free to use and share. The organization’s web page⁵ defines its mission as “to create a more open world – a world where all non-personal information is open, free for everyone to use, build on, and share.” (Open Knowledge Foundation, 2019). OKF describes open data as the building blocks of open knowledge and considers the Open Definition⁶ the benchmark for open data and content (Open Knowledge Foundation, 2023).

³ <https://www.opengovpartnership.org/process/joining-ogp/open-government-declaration/>

⁴ <https://www.oecd.org/governance/open-government-data.htm>

⁵ <https://okfn.org/about/>

⁶ <https://opendefinition.org/>

Opening government data and making it available to the public is made possible by advances in information and communication technologies (ICT), the Internet, and Web 2.0 technologies. The development of these technologies introduced structural changes to the relationship between government and citizens by making information and data availability essential to doing business. The United Nations organization describes data as “the lifeblood of decision-making and the raw material for accountability”. Without high-quality data providing the right information on the right things at the right time, designing, monitoring, and evaluating effective policies becomes almost impossible (U.N. Independent Expert Advisory Group, 2014).

Founding and new members of OGP began adopting and implementing open government data initiatives in 2011(Ruijter et al., 2020), legislating policies and regulations (ODB, 2016), and publishing government data in an open format. Between 2010-2020 the building of open government “*data portals*” increased internationally, making open datasets available to the public in growing numbers. The movement to open government data and the pervasive availability of data on the Internet attracted practitioners and scholars to study the trend and understand its motivations and potential. As a result, the term “*Open Data*” became synonymous with innovation and value creation (Magalhaes & Roseira, 2017). Furthermore, it was referred to as the fuel of the new century that will foster economic growth and contribute to social welfare (Kundra, 2012; Manyika et al., 2013; The Economist, 2017).

Making government data available to the public, license-free, and in an open format was projected to contribute to enormous economic growth. The European Commission (E.C.) estimated the annual data market in billions of USD (Berends, Carrara, & Radu, 2017). Another report

by McKinsey Global Institute indicated that open data has the potential to unlock as much as 3 trillion USD in additional value annually (Manyika et al., 2013). PIRA International estimated the value of public sector information (PSI) in Europe to lie between 28 billion EUR and 134 billion EUR per annum (Pira International Ltd., 2000). These reports and others raised the expectations of open government data's potential economic and social value and contributed to a euphoria that lacked a strong foundation.

After ten years of adopting and implementing OGP recommendations on open government by many countries worldwide, information about the economic return generated from open government data is scant and limited to a few studies by international organizations (Gurin, 2019; Huyer & Van Knippenberg, 2020; Stott, 2014). Furthermore, in a report on the economic impact of open data published by the European Commission in 2020, Huyer and Van Knippenberg indicated that questions are being raised on the validity of the economic expectations of open government data, its uses and users, and the generated value (Huyer & Van Knippenberg, 2020).

This thesis examines the economic value generated by open government data intermediary firms using open government data in Canada. The economic value will be examined through the firm's revenue and job creation aspects.

1.1 Terminology

This dissertation uses the following terminology.

Open: The term open is defined concerning data to mean “anyone can freely access, use, modify, and share data for any purpose” (Open Knowledge Foundation, 2017).

Data: Data is described as “raw material produced by abstracting the world into categories, measures and other representational forms – numbers, characters, symbols, images, sounds, electromagnetic waves, bits – that constitute the building blocks from which information and knowledge are created” (Kitchin, 2014 p. 1).

Open Data (OD): There are numerous definitions of open data built on the principles of availability, accessibility, reuse and redistribution, universal participation, and computability. The government of Canada portal defines open data as “structured data that is machine-readable, freely shared, used and built on without restrictions” (Government of Canada, n.d.).

Open Government Data (OGD): Open government data is a subset of open data. The Organization for Economic Co-operation and Development (OECD) defines open government data as: “a philosophy- and increasingly a set of policies - that promotes transparency, accountability and value creation by making government data available to all” (Ubaldi, 2013).

Open Government: The Canadian government portal states, “Open Government is about making government more accessible to everyone, giving greater government data and information access to the Canadian public and the business community (Government of Canada, n.d.).

E-Government: E-government refers to the use of information and communication technology in public administration and government operations to deliver government services to citizens and businesses more effectively and efficiently (OECD, 2002; UN, n.d.).

Data portal: A data portal is described as “a web application, website, or page that holds data from different sources, organized under subsets or categories to make it simple for the site users to find” (ZUAR, 2022).

Open Data Portal: A web-based interface designed to make it easier to find reusable information, like library catalogues, contains metadata records of datasets published for reuse, mainly relating to information in the form of raw, numerical data rather than textual documents (European Commission, n.d.).

Data intermediary (DI): An individual or organization that has selected data as its domain of operation. “Data intermediary serves as a mediator between those who provide data and those who seek to leverage that data” (H. Janssen & Singh, 2022, p. 1).

Open Data Intermediary (ODI): Someone positioned at a point in a data supply chain that incorporates an open dataset, who is located between two others in the supply chain, and who facilitates the use of open data that may otherwise not have been the case (Schalkwyk et al., 2015, p. 6).

Business Model: The business model concept has various definitions that converge to describe how a firm creates and captures value (Lahti et al., 2018). It is a conceptual tool containing a set of objects, concepts and their relationships to express the business logic of a specific firm (Osterwalder et al., 2005).

Value: This is the monetary, material, or assessed worth of an asset, goods, or service.⁷ Bowman and Ambrosini (2000) suggested that value should be evaluated in terms of “*use value*” and “*exchange value*”. They defined “*use value*” as referring to the specific qualities of the product perceived by customers’ needs and “*exchange value*” to refer to price, the monetary amount realized at a single point in time when the exchange of the good takes place (pp. 3-4). This thesis is concerned with the exchange value ODIs realize from selling their products/services.

Value Creation: Value creation is the process of the combination and deployment of labor and resources into something that meets customers’ needs (Bowman & Ambrosini, 2000).

1.2 Background

Government data is a core component of the open government movement intended to increase governments’ transparency, accountability, and citizen participation. The drive to access governments’ data is as old as the Athenian democracy (Hedrick, 1999). However, in contemporary history, the campaign to access government information started in the U.S. during the 1950s and culminated in 1966 when U.S. legislators adopted the Freedom of Information Act (FOIA). The FOIA was intended to increase government transparency and accountability and institute a system whereby U.S. citizens, foreign nationals or organizations could make a FOIA request. Following the U.S. FOIA initiative, democratic countries began to follow its path worldwide. For example, Canada’s Access to Information Act was enacted in 1983. Legislators defined the act’s purpose as “to enhance the accountability and transparency of federal institutions to promote an

⁷ <https://www.investopedia.com/terms/v/value.asp>

open and democratic society and enable public debate on the conduct of those institutions” (Justice Law, 2014).

The commercialization and mainstreaming of the Internet in the last decade of the 20th century, coupled with advances in information and communication technologies (ICT) and Web 2.0 technology, led to transforming governments’ operations from a paper-based structure into a digital-based one (Fountain, 2001). The concept of e-government refers to the use of information and communication technology, including the Internet, for exchanging information, providing services, and transacting with citizens, businesses, and other arms of government, among other internal activities, to build more efficient operations and processes (Curtin, Sommer, & Vis-Sommer, 2013; Sandoval-Almazan et al., 2014; Silcock, 2001). The increased use of ICT and Internet technologies by democratic governments resulted in a vast amount of data being collected and stored in electronic form in computer storage devices.

A report prepared for the European Commission in September 2000 recognized that “public bodies” were the largest information producers in Europe and that this information was an under-exploited asset (Pira International Ltd., 2000). The European Union (E.U.), recognizing the importance of public sector information, adopted, and issued directive 2003/98/E.C. on the re-use of public sector information in December 2003. The Directive called on all E.U. states to open their data and make them accessible to the public (European Union, 2003). Following the European initiative to open government data, 30 thought leaders and Internet activists in the U.S. met in Sebastopol, California, in December 2007. At this meeting, they adopted eight principles that defined and evaluated open public data and called on U.S. presidential candidates to adopt them

(Chignard, 2013; Tauberer, 2007). The Sebastopol declaration⁸ considered the government data to be open if it was made public and complied with the following principles: (a) Complete, (b) Primary, (c) Timely, (d) Accessible, (e) Machine-processable, (f) Non-discriminatory, (g) Non-proprietary, and (h) License-free. Table 1 explains the meaning of each concept.

Table 1: *Open Data Definition*

Item	Characteristic	Meaning
a	Complete	All public data is made available. Public data is data that is not subject to valid privacy, security, or privilege limitations.
b	Primary	Data is as collected at the source, with the highest possible level of granularity, not in aggregate or modified forms.
c	Timely	Data is made available as quickly as necessary to preserve the value of the data.
d	Accessible	Data is available to the widest range of users for the widest range of purposes.
e	Machine processable	Data is reasonably structured to allow automated processing.
f	Non-discriminatory	Data is available to anyone, with no requirement of registration.
g	Non-proprietary	Data is available in a format over which no entity has exclusive control.
h	Licence-free	Data is not subject to any copyright, patent, trademark, or trade secret regulation. Reasonable privacy, security, and privilege restrictions may be allowed.

Note: Retrieved from Open Government Data Principles (https://public.resource.org/8_principles.html)

The push for open government movement gained momentum when U.S. President Obama signed a Memorandum on Transparency and Open Government on his first day in office, January 21, 2009. The memorandum committed his administration to create an “unprecedented level of

⁸ <https://opengovdata.org/>

openness in government.” In addition, he called on his administration to ensure public trust and establish a system of transparency, public participation, and collaboration (Obama, 2009).

The push began when seven countries (Brazil, Indonesia, Mexico, Norway, the Philippines, South Africa, and the United Kingdom) joined the U.S. in 2011 to form the Open Government Partnership (OGP). The organization’s main objective is to work with civil society and the private sector to co-create and implement ambitious open government reforms anchored in four guiding principles:

1. Increase the availability of information about governmental activities.
2. Support civic participation.
3. Implement the highest standards of professional integrity throughout the administration.
4. Increase access to new technologies for openness and accountability. (OGP, 2011)

On June 18, 2013, the leaders of the G8 countries, namely Canada, France, Germany, Italy, Japan, Russia, United Kingdom and the United States of America, signed what came to be known as the “Open Data Charter” to promote transparency, innovation and accountability (Charter, 2013). The Open Data Charter⁹ sets out five strategic principles that all G8 members are committed to acting up on:

1. Open data by default.
2. Quality and quantity.
3. Usable by all.

⁹ <https://www.gov.uk/government/publications/open-data-charter>

4. Releasing data for improved governance.
5. Releasing data for innovation.

According to the Open Data Charter (ODC) website¹⁰, the five principles were welcomed by many nations. However, during the international open data conference in Ottawa at the end of May 2015, the participants had a general sense that the principles could be refined and improved to support the broader global adoption of open data principles. More than 350 comments from all over the world were registered. Subsequently, the International Open Data Charter was launched in 2015 to address the recommendations for refining the Open Data Charter. Following a global consultation with governments and civil societies, and experts from around the world, six ODC principles were developed in 2015 to represent “a globally-agreed set of aspirational norms for how to publish data” (ODC, 2015):

1. Open by default.
2. Timely and comprehensive.
3. Accessible and usable.
4. Comparable and interoperable.
5. Improved governance and citizen engagement.
6. Inclusive development and innovation.

1.3 Canada Open Government Strategy

Canada’s open government strategy was launched in 2011 with the announcement that the government of the day was launching an open government initiative on three main streams: open

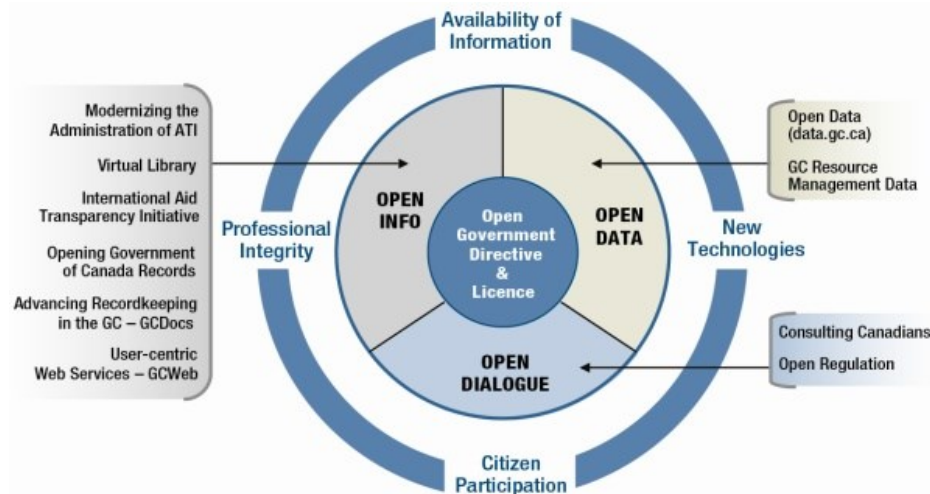
¹⁰ <https://opendatacharter.net/our-history/>

information, data, and dialogue. The same year, the government started an open data pilot project and announced its intention to join the Open Government Partnership (OGP). Canada formally joined the organization in April 2012, adopting the Open Government Declaration and committing to its principles. Subsequently, Canada released its first National Action Plan (NAP) in 2012. The introduction to the first National Action Plan states:

Canada's commitment to open government is part of the federal government's efforts to foster greater openness and accountability, to provide Canadians with more opportunities to learn about and participate in government, to drive innovation and economic opportunities for all Canadians and, at the same time, create a more cost-effective, efficient and responsive government. (Treasury Board Secretariat, 2012)

Canada's first action plan focused on 12 key commitments; two core commitments considered foundational to the success of the open government strategy and ten commitments split between the three open government initiative streams. The development of the pledges was informed by consultations with citizens across the country, members of civil society, the private sector, key federal departments and agencies, and other levels of government¹¹. Figure 1 is a representation of Canada's 12 commitments.

¹¹ <https://open.canada.ca/en/canadas-action-plan-open-government>



Note: Figure reprinted from Canada's Action Plan on Open Government (Treasury Board Secretariat, 2012)

Figure 1: *Canada's First National Action Plan Commitments*

Canada unveiled its first open data pilot portal (data.gc.ca) in 2011 and relaunched it in 2013 under the name open government portal (open.canada.ca). By January 2023, the open data portal had grown to contain 31,000 open datasets covering economics and industry, agriculture, health and safety, labour, transport, and society and culture, to name a few¹². It also reported more than 260,000 visits in January 2023, with the number of downloads reaching 130,000 in the same month (Government of Canada, 2022). However, this vast increase in the number of open datasets available and the number of downloads of datasets is not reflected in any study on the Canadian data intermediaries' use of open data. Hence, this thesis examines how Canadian data intermediaries use the government's open datasets, studies what they do with them, and discovers if and what value is created from open government data.

¹² <https://search.open.canada.ca/opendata/>

Canada's commitment to open government was confirmed in its directive, released on October 9th, 2014, that sets out the government's objective of releasing government information and data of business value: "to support transparency, accountability, citizen engagement, and socioeconomic benefits through reuse, subject to applicable restrictions associated with privacy, confidentiality, and security" (Directive on Open Government - Canada.ca, 2014). The directive listed six requirements (see Table 2) relating to the release of open data and open information assigning the responsibility for their implementation to senior departmental information management officials.

Table 2: *Canada's Directive on Open Government Requirements*

Item	Requirement
1	Maximizing the release of Government of Canada open data (structured data) and open information (unstructured documents and multi-media assets) under an open and unrestrictive licence designated by the Treasury Board of Canada Secretariat
2	Ensuring that open data and open information is released in accessible and reusable formats via Government of Canada websites and services designated by the Treasury Board of Canada Secretariat.
3	Establishing and maintaining comprehensive inventories of data and information resources of business value held by the department to determine their eligibility and priority, and to plan for their effective release.
4	Developing, posting to the designated website, implementing, and annually updating a departmental Open Government Implementation Plan (OGIP)
5	Maximizing the removal of access restrictions on departmental information resources of enduring value prior to transfer to Library and Archives Canada as part of planned disposition activities.

Item	Requirement
6	Ensuring that open government requirements in Sections 6.1 to 6.5 of this directive are integrated in any new plans for procuring, developing, or modernizing departmental information applications, systems, or solutions in support of the delivery of programs and services.

Note: Reprinted from the Directive on Open Government (Canada Treasury Board, 2014).

In making government data available to the public, the government hoped that citizens would access and use the data and thereby increasing government transparency, accountability, participation, business activities and job creation and stimulating economic growth (Canada Treasury Board, 2014). Canada has released five National Action Plans since 2012, meeting the OGP's requirements.

1.4 Open Data Intermediary (ODI)

An *Open Data Intermediary* is an actor who selects open government data as the primary resource for building products/services to solve a problem, answer a question, or provide services to other actors, individuals, businesses, or public organizations (Dumpawar, 2015; Schalkwyk et al., 2015). The term “open data intermediary” indicates that the actor is an intermediary who acts between two or more entities and uses open government data to create products/services and conduct business.

Davies (2014) described the role of intermediaries in the open data ecosystem as “not just important as users and re-users of open data. They often play a vital role in articulating the demand for data and working with the government to implement the supply of open data, translating data into information, knowledge and action” (p. 15). Schalkwyk et al. (2014) described the role open

data intermediaries play in the ecosystem as follows: “Increases the accessibility and utility of data, provide both supply-side as well as demand-side value, enabling role of mediation and innovation within the ecosystem, and may assume the role of a keystone species in a data ecosystem” (p. 4). Open government data is assumed to have great potential to foster economic growth and create social value. However, data has no value until used (Attard et al., 2016; Janssen et al., 2012). Thus, open data intermediaries are needed to unlock this potential of open government data for generating various economic, cultural, and social values. Davies (2014) described the efforts to extract benefits from open data as a linear theory of change consisting of two variables: *data + intermediaries = impact* (p. 5). The two variables must exist together to generate value from open government data.

1.5 Purpose of the Study

Since 2011, the Canadian federal, provincial, and local governments have built open data portals to make open datasets available to the public. As a result, thousands of datasets have been published by the three levels of government under various categories. In January 2023, the number of datasets listed on the open data portal reached 35,566, comprising 31,001 open datasets and 4,565 open information. The open datasets include 27,754 federal, 7,042 provincial and 770 municipal data sets¹³.

As the number of open datasets made available grows in Canada and abroad, questions abound about whether this data delivers on the premise associated with its economic potential. A report published by the Open Data Barometer¹⁴ (ODB), titled “From Promise to Progress” (2018),

¹³ <https://search.open.canada.ca/opendata/>

¹⁴ Report produced by the World Wide Web Foundation (Webfoundations.org, 2018)

states that at the time, little historical evidence of tangible benefits from open government data initiatives had surfaced, particularly for social impact (Open Data Barometer, 2018).

In a series of three opinion pieces on “Open Data – The Promises, The Disillusion and The Panacea,” published by the Global Digital Foundations¹⁵ and by the medium.com website, Dr. Audrey Lobo-Pulo, a then senior advisor in the Australian Public Service, questioned the economic impact of open data claiming that the consumption of open data and how it addresses specific economic and social issues are still being researched, and while there have been case studies examining the use of open data for specific companies, a broader view of the global impact is yet to be fully explored (Lobo-Pulo, 2017).

The literature on open data is rich with studies that address the impact of open data from different perspectives: Citizen’s trust (Purwanto et al., 2020), business models (Germano et al., 2016), benefits (M. Janssen et al., 2012), ecosystems (Reggi & Dawes, 2016b), innovation (Gurin, 2014), and value creation (Gustavsen, 2017). Safarov et al. (2017) indicated that a few empirical studies approached the subject from a utilization perspective, with most focusing on the provision of OGD and assuming but not empirically testing to identify the presence of various forms (p. 1).

The dearth of empirical studies examining open government data use in Canada and the world points to a gap in knowledge about open government data usage, the types of products/services developed, and the economic value created. This gap is the primary motivator for this

¹⁵ <https://www.globaldigitalfoundation.org>

exploratory study that aims to discover how open data intermediaries use OGD, the products/services they create, and the economic value they generate.

This thesis examines how a sample of Canadian open data intermediaries use open government data in Canada, the products/services created, and how these products/services contribute to economic value creation. The study is based on an inductive qualitative research approach to empirically find instances in which firms realized economic benefits using open government data. In addition, it seeks to expand the knowledge and understanding of the field of OGD, and the economic benefits projected from its use. An in-depth analysis of open government data intermediaries' firms is necessary to uncover the motivations, resources, and processes that may contribute to generating economic value from open government data. Thus, the following question and sub-questions are used to guide the research.

RQ: How do data intermediaries create value using open government data?

- a. How are open data intermediaries using open government data?
- b. What are the resources and capabilities vital to open data intermediaries?
- c. What are the challenges and barriers faced by open data intermediaries?
- d. What economic values are created, and how?

1.6 Significance of the Study

The significance of this study lies in being the first empirical research study on the applied use of open government data by business-oriented data intermediaries in Canada.

The concept of intermediation is as old as humanity. However, political, economic, and social considerations shaped it. The digital transformation impacting society since the mid-1990s has created new opportunities for intermediaries in the information and communication technology space and, subsequently, in the data space that appeared following the permeation of the Internet and Web 2.0 technologies. The data space has expanded massively during the past two decades and led to increased global data availability made possible by the Internet (Heath & Bizer, 2011; Staab & Bizer, 2009). Consequently, the vast volume of ubiquitous data led to a new type of data intermediary actors that saw opportunities to create products/services out of the data using their knowledge and technical skills. The actors used open government data when open data began to appear in 2011 (Richter & Slowinski, 2019; Yoon, Copeland, & McNally, 2018).

In studying Canadian ODIs, the researcher aims to identify the traits that distinguish these data intermediaries and drive them to undertake challenging initiatives using open government data. Furthermore, the thesis seeks to identify the resources vital to data intermediaries' success, the challenges and barriers impacting their efforts, and the value they created using open government data. The study is expected to expand the knowledge of open government data use in Canada and reveal what value is created. It will also inform the critical elements of the successful transformation of open government data into products/services and the economic benefits generated.

The outcomes of this thesis are expected to influence policymakers' decisions on expanding the use of open government data, companies exploring open government data and academics and researchers in the fields of open government data and data intermediaries. In addition, the

researcher expects the outcomes to provide insights into the state of Canada's open government data ecosystem, the stakeholders, and the value anticipated.

1.7 Structure of the Thesis

This thesis comprises six chapters. This introductory chapter sets the background and context for the following discussion.

Chapter 2 reviews the literature about open government data, open data intermediaries and value creation, as well as theoretical frameworks and methods suitable for analyzing the work and performance of open data intermediaries. The review addresses the use of open government data by data intermediaries, the business model and value creation, and the resource-based view (RBV) and dynamic capabilities (DC) theories.

Chapter 3 covers the research methodology and design and describes the rationale behind its selection. It presents the data collection method, the identification of participants and the selection of firms, the interview process and data coding and analysis.

The research findings are set out in Chapter 4, where the participating open data intermediaries' motivations, approaches and use of open government data are presented. Chapter 4 shows the products/services created, the resources needed, the challenges and issues reported, and the economic value reportedly generated from open government data.

Chapter 5 discusses the research findings and answers the research questions. Finally, Chapter 6 presents the thesis conclusions, implications, limitations, and future work.

Chapter 2. Literature Review

Open government data is made available to the public based on the assumption that it can contribute to creating economic and social value. However, the mere availability of data does not establish its value credential; data have to be manipulated, processed and/or used to create value (Attard et al., 2016). The discussion in this chapter reviews the academic literature on the role data intermediaries play in the open data ecosystem, the methodologies and processes reportedly used to exploit open government data, and some of the business models and theoretical frameworks associated with creating economic value.

In Section 2.1, definitions for open government, open data, and open government data are recalled, their objectives and potential users are explored, and their hypothesized economic impact is discussed. Section 2.2 reviews the concept of data intermediaries and open data intermediaries and discusses the different types of open data-related work conducted by these entities. Section 2.3 examines the concept of “value” in relation to the open data ecosystem, while Section 2.4 discusses open government data value creation. Section 2.5 reviews the business model concept of the operations of different data intermediaries. Section 2.6 reviews theoretical frameworks that can directly or indirectly be used to evaluate the work of open data intermediary firms.

2.1 Open Government, Open Data and Open Government Data

Open government data is based on two concepts: open government and open data. The discussion in this section examines the evolution of open government data, its often-associated economic objectives, and the hypothesized economic potentials attributed to it.

2.1.1 Open Government

The concept of open government is rooted in the associated liberal democracy with transparent and accountable government. In the introduction to their book *Open Government - Collaboration, Transparency and Participation in Practice*, Lathrop and Ruma (2010) described the basis of open government as “the notion that the people have the right to access the documents and proceedings of government” (p. xix). They further noted that the concept is evolving and is influenced by different factors, including the open-source software movement and citizen participation. The notion of citizen participation includes monitoring and influencing the government (Meijer et al., 2012). Having synthesized various definitions of open government, Wirtz and Birkmeyer (2015) defined open government as “a multilateral, political, and social process, which includes transparent, collaborative, and participatory action by government and administration” (p. 382). The Organization for Economic Co-operation and Development (OECD) advanced a definition that sees open government as “a culture of governance based on innovative and sustainable public policies and practices inspired by the principles of transparency, accountability, and participation” (p. 1).

Following the end of World War II, in the United States, the idea of open government began to gain traction in the 1950s (Yu & Robinson, 2012). The push for more open government achieved its first success by enacting the Freedom of Information Act (FOIA) in 1967. However, the most potent enabler to open government came in the mid-1990 with the revolution in information and communication technology (ICT).

The permeation of personal computers among the population and the evolution of e-government throughout the 1990s accelerated the demand for government transparency and

accountability (Evans & Campos, 2014; Harrison et al., 2012; Meijer et al., 2012). The European Union (EU) directive on the reuse of Public Sector Information, known as the PSI Directive, is a direct result of the open government movement. The directive focused on the economic aspects of the reuse of information and encouraged the member states to make as much information available for reuse (European Commission, 2003).

On December 9, 2009, the Obama Administration issued a directive on open government, calling on the heads of executive departments and agencies in the federal public service to take specific steps toward creating a more open government (OMB, 2009). The first item in the directive linked the publication of government information online to increasing accountability, promoting informed participation by the public, and creating economic opportunity. The directive also instructed each agency to proactively expand access to information by making it available online in open formats (OMB, 2009).

Canada published its Directive on Open Government on October 9, 2014¹⁶. The Directive's objective was to "maximize the release of government information and data of business value to support transparency, accountability, citizen engagement, and socio-economic benefits through reuse, subject to applicable restrictions associated with privacy, confidentiality, and security" (Canada Treasury Board, 2014). Common among the Canadian and US directives is a shared emphasis on the socio-economic benefits that the release of open government data would generate and the establishment of open data portals in the US (data.gov) and Canada (open.canada.ca) as repositories for open datasets.

¹⁶ <https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=28108>

2.1.2 Open Data and Open Government Data

2.1.2.1 The Open Data Concept. “Open data was just an idea, emerging as a rough consensus for action among pro-democracy practitioners, internet entrepreneurs, open-source advocates, civic technology developers, and open knowledge campaigners” (Davies et al., 2019, p. 1).

The term “Open Data” developed into both a movement and a label applied to a large volume of data, from genomics and geospatial data to land registers, contracting, and parliamentary voting (Davies et al., 2019). According to Chignard (2013), the term “open data” first appeared in 1995 in a document from an American scientific agency that dealt with geophysical and environmental data disclosure¹⁷. Dumpawar (2015) argues that the term *open* has become a general descriptor in various rights to access movements such as open-source, open content, and open science (p. 20). In addition, the notion of an open government movement adds another dimension to that term by calling upon governments to open their data and make it available to the public for use and reuse.

The Open Knowledge Foundation (OKF) defines open data as: “Data that can be freely used, shared and built-on by anyone, anywhere, for any purpose” (James, 2013). The OKF considered the term open to apply to information from any source and about any topic released by organizations, public and non-public, under an open license for free use by and benefit to the public. It associated the term open with knowledge and advanced the following definition: “Knowledge is open if anyone is free to access, use, modify, and share it — subject, at most, to measures that preserve provenance and openness” (Open Knowledge Foundation, 2017). The

¹⁷ <https://www.paristechreview.com/2013/03/29/brief-history-open-data/>

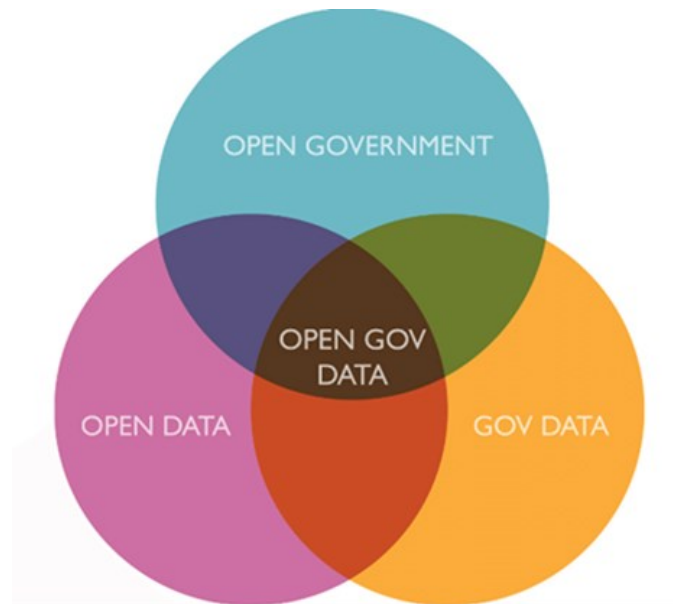
“work”, a term used by the OKF to denote the item or piece of knowledge being transferred, is considered open if it complies with four characteristics: (a) Open license, meaning the work must be in the public domain or provided under an open license, (b) Accessible and downloadable from the Internet without charge, (c) Machine-readable, meaning the work must be provided in a form readily processable by a computer and where the individual elements of the work can be easily accessed and modified, and (d) Open format, which places no restrictions, monetary or otherwise, upon its use and can be fully processed with at least one free/libre/open-source software tool (Open Knowledge Foundation, 2017).

Yu and Robinson (2012) recognized that the label open, as applied to various kinds of information, inherits both technological and philosophical meanings. These authors indicated that, at the technical level, “the term suggests using computers to handle information efficiently in place of manual human processing” (p. 188). In contrast, at the philosophical level, “the term suggests participation and engagement” (p. 189).

The definition of open government data became the standard to follow when opening government-held data. The only difference between the terms open data and open government data is the sources of their availability. The former refers to data from different sources globally, the latter to subsets of government-held data made publicly available without restriction on its use, and which is predominantly accessed by specific government-sanctioned data portals.

The terms open data and open government data are used in this thesis interchangeably to refer to “structured data that is machine-readable, freely shared, used and built on without any

restriction” (Government of Canada, n.d.). Figure 2 illustrates the environment of open data and open government data.



Note: Retrieved from “Guidelines on Open Government Data for Citizen Engagement” (UNDESA, 2013).

Figure 2: *The Data Environment*

2.1.2.2 The Open Government Data Concept. The Organization for Economic Co-operation and Development (OECD) defined open government data as “a philosophy- and increasingly a set of policies - that promotes transparency, accountability and value creation by making government data available to all” (OECD, n.d.). The organization considered public institutions becoming more transparent and accountable to citizens by opening their data. In addition, governments promote business creation and innovative, citizen-centric services by encouraging the use, reuse, and free distribution of datasets.

The United Nations defined OGD as a composition of two elements: (a) Open data is material anyone can use without restrictions, (b) Government data or public sector information is any data and information produced or commissioned by public sector bodies (UNDESA, 2013). Similarly, OKF defined open government data as: “Data produced or commissioned by the government or government-controlled entities, and data that is open as defined in the Open Definition – that is, it can be freely used, reused and redistributed by anyone” (OKF, 2019).

Open government data is a subset of open data and is simply government-held data made open to the public (Kučera, Chlapek, & Nečaský, 2013). The movement to open public sector information was related to e-government inquiries addressing democratic theory, voter participation, democratic deliberation, and open government in a broader context (Lakomaa & Kallberg, 2013). PSI is the name given to government data before the “open data” term was coined later in the US. The European Union (EU) became the first public organization to issue a directive promoting the re-use of public sector information, considering it an instrument for both extending the right to knowledge and permitting European companies to exploit its economic potential (EU Directive, 2003). Six years later, the Obama Administration published its Directive on Transparency and Open Government, followed by the UK government, which set out its objectives for opening the data it held and for the development of an open data portal focusing on four key dimensions (Davies, 2010):

1. Promoting transparency and accountability.
2. Empowering citizens to drive public sector reform.
3. Releasing the economic and social value of information.
4. Putting Britain at the leading edge of semantic-web developments.

As the movement to open government data spread, it, along with its hypothesized promises and potentials, became the focus of research from a wide range of perspectives: users acceptance (Zuiderwijk & Cligge, 2016), benefits (Janssen et al., 2012), business models (Germano et al., 2016; Chambers & Patrocinio, 2011), value creation (Gustavsen, 2017), social and technical barriers (Beno et al., 2017), as well as adoption (Maccani, 2016), social and economic impact (Koski, 2015), user acceptance (Weerakkody et al., 2017), and its ecosystems (Reggi & Dawes, 2016a).

Janssen et al. (2012), who examined the benefits, barriers and myths of both open data and open government, identified three clusters of benefits: (a) political and social, (b) economic and (c) operational and technical benefits. The authors reported that “economic growth and contribution to public values, transparency and accountability, are the overarching arguments for stimulating open data” (p. 260). They listed seven economic benefits of publishing open data:

1. Economic growth and stimulation of competitiveness.
2. Stimulation of innovation.
3. Contribution toward the improvement of processes, products and/or services.
4. Development of new products and services.
5. Use of the wisdom of the crowds: tapping into the intelligence of the collective.
6. Creation of a new sector adding value to the economy.
7. Availability of information for investors and companies.

Tinholt (2013) classified the projected benefits into three categories: (a) revenue generation through multiple areas, (b) cost saving and increased efficiencies, and (c) job creation and skills development (p. 7). Complementing this view, Davies and Perini (2013) classified the projected

benefits into three broad, complementary categories: (a) transparency and accountability, (b) innovation and economic development, and (c) inclusion and empowerment (p. 16). They summarized the characteristics of each category as follows:

- Transparency and accountability: open data will bring about greater transparency, which brings about greater accountability of key actors, leading to them making decisions and applying rules in the public interest.
- Innovation and economic development: Open data will enable non-state innovators to improve public services or build new products and services with social and economic value. Open data will shift specific decision-making from the state into the market.
- Inclusion and empowerment: Open data will remove power imbalances resulting from asymmetric information and bring new stakeholders into policy debates, giving marginalized groups a more significant say in creating and applying rules and policy.

In 2014 the US Department of Commerce estimated the revenues of 13 selected companies in the government data-intensive sector to exceed USD 24 billion and the number of employments to reach 90,000 at the lower bound. It projected the upper-bound revenue to exceed USD 220 billion (US Department of Commerce Economics and Statistics Administration, 2014, p. 40-41).

In a report published by the European Data Portal on the economic benefits of open data (Analytical Report No. 9), the authors estimated the market size and value of open data in the 28+ European countries to reach 325 billion EUR in 2020, with 100,000 jobs in open data created in 2020 generating some 1.7 billion EUR in cost-saving (Berends et al., 2017). In a follow-up report in January 2020 titled “The Economic Impact of Open Data: Opportunities for value creation in

Europe,” the authors forecast an open data market size for Europe in 2025 to be in the range of 199.51 to 344.20 billion EUR, creating some 1.12-1.97 million open data employees (Huyer & Van Knippenberg, 2020).

While there is a general agreement among academics and practitioners that open government data affords the potential to generate economic and social value, there are many barriers to realizing this potential. Obstacles are manifest on both the providers’ side (Barry & Bannister, 2014) and the users’ side (M. Janssen et al., 2012), ranging from opening and releasing the data to its adoption (Maccani, 2016) and using it for any array of purposes (Smith, Sandberg, Viktoria, & Gothenburg, 2018). In a study on the socio-technical impediment of open data, Zuiderwijk et al. (2012) identified 118 socio-technical constraints impacting the use of open data. They classified the barriers into ten categories, from data finding and accessibility to data quality and metadata problems. The authors argued that “scant attention is paid to the user perspective, whereas the user needs to generate value from the open data” (p. 167).

Despite the barriers and impediments found in implementing and using open government data, scholars and practitioners believe that opening government data and releasing it to the public in an open, accessible, and machine-readable format for reuse will increase government transparency, accountability, and public participation. Moreover, they concur that open government data stimulate innovation, create business opportunities, and create social-economic value (Alexopoulos et al., 2016; Attard, Orlandi, & Auer, 2016; Davies & Perini 2016; Kassen, 2018; Lakomaa & Kallberg, 2013; Magalhaes & Roseira, 2017; Zeleti, Ojo, & Curry, 2016).

2.2 Data Intermediary (DI) and Open Data Intermediary (ODI)

2.2.1 Data Intermediaries

Data intermediaries use data to create social and economic value through innovative products/services that solve problems, answer questions, and extract insights and knowledge (Elcock, 2021). The term data intermediary is associated with the explosion of data coinciding with the commercialization and mainstreaming of the Internet in the mid-1990. However, the concept of an intermediary is old. Depending upon their operations' context and roles therein, intermediaries have been described as innovators (Agogu   et al., 2017; Hargadon & Sutton, 1997), entrepreneurs (Peng, Lee, & Hong, 2014a; Poppen & Decker, 2018), brokers (Ye, Kankanhalli, & Yang, 2012), agents (Howells, 2006a; Treuhaft, 2006) and support organizations (Brown & Kalegaonkar, 2002).

In the context of innovation, Howells (2006) describes an intermediary as an “organization or body that acts as agent or broker in any aspect of the innovation process between two or more parties” (p. 720). In addition, Stewart and Hyysalo (2008) advanced the following context for innovation intermediaries:

Innovation intermediaries can be identified by their engagement in activities that gather, develop, control, and disseminate knowledge and collect and disseminate financial, technical and institutional resources. In addition, they support users and sponsors and attempt to regulate uses, developments, participation and the actions of others in the innovation networks. (p. 297)

The authors conferred the term “intermediaries” to individuals and/or organizations that are “grounded in an institutional, technical and often physical context that facilitates their activities which attempt to configure the users, the context, the technology and the content (Stewart & Hyysalo, 2008, p. 297). Peng et al. (2014) argued that entrepreneurship could be viewed as intermediation. They recommended extending the definition of entrepreneurs to include “Individuals or organizations that discover and create entrepreneurial opportunities somewhere in the value chain in an industry” (p. 6).

Publishing open government data contributed to the emergence of specialized intermediaries that seek to transform this resource into cultural, economic, environmental, political, and social products and services. These relatively “new” entities are the focus of discussion in the next section.

2.2.2 Open Data Intermediary (ODI)

In their study of open data in the governance of South African higher education, Van Schalkwyk et al. (2014) describe the efforts and activities of open data intermediaries as “providing both supply-side and demand-side value, assuming the role of a keystone species in a data ecosystem, and potentially democratizing the impacts and use of open data (p. 47). The importance of ODIs in the open data ecosystem was highlighted by Davies (2014), who describes intermediaries as “vital to both the supply and use of open data... Intermediaries can create, articulate demand for data, and help translate open data vision from political leaders into effective implementations” (p. 2). ODIs are also seen as “*barrier removal*” by unlocking the potential of data that permits its

use by communities with limited access to existing data infrastructures (Yoon et al., 2018).

Dumpawar (2015) classified the functions of ODIs into eight categories:

- aggregation,
- dissemination,
- rectifications,
- interpretation,
- representation,
- dissemination,
- augmentation,
- connection.

Magalhaes et al. (2013) provided an analysis of the literature about open government data intermediaries and reported three types: (a) civic startups, (b) open data services, and (c) infome-diaries (p. 336). In addition, a 2012 study by Deloitte Analytics of 230 companies and not-for-profit organizations in the UK and overseas identified five archetypes of intermediary business models in the open data marketplace (Deloitte Analytics, 2012):

- suppliers,
- aggregators,
- developers,
- enrichers,
- enablers.

Drawing from synthesis research on the role of open government data, Van Schalkwyk et al. (2015) concluded that the literature on open data intermediaries lacks a comprehensive definition of ODI. Thus, they put forward their own definition describing open data intermediaries:

An agent (i) positioned at some point in a data supply chain that incorporates an open dataset, (ii) positioned between two agents in the supply chain, and (iii) facilitates the use of open data that may otherwise not have been the case. (p. 6)

As should be clear from the above, the concept of ODI is broad and seemingly fluctuates according to the contexts in which it is applied. Nevertheless, there appears to be consensus about the crucial role open data intermediaries play in leveraging open government data to create economic value through their multiple data practices (Davies, 2014).

2.3 The Concept of Value

The concept of value has been debated and theorized in the literature since the days of Aristotle (Vargo et al., 2008). According to Aristotle, use-value derives from a thing's or a service's capacity to be productive for a person's good (Grönroos, 2011). Bowman & Ambrosini (2000) noted that the concept of value has a definition problem with a tendency in the literature to use the term to refer to different phenomena (Bowman & Ambrosini, 2000). In the classical view, value is thought to be generated in exchange for goods for money. However, this understanding began to change in the late 20th century (Bowman & Ambrosini, 2000; Grönroos, 2011; Vargo et al., 2008). Vargo et al. (2008) contend that value is “fundamentally derived and determined in use, not in exchange” (p. 145) because, in their view, “value” results from integrating and applying

resources in a specific context embedded in firm output and captured by price. In a study aimed at alleviating some ambiguity associated with value, Gummerus (2013) identified two main streams of research about value in literature: One focusing on value creation processes and the other on value outcomes. Gummerus described the value creation process as entailing “the parties, activities, and resources involved in value creation” and the value outcomes as “the type of value outcomes customers perceive” (p. 19). The author distinguished between three creation processes:

- (a) Firm-created value refers to the resource firms possess and their internal processing of these resources into products/services.
- (b) Co-creation of value between consumers and providers. This notion is rooted in the premise that understanding value requires understanding the interface between a firm’s internal processes and external environment.
- (c) Customer value creation, wherein the value is seen as residing with consumers who determine what to do with the product/services (pp. 22-24).

In his study of customer perspectives on value creation, Priem (2007) stated that “customers are the arbiters of value” (p. 219) and emphasized the role of consumer experience in determining the benefits generated by a product or a service. He further contends that “consumers must be an important consideration in strategy formation because consumers experiencing benefits are essential to company success” and added that “value creation by a firm is a precondition for value capture” (p. 219).

2.4 OGD Value Creation

As the introduction discusses, open government data is hypothesized to generate economic and social value. However, academic research has not agreed on the value created and how.

Janssen et al. (2012) recognized the difficulty of measuring the value of open data. They attributed the problem to the fact that data “has no value in itself; it only becomes valuable when used” (p. 260). Stott (2014), on the other hand, maintains that the combination of the data with one or more other factors, such as an innovative idea, the inadequacy of existing services, the availability of new techniques to process data or new technologies for the delivery services, including the use of smartphones or even the Internet of Things (p. 8) can account, in part, for processes of economic value creation.

Data as a concept has often been referred to as raw material. Kitchin (2014) notes that data are commonly understood to be “the raw material produced by abstracting the word into categories, measures and other representational forms – numbers, characters, symbols, images, sounds, electromagnetic waves, bits – that constitute the building blocks from which information and knowledge are created” (p. 1). The World Economic Forum, likewise, defines data as “simple, raw and unorganized facts and consists of basic subscriber information (BSI), transactional data and content data. When this data is organized, processed and given a context, it can be termed information” (Thirani & Gupta, 2017).

The consensus among scholars that data has no value until it is exploited (Attard et al., 2016; Janssen et al., 2012; Kitchin, 2014) is the foundation this thesis builds on. Regarding the

economic value of data, it is only realized when data is transformed into products or services (Walker & Simperl, 2018). However, exploiting data and building products/services require investment in resources “to locate, extract and refine it before it yields value” (Deloitte Analytics, 2012, p. 1).

2.5 Business Model (BM)

Business models expanded in late 1990 as a result of the revolution in information and communication technologies (ICT) and the commercialization of the Internet that brought structural change to the way business operates, conducts transactions and reaches its customers (Charalabidis et al., 2018; Hanna, 2010; Magretta, 2002). Chambers & Patrocínio (2011) and Nielsen, Lund, & Bukh (2014) noted that despite common use in academic and practitioner domains, there is no consensus about nor broadly accepted definition of this term. Regardless of the definition used, value creation usually lies at the heart of this concept (Charalabidis et al., 2018). For example, Osterwalder and Pigneur (2010) describe the business model as “the blueprint of a company strategy to earn money” (p. 15). Chamber and Petrocino (2011) view it as “the logic of how an organization operates and creates value for its stakeholders” (p. 56); Kitsios et al. (2017) maintain that business models are “conceptual tools for characterizing how value is created for customers and how revenue is achieved” (p. 399). Magalhaes and Roseira (2017) echo the latter view suggesting that business models are “a strategic conceptual tool” (p. 2), as does Teece (2018), who describes business models as “an architecture for how a firm creates and delivers value to customers and the mechanisms employed to capture a share of that value” (p. 40).

The global expansion of open government data has created business opportunities for data intermediaries worldwide, giving rise to the emergence of new business models (Zeleti et al., 2014). Yu (2016) describes research investigating value creation and business model design for open data applications as being in its infancy. He suggested several components of business models in this domain, including actors and their roles, the value proposition, value creation and capture, assessment mechanisms, and activities and processes employed by firms (p. 81).

Despite OGD being a new phenomenon, many open government data business models emerged since the mid-1990s. Zeleti and Ojo (2017) reviewed the literature about open data business models and consolidated elements of the various business model frameworks. They proposed a business model framework for open data based on six values: value proposition, value-adding process, value network, value in return, value capture, and value management (pp. 224-226). Zott et al. (2011) found disagreement among scholars about what a business model represents, which they interpreted as indicating that the literature is developing in silos, in accord with individual interests (p. 1019). They did, however, identify four common themes: (a) the business model was, at the time of their study, emerging as a new unit of analysis, (b) business models emphasize a system-level, holistic approach to explaining how firms do business, (c) firms' activities play an essential role in the various conceptualizations of business models, and (d) business models seek to explain how value is created, not just how it is captured (p. 1019). Al-Debei and Avison (2010) examined 22 scholarly definitions of the business model to identify the primary components of the concept. From this, they established what they called a "unified business model ontological framework" based on four dimensions: (a) value proposition, (b) value network, (c) value finance, and

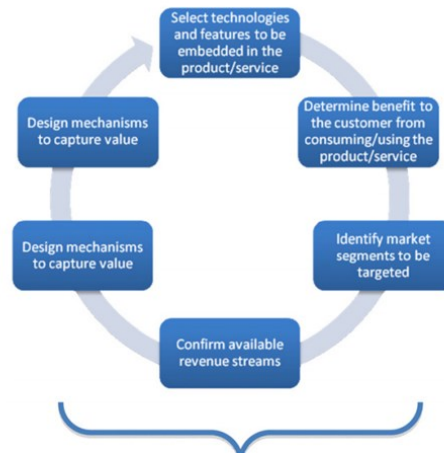
(d) value architecture (pp. 372-373). The various business models presented above suggest a critical link between business models and firms' economic value creation.

A critical business model component is the firm's value proposition to their customers. Yu (2016) found value proposition to be the core construct of a business model, and value creation to be the expected outcome (p. 81). Likewise, Magalhaes et al. (2014) considered the value proposition as "a statement of benefits that the firm delivers to its external constituencies" (p. 368).

Teece (2010) described the business model as a conceptual model that embodies a business's organizational and financial architecture in these words:

The business model makes implicit assumptions about customers, the behavior of revenues and costs, the changing nature of user needs, and likely competitor responses. It outlines the business logic required to earn a profit and, once adopted, defines how the enterprise goes to market. (Teece, 2010, p. 173)

Teece (2010) described the business model as a conceptual rather than a financial model. In his view, a business model "makes implicit assumptions about customers, the behavior of revenues and costs, the changing nature of user needs, and likely competitor responses" (Teece, 2010, pp.173-174). Figure 3 depicts the elements of the business model design.



Note: retrieved from Business Models, Business Strategy and Innovation (Teece, 2010, p. 173).

Figure 3: *Elements of a Business Model*

Business models are central to the operations and performance of firms and vital to their ability to create economic value and survive in the marketplace. In the next section, the attention turns to two commonly used frameworks for examining value creation by firms: the Resource-Based View (RBV) of a firm and Dynamic Capabilities.

2.6 Theoretical Frameworks and Methods

In the past decade, the phenomena of open government data and data intermediaries received the attention of scholars and practitioners who studied these phenomena and developed and proposed conceptual frameworks to analyze the use of open government data and the role of data intermediaries in the process of creating value from the data. The frameworks provided valuable insights but failed to offer a holistic theoretical framework that assesses the role of data intermediaries in the open government data ecosystem and discovers the functions they perform, the products/services they develop, and the economic value created through the use of OGD.

The Resource-Based View (RBV) of a firm and Dynamic Capabilities are perspectives rooted in strategic management. They provide conceptual frameworks to examine firm performance and how they seek to create economic value (Barreto, 2009; Eisenhardt & Martin, 2000b; Halawi et al., 2005; Kraaijenbrink et al., 2010; Zahra, 2021).

2.6.1 Resource-Based View (RBV)

The Resource-Based View (RBV) is concerned with firms' internal resources and capabilities to create economic value that enables them to achieve and sustain competitive advantage (Amit & Schoemaker, 1993; Barney, 1991a; Connor, 2002; L. Halawi & McCarthy, 2006; Ramon-Jeronimo, Florez-Lopez, & Araujo-Pinzon, 2019; Wernerfelt, 1984). RBV was first set out by Barney (1991) in a seminal article titled "Firm Resources and Sustained Competitive Advantage". Drawing influence from Penrose's (1959) and Wernerfelt's (1984) works, Barney suggested that firms obtain a sustained competitive advantage when implementing a strategy that exploits their internal strengths to respond to environmental opportunities while neutralizing external threats and avoiding internal weaknesses. For Barney, a firm's resources refer to "all assets, capabilities, organizational processes, attributes, information, and knowledge controlled by a company and classify them into physical, human, and organizational capital resources" (p. 101). Starting from the assumption that strategic resources are heterogeneous across firms, Barney (1991) posited that these resources had four characteristics commonly referred to by the acronym VRIN. The four characteristics are:

- (a) Valuable if the resources enable a firm to conceive or implement strategies that improve efficiency and effectiveness.
- (b) Rare if the resources are only possessed by few firms.

- (c) Inimitable if the resources are imperfectly imitable or difficult for other firms to obtain.
- (d) Non-substitutable, as there are few other strategically equivalent resources.

Barney considered the imperfectly imitable resources to be the results of the unique historical conditions, causal ambiguity, and social complexity on which a firm is based. In other words, he viewed the life and experience accumulated by firms as a determinant of their performance. In addition, the ambiguity in the link between a firm's resources and its sustained competitive advantage leads to imperfect imitability contributing to a sustained competitive advantage when not understood by potential competitor firms. In turn, Barney hypothesized social complexity to further contribute to the imperfect imitability of resources.

The RBV framework has been subject to criticism and calls for further development. Priem and Butler (2001) called it “not yet theory” and indicated that additional steps are necessary to become one (p. 57). The applicability of RBV to small businesses was questioned by Connor (2002), who argued that RBV only applies to large firms with significant market power. Connor also claimed that the usefulness of RBV was to be found in terms of its descriptive rather than explanatory nature, adding, “it does not equip strategists with practical competitive advantage-building propositions” (p. 312). As the RBV framework is about the importance of strategic resources in enabling firms to achieve sustained competitive advantage (SCA), Zahra (2021) questioned its applicability to entrepreneurial firms and startups who lack resources and experience in accessing and managing them. Furthermore, Eisenhardt and Martin concluded that traditional RBV misidentifies the locus of long-term competitive advantage in dynamic markets (Eisenhardt & Martin, 2000).

Following an empirical test of RBV, Newbert (2008) established that “improving performance is not directly a function of the value or rareness of a firm’s resource-capability combinations but rather the advantages it creates from its exploitation” (p. 762). This conclusion by Newbert emphasized the exploitation of resources rather than the quality of resources available.

RBV emphasizes that firms’ internal resources determine their ability to create value and achieve sustainable competitive advantage. It demonstrates how economic value creation can be achieved and the resources required to accomplish the task. This thesis focuses on how open data intermediaries create economic value from open government data. Hence, studying data intermediaries’ resources and capabilities is expected to provide insights into their strengths, weaknesses, and success.

2.6.2 Dynamic Capabilities (DC)

The Dynamic Capabilities (DC) perspective is an extension of RBV that was initially proposed by Teece et al. (1997) in their seminal article “Dynamic Capabilities and Strategic Management.” The authors defined DC as the “firm’s ability to integrate, build and reconfigure internal and external competencies to address rapidly changing environment” (p. 516). They explained that DC “reflects an organization’s ability to achieve new and innovative forms of a competitive advantage given path dependencies and market positions” (p. 516). For them, the term dynamic refers to “the capacity to renew competencies to achieve congruence with the changing business environment” (p. 515). They also associated “capabilities” with the role of strategic management: “the term ‘capabilities’ emphasizes the key role of strategic management in appropriately adapting, integrating, and reconfiguring internal and external organizational skills, resources, and functional

competencies to match the requirements of the changing environment” (p. 515). By contrast, Eisenhardt and Martin (2000) described dynamic capabilities as “a set of specific and identifiable processes such as product development, strategic decision making, and alliancing” (p. 1) that create value for firms within active markets through manipulating resources into new value-creating strategies. However, they emphasized that a firm’s competitive advantage does not come from capabilities but through resource configurations.

Teece et al. (1997) indicated that the competitive advantage of firms is dependent on three conditions: (a) managerial and organizational processes, which are the routines of current practice and learning at the firm; (b) specific asset position, which refers to the firm’s current state of technology, intellectual property, complementary assets, customer base, and external relations with suppliers and (c) the path available to the firm in the form of alternatives. In addition to the three conditions, the authors emphasized the importance of learning at the organizational and individual skills levels, reflected in the firm’s routines or new logic of the organization (p. 520).

Winter (2003) described dynamic capabilities as a construct that cannot be directly measured, touched or seen. Bowman & Ambrosini (2003) considered dynamic capabilities to focus not on the firm’s resources but on its ability to renew and alter them in line with environmental changes. Easterby-Smith et al. (2009) likened dynamic capabilities to RBV in its focus on core competencies and firm performance. They, however, noted that the difference between these two perspectives rests in the emphasis DC places on the dynamic of the market and changes in the business environment, something RBV did not consider. They further point out that DC is difficult to measure empirically, as are the underlying operational processes and the relationship between

DC and firm performance. Helfat et al. (2007) defined dynamic capabilities as “the capacity of an organization to purposefully create, extend or modify its resource base” (p. 1). In line with this definition, Ambrosini et al. (2009) stressed the role of managers in recognizing changes in their business environment and their responsibility to foresee and predict coming changes.

The concept and its definitions have been debated rigorously (Ambrosini et al., 2009; Eisenhardt & Martin, 2000a; Helfat et al., 2007; Teece, Pisano, & Shuen, 1997b; Winter, 2003) and literature on DC called for a standard definition that improves the acceptance and application of dynamic capabilities.

Zahra et al. (2006) argued that “the emergent literature on dynamic capabilities and their role in value creation is riddled with inconsistencies, overlapping definitions, and outright contradictions” (p. 917). Augier & Teece (2007) criticized the concept of dynamic capabilities because it was initially designed for large multinational firms without consideration for small firms. Finally, Ambrosini et al. (2009) argued that the impact of DC on firm performance may be negative, stating that “while the regenerative dynamic capabilities may allow a firm to change its dynamic capabilities, this does not ensure that the organization will be successful or even survive” (p. S20).

The primary objective of open data intermediary firms is to create economic value through alliances and product/service development. These two characteristics are among the components of dynamic capability organizational and strategic processes identified by Eisenhardt and Martin (2000) to enable the manipulation of resources into value creation strategy. Thus, these two

characteristics will be investigated to determine how they contribute to open data intermediary firms' successes and economic value creation.

Chapter 3. Methodology

This study aims to learn and understand the role open data intermediaries (ODIs) play in the open data ecosystem. Specifically, it seeks to discover how open government data (OGD) is used, what products and services are created, what resources are required, what challenges exist, and what value is created using OGD. The study employs inductive research to explore the works and experiences of ODIs concerning the OGD phenomenon. Accordingly, this study's design has selected a qualitative research approach based on grounded theory.

Corbin and Strauss (2012) defined qualitative research as “a process of examining and interpreting data to elicit meaning, gain understanding, and develop empirical knowledge” (p. 1). The authors advised researchers to use a qualitative research approach when two conditions are met: (a) when the research question dictates that a qualitative research approach be used, and (b) when researchers want to get at the inner experience of participants to determine how meanings are formed through and in culture and to discover rather than test variables (p. 12).

Qualitative research has many definitions; however, all characterize it as an interpretive methodology, inductive, seeking to explore things and collect data in their natural setting and oriented toward interpreting phenomena in terms of the meanings people bring to them. Furthermore, scholars agree that qualitative research aims to generate understanding from those being studied and promises to contribute significantly to the social entrepreneurship knowledge base and practice (J. Corbin & Strauss, 2012a; Creswell, 2007; Denzin & Lincoln, 2018).

Creswell (2007) described the qualitative research approach as one that “begins with assumptions, a worldview, and possible use of a theoretical lens” (p. 37). The author advised researchers to use a qualitative approach to inquiry, collect data in a natural setting, and perform inductive data analysis that establishes patterns and themes.

Because ODIs are a new phenomenon that has not been explored empirically, a qualitative research approach will be used to design the current study. Zeleti and Lehmann (2005) argued that emerging phenomena lack theories grounded in empirical data and emphasized the importance of selecting an appropriate research method when there is an issue with extant theories.

3.1 Research Design

Research designs are a type of inquiry within qualitative, quantitative, and mixed methods approaches that provide specific directions for procedures in research (Creswell, 2014). According to Glaser and Strauss (1967), the research process guides the researcher toward examining all possible rewarding avenues of understanding: “a research method is one of discovery and grounds a theory in reality” (Glaser & Strauss, 1967 as cited in Corbin & Strauss, 1990). The literature on qualitative research methods is rich. Creswell (2009) examined several classifications and typologies and selected five methods that reflected the types of qualitative research he saw in the social, behavioural, and health science literature: narrative, phenomenological, grounded theory, ethnographic, and case study research. This study employs a qualitative case study research approach guided by grounded theory principles. Case studies are used to develop an in-depth analysis of a case, whether a program, event, activity, process or one or more individuals. Cases are bound by

time and activities, and researchers collect detailed information using a variety of data collection procedures over a sustained period (Creswell, 2009).

ODIs and OGD are new phenomena that emerged during the second decade of the 21st century. Fernandez and Lehmann (2005) described emerging phenomena as “new, with little or no prior theoretical studies and/or frameworks that lack theory grounded on empirical data obtained from real participants in the substantive field of the phenomena” (p. 83).

Following a thorough examination of the qualitative methods discussed in the literature, the researcher selected a case study approach informed by grounded theory (GT), as defined by Corbin and Strauss (1990) and Glaser and Strauss (1967), as the overarching method to study data collected from case studies of ODIs. As Creswell (2013) noted “good design when a theory is not available to explain a process” (p. 88). According to Glaser (1978), grounded theory accepts qualitative, quantitative, and hybrid data collection from surveys, experiments, and case studies (Glaser, as cited in Fernández, 2004). Fernandez (2004) recommended that the researcher specify the method driving their investigation to avoid conflict between the grounded theory and case study method.

Fernandez and Lehmann (2005) considered the discovery aspect of the grounded theory method a critical factor in its relevance and success. They advised researchers to immerse themselves in data and follow a rigorous approach in searching for patterns, similitude, and contradictions to conceptualize what is going on in their field of interest. Grounded theory principles are utilized in the data collection and analysis within this qualitative case study research.

Note that in this thesis, although GT was used within the case study research design, the theory generation could not be entirely completed due to the low number of data point collected and analyzed, which results in part from the recency of OGD and a general lack of public knowledge.

3.2 Case Study Approach

The case study approach has been adopted as the method for data collection in this research. Like grounded theory, this approach uses interviews for data collection (Halaweh et al., 2008), as both the case study and grounded theory consider the interview the primary data source (J. Corbin & Strauss, 1990; Yin, 1994). This approach is one of Creswell's (2007) five widely used qualitative research approaches (see Section 3.1 for information about the other four approaches). The author has defined a case study as "a qualitative approach in which the investigator explores a bounded system or multiple bounded systems over time, through detailed, in-depth data collection involving multiple sources of information and reports a case description and case-based themes" (Creswell, 2007, p. 73).

The thesis author followed Yin's (1981) suggestion regarding case study selection as a method for data collection when two conditions are met: (a) an empirical inquiry examines a contemporary phenomenon in its real-life context, and (b) the boundaries between phenomenon and context are not clearly evident. The selection of case studies also followed Yin's (2003) four principles: (a) the focus of the study is to answer "how" and "why" questions, (b) the researcher cannot manipulate the participants' behaviour in the study, (c) the researcher wants to cover contextual

conditions because they believe they are relevant to the phenomenon under study, and (d) the boundaries are not clear between the phenomenon and context (Yin, as cited in Baxter & Jack, 2008).

OGD is a nascent phenomenon built on the premise that it will generate economic and social values. Data intermediaries are vital players in the OGD ecosystem. They use open data to create products/services that theoretically contribute to value creation. Exploring the complex relationship between the open data ecosystem, data intermediaries' objectives and motivations, and the capabilities and resources needed to achieve their objectives necessitated adopting the case study approach. Crowe et al. (2011) described a case study as "a research approach used to generate an in-depth, multi-faceted understanding of a complex issue in its real-life context" (p. 1). The authors argued that case studies are valuable when there is a need to obtain an in-depth appreciation of an issue, event or phenomenon of interest in its natural, real-life context (Crowe et al., 2011).

3.3 Case Selection

Eisenhardt (1989) argued that selecting cases is essential to building theory from case studies. In addition, the author emphasized the crucial concept of "population" as it defines the set of entities from which the research sample is drawn. Following Eisenhardt's recommendations, this researcher selected ODIs operating in Canada as the focus of this study to avoid any extraneous variation and help to define the limits for generalizing the findings.

The literature on case studies shows an agreement among scholars (Creswell, 2009; Eisenhardt, 1989; Yin, 2009) on the design of the case study. Emphasis is placed on the research question, which determines the type of case study, the boundary of the case, and the unit of analysis

(Baškarada, 2014; Baxter & Jack, 2008; Creswell, 2009; Crowe et al., 2011; Eisenhardt, 1989; Yin, 2003).

The research question in this study aims to explore the roles of open data intermediaries in the open data ecosystem. It is constructed as an overarching question with four sub-questions:

- Q. How do data intermediaries create value using open government data?
 - a. How are open data intermediaries using open government data?
 - b. What resources and capabilities are vital for open data intermediaries?
 - c. What are the challenges and barriers faced by open data intermediaries?
 - d. What values are created, and how?

The adverbs “how” and “what” are used in the questions to explore open data intermediaries’ motivations, objectives, and required resources and to understand their functions. Rowley (2002), referring to Yin (1994), stated that the how or why questions are asked about a contemporary set of events over which the investigator has little or no control” (p. 17). In the case of this thesis, the exploratory question and sub-questions led to the selection of the case study method for data collection for this research.

Scholars recommend setting boundaries to case studies to prevent researchers from going wide in their research and attempting to answer broad questions with multiple objectives. Time, event, place, activity, definition, and context boundaries are typically suggested (Baxter & Jack,

2008; Creswell, 2013; Miles & Huberman, 1994; Yin, 2003). Consequently, this researcher has chosen to study ODI firms in Canada and has limited the period studied to 2010–2020.

The unit of analysis in this study is the open data intermediary. A data intermediary is an individual or an organization that elects to use open data as a resource to develop products and services, anticipating creating economic and social value. Open data intermediary firms are the cases that will be studied. As intermediary firms operating in the open data ecosystem are developing, and research on their roles and functions is limited, the researcher opted to use a multi-case research approach instead of a single case. Yin (1994) argued that multiple-case designs should follow a replication, not a sampling, logic. The case studies should serve as multiple experiments, with similar results (a literal replication) or contrasting results (a theoretical replication) predicted explicitly at the outset of the investigation (Yin, 1994). Eisenhardt (1989) explained that “in replication logic, cases which confirm emergent relationships enhance confidence in the validity of the relationships; conversely, cases which disconfirm the relationships often can provide an opportunity to refine and extend the theory” (p.542). Applying that logic to this research will help to identify common patterns and themes contributing to theory building.

In deciding to use multiple case studies, this researcher considered the scholars’ recommendations, the studied phenomenon’s specificity, and the targeted stakeholders. The cases in this study are individuals/organizations representing a new phenomenon in an emerging initiative. The researcher was concerned that a single or a small number of cases would not provide a comprehensive view of the roles and functions of a data intermediary startup firm, so it was decided to use multiple cases.

Eisenhardt and Graebner (2007) justified using multiple cases because they create a more robust theory, as “the propositions are more deeply grounded in varied empirical evidence” (p. 27). Yin (2003) described multiple-case designs as having distinct advantages and disadvantages compared to single-case methods (p. 45). However, Yin determined that the evidence from multiple cases is often considered more compelling, and the overall study is more robust.

3.4 Data Collection

Yin (2009) listed six sources for data collection relevant to case studies: documentation, archival records, interviews, direct observations, participant observations, and physical artifacts (Yin, 2009). This research relied on interviews and documentation, where it existed, supplied by the informants and/or collected from the Internet.

The data collection phase was planned to start immediately after the researcher received the University of Ottawa Research Ethics Board’s (REB) approval for this study. Unfortunately, the date the certificate arrived coincided with the spread of the COVID-19 virus in Canada, which led provincial governments to take various measures, including lockdowns, to stop that spread. As a result, interviews planned to be in person were replaced by online interviews using available technology-based conference tools (Zoom, Cisco, and others). The lockdown also limited the possibility of conducting observations at the firms’ offices; accordingly, these planned activities were cancelled.

3.5 Identification of Participants

The open data intermediary firms selected for this research were startup organizations that had emerged during the past ten years. The search for organizations working with open data in Canada was complex and lengthy. The researcher first looked at the Government of Canada's data portal, which provides a database of applications developed using open data for the public sector¹⁸. However, the applications are mainly published by government agencies and a limited number of individuals and small businesses. Consequently, the information found was insufficient to populate the database of potential firms for this research. Therefore, the researcher conducted an Internet scan of startups in the Canadian open data ecosystem, which resulted in finding a website¹⁹ that included a list of Canadian companies (Open Data 150). The web page described Open Data 150 as a joint project between The GOVLAB and Canada's Open Data Exchange (ODX), supported by Open North, the Federal Economic Development Agency for Southern Ontario, and Thomson Reuters. The GOVLAB²⁰ is a non-profit, university-based action research lab whose mission is to improve people's lives by changing how the country is governed. The organization conducts mixed-methods research, builds civic technology, and undertakes coaching to strengthen the ability of institutions to make decisions and solve public problems more effectively and legitimately.

The ODX is a national initiative to strengthen Canada's technology brand and accelerate commercialization. It connects those who need data with those who have data, develops tools to share data, provides training and networking opportunities, and influences policy and standards. The establishment of the ODX organization was the third commitment in Canada's second action

¹⁸ <https://open.canada.ca/en/open-data>

¹⁹ <https://canada.opendata500.com/>

²⁰ <https://thegovlab.org/>

plan for 2014–2016. The action plan announced an investment of \$3 million for three years by the Government of Canada to launch the ODX²¹. The third Canada action plan, 2016–2018, included a commitment (No. 15) to stimulate innovation through the ODX. One of the milestones attached to this commitment was the completion of the comprehensive mapping of 150 Canadian companies using open data. Another milestone was launching an online platform at www.opendata500.com/ca to showcase Canada’s “Open Data 150.” The plan also announced the end of ODX as an independent organization and its relocation to be part of the Communitex Data Hub²². Open North²³ was described as a “Canadian not-for-profit social enterprise that specializes in the development of innovative online tools that promote greater government transparency and citizen participation” (Open Data 500: Canada, 2020).

The Open Data 150 website provided a list of companies using open data in 2016, along with their business categories, years of establishment, and locations. The list served as a first stage in identifying organizations for this study²⁴. Another Internet scan led to the Open Data Impact Map²⁵ website, a public database of organizations worldwide that use open data. The website listed 52 Canadian companies that use OGD. The Open Data Impact Map, described as “a project of the Open Data for Development Network (OD4D),” is a public database of organizations that uses open government data worldwide. The Open Data Impact Map is supported by funding from the International Development Research Centre (IDRC) and the World Bank. It is licensed under a

²¹ <https://www.opengovpartnership.org/members/canada/commitments/CA0032/>

²² <https://open.canada.ca/en/commitment/mtsar/2016-2018/commitment-15-stimulate-innovation-through-canadas-open-data-exchange-odx>

²³ <https://opennorth.ca/>

²⁴ <https://canada.opendata500.com/list-page.html>

²⁵ <https://opendataimpactmap.org/map>

Creative Commons Attribution-Share Alike 4.0 International License. (See Appendix F for a combined list of unverified ODX150 and Open Data Impact’s open data intermediaries.)

Merging the two databases and eliminating duplicates and organizations established before 2010 was the second phase of identifying firms for the case studies. The third phase involved verifying business categories and eliminating all non-profit organizations and organizations with unknown business functions. At the end of this phase, the researcher arrived at a list of 144 cases.

Phase four involved verifying the existence of the organizations, identifying senior executives and their contact details, and refining the population database further. To do this, the researcher relied on the companies’ URLs included in the ODX 150 list, the Open Data Impact Map list, Canada’s corporation database, provincial business registries, and other Internet websites (for a complete list, see Table 3). At the end of phase four, the population database stood at 76 startups.

Table 3: *List of Companies Used for Verifying Open Data Intermediaries’ Legal Status*

Organization Name	URL
PitchBook	https://pitchbook.com/
Crunchbase	https://www.crunchbase.com
ZoomInfo	https://www.zoominfo.com
The GOVLAB Open data 500	https://canada.opendata500.com/
Government of Canada	https://www.ic.gc.ca/
Ontario Business Registry	https://www.ontario.ca/page/ontario-business-registry
Government of Quebec	http://www.registreentreprises.gouv.qc.ca/en/consulter/rechercher/default.aspx
Velocity	https://velocityincubator.com/velocity-founders-and-companies/

Organization Name	URL
Communitech	http://news.communitech.ca/od150-first-out-of-the-gate/
RocketReach	https://rocketreach.co/terrene-inc-profile-file_b453446dfc816123
Bloomberg	https://www.bloomberg.com/profile/company/1516699D:CN
Facebook	https://www.facebook.com/
LinkedIn	https://www.linkedin.com/
OpenCorporates	https://opencorporates.com/

3.6 Theoretical Sampling

The data collection process in this study was preceded by a theoretical sampling of homogeneous cases of ODI firms assumed to use OGD to create products/services that generate economic value. Theoretical sampling has been described as follows:

The process of data collection for generating theory is whereby the analyst collects, codes, and analyzes his data and decides what data to collect next and where to find them to develop the theory as it emerges. (Glaser & Strauss, 1967, p. 45)

Urquhart (2017) considered theoretical sampling a critical element of the grounded theory method that needs to be considered for a theory-building design. Corbin and Strauss (2012) indicated that the purpose of theoretical sampling is “to collect data from places, people, and events that will maximize opportunities to develop concepts in terms of their properties and dimensions, uncover variations and identify relationships between concepts” (p. 143). Building on the definition and guidelines provided by Corbin and Strauss (2012), Urquhart (2017), and Glaser and

Strauss (2017), the theoretical sampling undertaken in this study was based on a homogeneous sample (Creswell, 2013) of 76 ODI firms assumed to be using OGD.

3.7 Ethical Considerations

This research followed the University of Ottawa Research Ethics Board guidelines and process as stipulated by the University of Ottawa policies and the Tri-Council Policy Statement (2018) For Research Involving Humans (Tri-Council, 2018).

As required, the researcher applied to the REB to acquire the appropriate certification before starting the data collection phase (see Appendix C for a copy of the certificate). The application included copies of the invitation letter and the consent form sent to identified participants. The invitation letter (Appendix B) introduced and described the project and its purpose. It informed the recipient of the voluntary nature of participation and other personnel at the organization. When a positive reply was received, a consent form was exchanged for a signature (Appendix A). The consent form informed the participant of the method of participation and what was expected, recounting his/her rights and the risks and benefits of participation, and confirming the participant's privacy and confidentiality of exchange information. Recruitment of participants began after signed consent forms were exchanged between the participants and the researcher.

If no response was received after the email invitation was repeated, the participant details were removed from the potential list to a "No answer" list. The researcher then selected another possible case from the list and repeated the recruitment process.

3.8 Selection and Recruitment of Participants

Upon receiving the REB certificate, the researcher started preparing to invite targeted firms to participate in the study. The first attempt to reach participants was an invitation letter sent by regular mail to a selected group of 10 organizations. When the researcher did not receive any response within three weeks, a change in the communication method was initiated. Electronic mail (email) was sent to invite each selected firm's executives. When the researcher received a positive reply to the invitation, the date and time for an interview were negotiated and set. Immediately after, a consent form was sent to the participant using the DocuSign electronic software, requesting his/her signature and return of the document. Upon receiving the signed consent form, the researcher signed it and sent an electronic copy to the informant. The participants were informed that the virtual interviews would be recorded and were given a choice to accept or decline. Participants were also told that their name and their organization's name would not appear in the research and would remain anonymous.

The database of potential participants included 76 startup firms across Canada, many of them in Ontario and Quebec, which happen to be the provinces with the most data intermediaries recorded. The recruiting process began in April 2020 and continued until January 2021. The first case selected as a pilot was based in Ottawa, ON. The intent of starting with a pilot case is to test and evaluate the interview questions, uncover any missing categories, and assess the researcher's interviewing capability and the feasibility of the online conference tools. According to Yin (2003), pilot case studies may reveal inadequacies in the initial design or help articulate it.

3.9 Pilot Case

The pilot case interview was with the SP06 firm (see Table 4), a small operation started as a university research project by one individual researcher who decided to commercialize the concept. The interview revealed that this early data intermediary did not start with open data, as those were not available publicly. In addition, the interview revealed that sticking to pre-prepared interview questions may not generate a complete picture of the subject under discussion. Therefore, an intervention on the researcher's part was needed to correct the impression and refocus the questions. Subsequently, the interview questions were adjusted to ensure the informant's understanding. Finally, the interview revealed a lack of awareness of OGD, and the support Canadian governments offer early-stage startup companies.

3.10 Interview Process

A semi-structured list of 10 questions was prepared to discover how and why data intermediaries got involved with OGD (see Appendix D for the interview protocol). The interview questions focused on understanding the research question and sub-questions of this study using a set of characteristics the research aimed to discover:

1. Perception and motivation
2. OGD influence
3. Product/service development
4. Knowledge and skills
5. Resources and competencies
6. Value creation
7. Barriers and challenges

The interview was designed to take 1 hour and to be followed by another session if the need to clarify specific points arose. In addition, interviewees were given the option to provide clarification by email. ODI interviewees were located across Canada. This broad geographic spread enriched the study by providing a general overview of the support services available to new startups in the corresponding provinces. The interviews were conducted using Zoom, Cisco, and Skype video conferencing. All interviews were recorded using the video conferencing tool and an external device to ensure the security and availability of the recorded materials.

Before each interview, the researcher reviewed the data intermediary website to learn about its executives, products, and services. Multiple internet searches were also performed to locate text, video, and audio records about the interviewee and the company. When found, audio and video recordings were collected, listened to, and studied before the interview and after to compare and add any missing information.

The researcher started each interview by introducing the purpose and objective of the study and repeating the interviewee's rights and privileges, as specified in the invitation letter and on the consent form. Next, the interviewee was asked to provide documents explaining his/her views, approaches, and philosophies regarding open data and value creation. Finally, the interviewee was asked to review the open government data initiative and its ecosystem at the end of the interview.

After the completion of every interview, transcription started. Once the researcher completed the interview transcript, a copy was sent to the interviewee with a request that they verify it

and add any missing information. The transcript was coded and encrypted before being sent to informants. A password to decrypt the document was communicated by a different method.

3.11 Data Coding and Analysis

The data analysis approach used in this study was based on Corbin and Strauss' (1990) three-stage approach to data coding: open coding, axial coding, and selective coding accompanied by constant comparison (see Figure 4). Creswell (2007) considered data analysis a process of preparing and organizing the data for analysis by reducing the data into themes through coding, condensing the code, and ending by representing the data in figures, tables, or a discussion (Creswell, 2007). Corbin and Strauss (1990) described data collection and analysis as interrelated processes that should begin as soon as the first bit of data is collected and be accompanied by comparative analysis and questions. Figure 4 illustrates the data collection and analysis process used in this thesis.

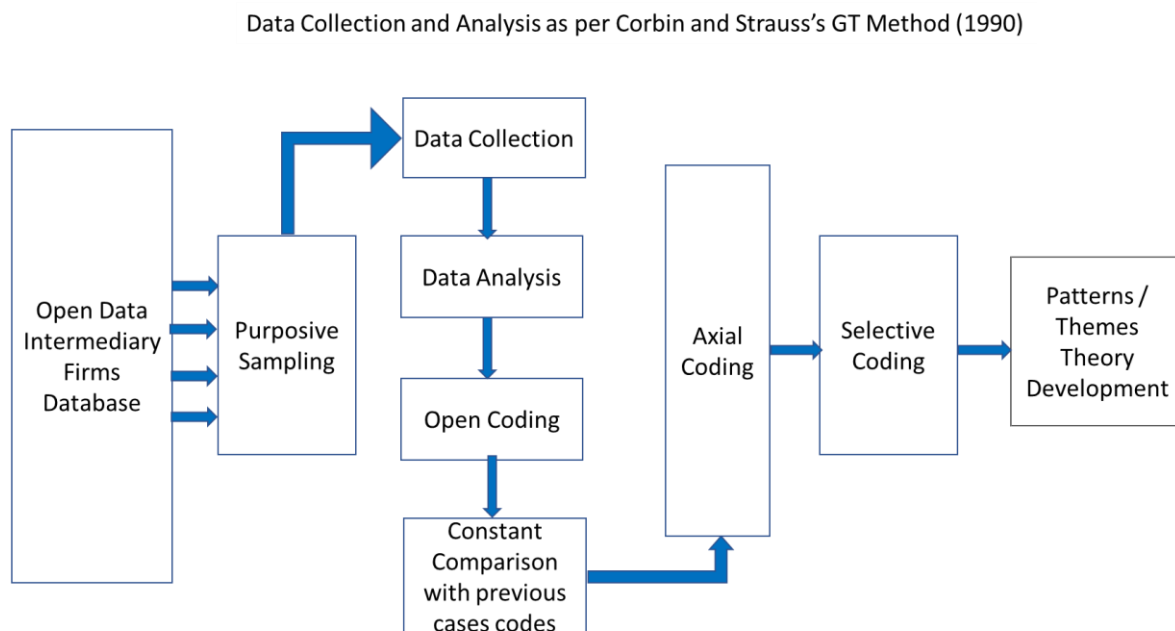


Figure 4: *Data Collection and Analysis Process*

3.11.1 Purposive Sampling

As this study is about a new and evolving phenomenon, open government data, which occurred in the last decade, triggering the emergence of a new type of users, open data intermediaries, the researcher chose to use the purposive sampling technique in selecting cases for this study. Purposeful sampling is a widely used concept in qualitative research. Patton (2002) described purposeful sampling as a technique commonly used to identify and select information-rich cases for the most effective use of limited resources. Furthermore, Creswell (2007) justified the choice of individuals and sites for study “because they can purposefully inform an understanding of the research problem and central phenomenon in the study” (p. 125).

When deciding on the number of cases to include in the research, the researcher considered Eisenhardt’s (1989) recommendation to select between 4 and 10 case studies while keeping in mind Creswell’s (2007) indication that no specific number is recommended. Considering that ODIs, the unit of analysis in this study, are new and their organizations are not large enough to inform the research and enhance the generalizability of the findings, the researcher purposefully selected ten such firms. The aim was to conduct 20–24 interviews. The case selection was based on the criteria that the organization should be (a) a startup; (b) commercially oriented; (c) operating in the data, technology, and service domains; and (d) based in Canada.

The above criteria were based on the newness of the phenomenon and the characteristics of users who chose OGD as the resource to create products and services that generate rents. They were also selected with firms developing products/services using information technology tools and software based in Canada in mind.

As the process of contacting targeted organizations proceeded, the researcher realized that obtaining more than one interview with each startup might not be possible in most cases due to the size of the startup and its current executives. The researcher then decided to purposely sample new firms and conduct interviews until saturation was achieved. Eisenhardt (1989) advised researchers to stop adding cases when theoretical saturation is reached. Saturation means that no additional data are being found, and the inquirer can then develop the properties of the category (Glaser & Strauss, 1967).

The concept of saturation is vaguely described in the literature. Bowen (2008) reflected on the notion of saturation by stating that “explicit guidelines for determining saturation are almost nonexistent in the literature on qualitative methodologies. Whether such guidelines are necessary remains open to debate” (p. 138). Glaser and Strauss (1967) discussed the theoretical saturation of the created categories, stating:

The criterion for judging when to stop sampling the different groups pertinent to a category is the category’s *theoretical saturation*. *Saturation* is achieved when no additional data are being found, whereby the sociologist can develop properties of the category. (p. 61)

The authors noted that theoretical saturation is reached by joint collection and analysis of data: “When one category is saturated, nothing remains but to go on to new groups for data on other categories and attempt to saturate these new categories also” (Glaser & Strauss, 1967, p. 61). The researcher applied the approach Glaser and Strauss (1967) recommended when identifying

categories within a dataset, using constant comparison to understand the relationship between them before creating a higher-level category. This process reached a saturation level when no additional categories were created.

The selection of potential ODI firm participants included consideration of similar business categories, different business categories, and atypical cases. As noted, the study was initially designed to examine ten firms. However, recruitment continued until the number of firms reached 17. These were coded with the first-level categories developed. The researcher considered that the identified categories had reached a saturation level when no new properties yielded additional ones.

The 17 firms came from different industries and provinces and provided diverse information about data intermediaries' motivations and their processes of capturing open data and developing products/services. Twenty semi-structured interviews were conducted with founders and/or company executives recommended by the founders. Additionally, the researcher collected publicly available documents (text, video, and audio) on the startups found on the startups' websites or the Internet or provided by the interviewees (see Table 4 for a list of the participating firms).

Table 4: *List of Participating Open Data Intermediaries*

Startup	Location	Type of Business
SP01	Montreal, QC	Data/consulting services
SP02	Fredericton, NB	Data/consulting services
SP03	Edmonton AB	Data/consulting services
SP04	National	Data/consulting services
SP05	Kitchener, ON	Information services
SP06	Ottawa, ON	Information services
SP07	Fredericton, NB	Solution services/consulting
SP08	Montreal, QC	Data services
SP09	Kitchener, ON	Data services
SP10	Toronto, ON	Data services
SP11	Vancouver, BC	Solution services/consulting
SP12	Vancouver, BC	Data services
SP13	Toronto, ON	Information services
SP14	Toronto, ON	Information services
SP15	Prince George, BC	Data services/consulting
SP16	Toronto, ON	Data services
SP17	Edmonton, AB	Solution services/consulting

3.11.2 Open Coding

Corbin and Strauss (1990) defined open coding as the interpretive process by which data are broken down analytically. According to the authors, in this process, events/actions/interactions are compared with others for similarities and differences and assigned conceptual labels in open coding. In the second stage, the conceptually similar events/actions/interactions are grouped into categories and subcategories, which become the basis for sampling on theoretical grounds (Corbin & Strauss, 1990).

Following Corbin and Strauss's (1990) suggested approach, the researcher began coding the transcripts line by line using NVivo software, analyzing every event, action, and interaction by

open data intermediary and creating initial categories. Initial categories are low-level concepts representing the processes taking place. The process continued sequentially after each interview, accompanied by constant comparison between the new concept identified and previously created concepts. When a similar action/interaction was found, it was assigned to an existing category. If no similarity was noted, a new category was created. The process continued until all interviews were coded and initial categories were established. The open coding process ended with a total of 1808 preliminary categories registered. Appendix G provides examples of the coding process based on the NVivo software application²⁶ and used in this thesis.

3.11.3 Axial Coding

Axial coding is the process of relating categories to their subcategories (Strauss & Corbin, 1998). The researcher commenced axial coding activities immediately after the open coding process. The latter was based on the initial open coding of identified categories and was intended to reduce the high volume of open coding into manageable, focused higher-level categories by linking categories at the level of properties and dimensions (Strauss & Corbin, 1998). The process involved comparing, analyzing, testing, and joining initial open coding categories to form a higher level of abstraction. As Creswell (2013) explained, “axial coding emerges in which the researcher identifies one open coding category to focus on (called the “core” phenomenon), and then goes back to the data and creates categories around this core phenomenon” (p. 86).

The axial coding process extensively reduced the number of categories to a more focused and manageable number. In addition, it permitted the theoretical properties of the categories to

²⁶ <https://lumivero.com/products/nvivo/>

emerge into themes related to the functions and roles of ODI. The axial coding process ended with 129 higher-level categories (see Appendix G).

3.11.4 Selective Coding

The final stage of selective coding is when the theory elements appear through integrating and refining processes. According to Glaser (1978), selective coding “is the stage at which coding is limited to only those categories that relate to the core category” (Glaser, as cited in Urquhart, 2017, p. 88). Corbin and Strauss (1990) defined selective coding as “the process by which all categories are unified around a ‘core’ category, and categories that need further explication are filled in with descriptive detail” (p. 424). Following Corbin and Strauss’s (1990) recommendations, questions were asked, and a more profound and broader analysis of higher-level categories was performed to understand each category’s dimensions. The researcher evaluated the higher-level categories established during the axial coding process, integrated, and consolidated the categories around the core theme of the study, and interrogated and analyzed the final categories in relation to the objective of the research and their possible conceptualization and contribution to the development of the theory. Following the integration and analysis, a storyline was developed about the roles and functions of open data intermediaries and the resources needed to ensure successful operations. Creswell and Brown (1992) indicated that the final stage is where the research finding can be conceptualized in different forms: a narrative statement, a visual picture, or a series of hypotheses or propositions. Following refinements and integration of similar categories, the final selected categories were reduced to nine (see Appendix G).

3.11.5 Constant Comparison

Constant comparison is a key tenet of the grounded theory procedures (Urquhart, 2017). The researcher used this method extensively at all coding levels to reduce redundancy and produce more focused higher-level categories related to the core phenomenon. The procedure of constant comparison consists of four stages: (a) comparing incidents applicable to each theme that emerges from the data, (b) integrating themes and their properties, (c) delimiting the theory, and (d) writing the theory (Glaser & Strauss, 1967). Corbin and Strauss (1990) suggested that constant comparisons enable investigators to break through subjectivity and bias and contended that the fracturing of data would help examine preconceived notions against the data themselves.

3.12 Trustworthiness

To ensure the validity of the findings, the researcher followed Creswell's (2014) recommendation to use multiple approaches. Therefore, a triangulation method (combining interviews, documents collected from participants, and information found on the Web) was used to determine the validity of the results generated from the data. In addition, data about categories from different sources were compared and analyzed to assess their agreement and disagreement. Finally, the constant comparison method in this study enabled the researcher to build coherent themes that represented the reality of the actions undertaken by ODIs.

The second method provided a profound description of the intermediaries' environment. Hence, a detailed description of ODIs' operations was delivered after a comprehensive validation of the intermediaries using multiple data sources. The researcher had no direct knowledge or experience with the participating ODIs or any stakeholders in the open data ecosystems that would

have affected his neutrality and objectivity. Furthermore, the researcher followed Creswell's (2014) advice on avoiding bias and ensuring the validity of the research. To guard against the influence of bias, the researcher kept a journal during all stages of the research project. To ensure proper understanding and correctness of the interviewees' data, the researcher sent a copy of the transcribed interviews to the interviewees for verification and confirmation.

3.13 Chapter Summary

This chapter reviewed the qualitative research approach and the selection of grounded theory and case studies as the data collection and analysis methods. Furthermore, grounded theory and case studies were reviewed and discussed, and their use was outlined and justified. The chapter also discussed the theoretical sampling of the participant firms, their fields of operation, the selection process, the number of cases selected, and the justification for that number. Finally, it discussed the data collection and analysis phase recommended in Corbin and Strauss's three stages model—open coding, axial coding, and selective coding—as well as the constant comparison method to build the corresponding categories.

Chapter 4. Findings

This chapter addresses the study's research questions and presents the work undertaken by Canadian ODIs. It examines the types of work performed, the ways OGD is used, the resources and competencies needed by ODIs to work with open government data, the products/services created, the value generated, and the challenges and barriers that impact ODIs' work.

4.1 Open Data Intermediaries Overview

The Canadian open data intermediaries participating in this research are startup firms that commenced operations during the past decade (2010–2020). The startup concept has garnered the attention of many scholars and practitioners, who have provided various definitions, each from different perspectives and with different dimensions (Blank, 2007; Csorba, 2020; Magalhães, 2019). For example, Kollmann et al. (2016) provided three characteristics of a startup: (a) is younger than ten years, (b) features highly innovative technologies and/or business models, and (c) has (strives for) significant employee and/or sales growth (p. 15). Writing for *Forbes Advisor* Baldrige & Curry (2021) defined startups as young companies rooted in innovation and founded to develop a unique product or service, bring it to market, and make it irresistible and irreplaceable for customers.

The founders of the startups participating in this research were individuals driven by their knowledge, experience, skills, and motivation to create social products/services to solve issues and problems and respond to the needs and opportunities found in the market. They relied on their resources and publicly available data (open and not) to establish their businesses and develop their offerings.

Eleven open data intermediary startups were created by two co-founders, while a single person established each of the other six. The interviews showed that one of the co-founders in the 11 startups had experience and knowledge in IT, data science, or business. Similarly, the single founders had a mix of experience in business entrepreneurship and technical knowledge and skills such as computer and data science, startup development, industry experience, and business development. One informant equated the impact of having a co-founder to “doubling the chance of success.”

The founders ranged from university dropouts to graduate students, researchers, public or private sector employees, open data enthusiasts, and technical and business entrepreneurs. Two ODI startups were driven by their social motivation to contribute to the communities’ social good by developing business applications that facilitated their target clients’ business operations. One startup, based on volunteer work, had no direct commercial objective. However, its services contribute indirectly to value creation at its clients’ operations. The rest of the ODI organizations were business-motivated, developing products/services and generating revenue. However, social value creation was also on their founders’ minds. Table 5 presents an overview of the ODI organizations.

Table 5: *Overview of ODI Organizations*

Case	Founders	Interviewee Position	Background	Establish Date	Type of Business	Target Market
SP01	Two co-founders	CEO & Technical Director	Experienced open data & open access	2010	Data services & consulting	NFP & private sector
SP02	Two co-founders	CEO & Director of Operations	Experienced high-tech startup	2016	Data services	Municipalities

Case	Founders	Interviewee Position	Background	Establish Date	Type of Business	Target Market
SP03	Two co-founders	President (Co-founder)	Research Data science	2008	Data services & consulting	Municipalities & private sector
SP04	Two co-founders	President (Co-founder)	Data science	2014	Consulting & data services	NFP & NGO
SP05	One founder	President (Founder)	Business	2015	Data services	Industrial energy sector
SP06	One founder	President (Founder)	University research	2010	Data services	Media monitoring sector
SP07	One founder (student)	CEO	Engineering school	2013	Data services	Municipalities, private sector
SP08	One founder	Director of Business Development	2D mapping services	2017	Data services 3D mapping	AEC, municipalities
SP09	Two co-founders	CEO	Private sector	2017	Digital mapping 3D mapping	Architectural engineering, construction
SP10	One founder	President	Government	2017	Digital mapping & data analytics	Real estate industries
SP11	One founder	President	Academic entrepreneurship	2010	Data services	Municipalities
SP12	Two co-founders	Owner (Founder)	University Private company	2014	Data services & consulting	Private & public sectors
SP13	Two co-founders	CEO (Co-Founder)	Private sector	2016	Data services	Travel industry
SP14	Two co-founders	President Co-Founder)	Government & information technology	2012	Data & information services	Travel industry
SP15	One founder	President	Private sector & government	2010	Data services & consulting	Private/NFP sector
SP16	Two co-founders	Technical Director	Open data/government	2014	Data services & consulting	Open to all public & private sectors
SP17	Two co-founders	CEO	Business & technical	2015	Data services & consulting	Municipalities & private companies

4.2 Open Data Intermediary Profiles

The 17 open data intermediaries were in different locations in the country and had different motivations and backgrounds. Table 6 provides information on each startup's location, motivation, and background.

Table 6: *ODIs' Locations, Motivations, and Backgrounds*

Startup	Location	Motivation	Background
SP01	Montreal, QC	Social/commercial	NFP/NGO/entrepreneur
SP02	Fredericton, NB	Commercial/social	Business/entrepreneur
SP03	Edmonton AB	Commercial	Academic research
SP04	National	Social	Private sector
SP05	Kitchener, ON	Commercial	Business
SP06	Ottawa, ON	Commercial	Academic research
SP07	Fredericton, NB	Commercial	Academic
SP08	Montreal, QC	Commercial	Business
SP09	Kitchener, ON	Commercial	Private sector
SP10	Toronto, ON	Commercial	Public service
SP11	Vancouver, BC	Social/commercial	Entrepreneur/public service
SP12	Vancouver, BC	Commercial	Academic
SP13	Toronto, ON	Commercial	Entrepreneur/public service
SP14	Toronto, ON	Commercial	Public service
SP15	Prince George, BC	Commercial	Public service
SP16	Toronto, ON	Commercial	Public service
SP17	Edmonton, AB	Commercial	Entrepreneur/public service

The business ventures of the 17 startups were based primarily on their founders' experiences, knowledge, skills, discoveries of opportunities/gaps in their markets, and personal philosophical beliefs. The ODIs' profiles, motivations, objectives, and methods of developing products/services from open data are presented below.

SP01's motivation to start its business was socially driven. The founder, an active community member interested in open access and open data, wanted to work in "the space of technology for a social mission". The company was established in 2010 to provide services to charities and not-for-profit (NFP) organizations through a data repository of funders in Canada intended to increase NFP efficiencies, save them time and effort, and improve their productivity. Additionally, the company expected the product/service to generate a financial return to sustain operations and contribute to its growth and expansion.

The founders bootstrapped the company, and software developers were hired in exchange for shares in the company instead of salaries due to difficulties attracting external financial investment. The informant attributed the unavailability of external investments to investors' lack of interest in organizations operating in the NFP sector and acknowledged that external investments would have made a difference in the company's growth and expansion.

Lack of funds affected the company's growth and chances of survival, which drove management to use their knowledge and data skills to offer consulting services to organizations inside and outside the NFP sector. At the same time, the company maintained the original commercial product, which continued to generate close to 50% of the annual revenue. The informant described the state of the company as follows:

I consider the company successful, but I do not consider the initial incarnation successful. Focusing on one product based on open data and client-based has not generated sufficient returns. Consequently, as we branch into other activities

related to open data and sometimes involve open data, yes, that has been very successful. (SP01)

The business model adopted by the company for its commercial product/service is based on information as a service offered through a subscription model. Potential clients choose from multiple options that provide different services with different price tags. The consulting services business model is based on direct interaction with potential clients and offering consultants' knowledge and expertise in open data to enhance existing products/services.

SP02 was a project-focused data intermediary firm established in 2016 to develop custom solutions for all levels of government in Canada and the US using open and non-open data. The company identified "communication" between governments, citizens, and businesses as a significant issue, especially when deciding on land use and transportation. To help communicate governments' decisions and people's understanding of the decisions, the company developed interactive predictive analytics tools, data visualization tools, and predictive models built on top of data.

Two co-founders with business experience and technical skills established the company to support local governments in communicating with their citizens. The company developed software tools that analyzed open and non-open data collected from its customers and shared results with the public. The informant described their sales approach as "finding real pain points" in potential customers' operations involving the government's ability to communicate complex issues to stakeholders. The solution was based on a predictive analytics tool that permits stakeholders to test and see the trade-off between different planning options and the costs associated with each option. The

informant described the solution in the following statement: “We show the basic principles in a way that most stakeholders can understand rapidly.”

The company’s business model is intended to optimize business processes through products and tools that enable citizens and other stakeholders to understand a government’s decisions and implications. It is based on a direct sales approach and interaction with government departments and agencies to understand their issues and objectives. The informant described the benefits of the direct sales approach as “ensuring a check is received upon signing the contract and before starting the project.” The direct sales approach ensures a continuous flow of revenue, eliminating the need for external investment and hiring resources until after contracts are signed. The informant revealed that the company’s primary business premise is linked to customers’ willingness to sign a check. “Getting a check, even a small check, before starting a project is a proof of need.”

The company acknowledged that its method and approaches had generated successful sales. However, it recognized that many of the activities performed are repetitive due to the uniqueness of each project. Therefore, it is considering a different business model based on software as a service (SaaS) for future product/service development. Meanwhile, the company generates satisfactory financial results and looks for “exponential growth” in the coming years. The informant described the company’s performance as a “growing success” and mentioned the ultimate objective of making the company a household name in its selected industry through the effective use of the Internet.

The company received financial support from government agencies that helped start and build the organization and first product. Since then, the sales method has ensured continuous income without exposing the company to higher expenses, enabling it to self-finance its growth and expansion without external investment.

SP03 was a research project at a local university working with municipalities and public agencies to solve problems through in-house solutions. One of its research partners suggested that its work had commercial value and would interest many organizations like his. Consequently, with a colleague and support from the university, the researcher responsible for the lab created the startup to offer their services to public organizations. The two co-founders developed a software product that used municipalities' data in addition to open data available on the web to perform analysis and provide their customers with data visualization and forecasting tools to enhance their decision-making. The informant described the company's business and experience:

We are a team of data scientists and information designers who help the customer decide, engage, and convince with data. We have deep experience in predictive, spatial, and prescriptive methods. We are adept at machine learning, statistics, and econometrics and own the adoption, and we make sure the analysis is useful and used. (SP03)

The business model used to promote and sell the software product/service is based on the SaaS model. It is considered a door opener to additional consulting services the company offers its clients. As a result, the product generates one-third of the company revenue, while the

consulting services generate the other two-thirds. The informant described the company's services: "We build intuitive, interactive tools to explore complex data. We specialize in dynamic mapping, web dashboards, and immersive installations."

The informant acknowledged the importance of open data for the company's solutions. He attributed the company's success to people's technical skills, knowledge, and open data availability: "It would be difficult, and we would not be nearly as successful without access to open data. So it is important to our business."

The company initially received support from the university through space to operate and small grants at the starting phase. However, it is presently self-funding its growth and has not yet decided on taking external investment. The informant acknowledged that investors had approached him and his co-founder but had declined their offers.

The success experienced with the core business at municipalities allowed the company to grow beyond the municipalities' base into the private sector, which led to an increase in the company workforce to 20 people. The informant expected the company to double or triple in size in the coming years through geographical expansion in Canada and the US.

SP04 was a national NFP organization, the brainchild of a data scientist who is socially motivated to give back to the community. With a co-founder and support from the data scientist community, the company was federally incorporated in 2014 to provide data services to NFPs and NGOs nationwide. In addition, the founder started a crowdfunding campaign to finance its

operation to generate funds that garnered excellent support from their friends, family, and professional network. The informant described the company's mission and objectives:

We are a collective of passionate socially minded people who empower social changemakers to be better by allowing their data to speak for them. We help NFPs and NGOs harness the power of their data to make more informed and better decisions in their quest to make their communities flourish. (SP04)

The organization grew to nine chapters around the country, with the interested population of volunteers reaching 8,500 individuals. The informant described the organization's volunteers as "heavily quantitative-focused statisticians with computer science backgrounds who tend to be higher-quality communicators, visualizers, and presenters of quantitative information."

Open data was considered an essential factor in the startup's services targeted to the organizations. The lack of quality data at the serviced organizations justified using open data. The informant described the benefits to NFP and NGO organizations in maintaining, cleaning, preparing, validating, testing, and working on high-quality data to optimize their operations. Moreover, the provided services were not standard but designed around NFP and NGO needs and requirements and aimed at helping the organizations expand their social impacts through data.

SP04 does not charge for its services and expects no revenue in return. However, it believes that the provided free services contribute to social value creation and indirectly to economic value. The organization operates virtually and is a heavy user of the Internet. All volunteers are gainfully

employed, and their services are provided outside their official working hours and on weekends. Additionally, the organization aims to become a self-sustainable and going concern.

SP05 was founded on the premise of increasing transparency in the energy sector. Complex contractual issues related to energy that faced a family farm led to the company's creation by a family member in 2017. The company's main objective is to help organizations control their energy usage and costs: "The fundamental issue was an unnecessarily complex energy market that made it extremely difficult for average businesses and consumers to make educated decisions."

The company is built on open and non-open data from an independent government agency and the web. The initial software product was created using external software developers' services while relying on the founder's funds. The company leveraged low-cost/free online tools to develop its software tools and product and preserve its limited funds. It used OGD to create a predictive algorithm that assists industrial energy consumers in planning their energy use. The predictions generated by the software help the energy market operate more efficiently declared the informant, who stressed the critical role of open data in the process and the company's existence: "The company in its current form would not exist without the data. A huge part of our business's value to our partners and customers is giving them access to data they would otherwise struggle to access."

The business model used to market and sell the product/service is based on the SaaS framework. It gives potential clients a choice between a primary service based on a freemium model and a premium model that uses a more rigorous modelling technique to develop the predictions available on a subscription basis. The freemium model is a business model in which a company offers

basic or limited features to users at no cost. Under the premium model, a service provider charges a premium for supplemental or advanced features.

The company reported profits until 2020, when the COVID-19 pandemic began and caused lockdowns and business closures. The revenue the company had generated was invested in growing its workforce to a strong team of 22 people, with a plan to double/triple its growth in the next three years. Furthermore, the company's success was recognized in the capital market and led to investment from venture capital and support from an accelerator in its operating region.

SP06 was built on the belief that the need for data and information will explode in the information society, which will require new types of organizations catering to people's needs and private- and public-sector organizations. The company was established in 2010 with the support of a co-founder to capture open and non-open data from global sources, extract insights and knowledge from the data, and sell the information to interested customers. The company used advanced technologies to identify, collect, and analyze data. The informant described the firm's target customers as "organizations that want to know what is happening around the world in their industries based on their company name, competitors and other factors" and named three business sectors in which the company operates: (a) media monitoring, (b) government monitoring, and (c) technology services and consulting.

The company started in the basement of the founder's house with one computer in 2010 and has since grown into more than 70 cloud servers that process about 3 billion documents a month, generating about 1.2 million unique, fully processed documents every day, according to

the founder. However, the growth in the business did not translate into organizational development. The informant acknowledged that the company's potential is more significant than the results it has generated and admitted that his conservative approach to growth and expansion was to blame:

As a company, we could do better, so we are exploring a new business model. However, we are still in the starting phase. If we had a clicking business model, we could take it to the market and generate some heat, but I would not be running this company. (SP06)

The company, a technology-driven organization, has built a good name, and its presence on the Internet is noticeable. The products/services it offers can be classified as “information and knowledge” generated from collecting and analyzing billions of documents. The informant considers the product/service provided to the company's customers as a saver of time and resources that offers them valuable insights at a lower cost. The business model used by the company to generate revenue is based on the information as a service model, in which clients connect online to the company's website and subscribe to the service they choose. The company also licenses access to its data portal to a limited number of clients and offers consulting services.

SP07 was founded by a dropout undergraduate student who saw an opportunity in small municipalities' limited resources and decided to intervene by developing application software that enhances municipalities' efficient management of parking spaces and increases their revenues using a digital product. The company secured funds and grants at the starting phase to build the

necessary infrastructure and hired the technical people to develop the software by relying on data provided by municipalities and collected from other government agencies. The informant described the data collected from the various government agencies as “open data” and explained that cities have difficulty creating open data portals for financial reasons: “Data is not something you can make money off on its own. Municipalities need to wrap it around a better service, and small municipalities do not have the resources.”

The company uses the SaaS business model to sell its products/services and offer optimization and customization services to meet clients’ requirements. It has been growing in terms of clients and human resources; 75 clients from around the country have subscribed to the service, including airports, universities, hospitals, private operators, and municipalities. As a result, the workforce has grown from one person to 22 in the same period. In addition, the informant expressed the desire to expand into the US, where the company has excellent potential to grow.

The company’s growth is self-financed using revenue generated from the sales of its products/services. It is based on the founder’s business approach of adding more people as more customers sign in, thus avoiding exposing the company to financial difficulties. The informant preferred natural growth over growing using external investment, leading to a loss of control over the company.

SP08 was established in 2017 as a spinoff of an existing company. It is based on the founder’s vision to offer 3D mapping data for the real estate industry. The founder invested a large sum of money in the new startup, which initially relied on the mother company’s technical

resources, expertise, and network of business contacts in the public and private sectors. The company's main objective is to offer the highest-quality mapping data to architects, engineers, and construction (AEC) experts, municipalities, and provincial and federal governments.

The company invested heavily in building the technological infrastructure and developing its software. It relied on data captured by its workforce and integrated with available open data from local, provincial, and federal governments. It also created a mobile unit to capture high-definition geographic data, processed and rendered in 3D using artificial intelligence (AI) software to detect specific objects in its geospatial datasets and map them. Meanwhile, two products were created: raw data for use in software and systems and a data repository to provide access to structured 3D data through the company's web-based services. The structured data is intended for applications in the smart-city, autonomous driving, telecom, and powerline industries. A large loan, government incentives, and revenue from the company's small jobs financed the product development cost. The informant acknowledged receiving significant financial support from all levels of government.

We received many subventions from different levels of government, mainly from the federal government, and we do some smaller projects implementation in cities that are coming through various research centres. But, still, our primary source of income is the federal government. (SP08)

The company's business model is based on a direct selling approach that offers optimization services on a project-by-project basis. However, it plans to provide its products/services online

on a subscription basis, allowing potential clients to select and subscribe to the service that fits their needs. In addition, the company is eyeing the growth in autonomous driving vehicles, and smart-city applications are considered a natural market for its high-quality 3D data.

SP09 was built to provide easy access to mapping data. It was founded in 2017 by an engineer with experience and expertise accessing open mapping data from different sources. The difficulty of finding and accessing the data highlighted inefficiency in the process that led the engineer to establish the company to offer better and more efficient access to available open data. With the assistance of an experienced software developer co-founder, the company was established to develop products/services that give people with no training in geographic information systems (GIS) or computer-aided design (CAD) software access to mapping data, allowing them to interact with it and leverage it for communication. The informant described the state of the online mapping data and the work the company does:

There are many excellent open data online and various open data products, but there is no centralization or coordinated efforts to standardize across all platforms. What does that mean for people like us? It requires much work to go and find the information we need to start any project. (SP09)

The company started at a hackathon, where the founders built a pivot that allowed them to win ten thousand dollars to establish the company. After that, they applied for and received a grant from the local government and an accelerator. The developed product is an easy-to-use CAD software application accessible online and requires no expertise. It targets industrial and construction

engineering spaces but is not limited to them. The informant described the product as “a tool kit for designing and leveraging digital earth data. It broadly applies to other industries such as environmental service, landscape architecture, and many others that have to get their hands on digital earth data.”

The company has also built a data portal where customers can access mapping data worldwide. The business model it adopted to sell its products/services are based on SaaS, in which interested companies pay for monthly subscriptions. However, the company’s product/service sales have not generated the revenue necessary to sustain and grow the operation. In addition, the informant acknowledged many challenges in securing government funds. Therefore, the company is considering accepting investments from the capital market.

SP10 was a startup company targeting the real estate industry. It was founded by an engineer motivated by the problems he encountered while working at one of the government agencies. The engineer quit his job and created a solution that simplified engineers’ work in the real estate industry, reduced the time needed to find and inspect planning and construction regulations, and made efficient decisions. The product/service developed was based on open data collected from municipalities and, to a lesser degree, from provincial and federal governments. The company has since developed an algorithm to collect specific data and clean, process, and visualize it to provide information that helps developers and planners in the real estate industry make informed decisions. The informant indicated that the solution slashes the time needed by a real estate developer to inspect planning and construction regulations from 1 or 2 weeks to a few minutes:

We pull planning data and visualize them onto a Google Maps platform where the planner and developer can see all these different colours. The various dots represent other developments going on. So, in a few seconds, a developer who wants to build a condo can understand real estate or have location intelligence, how far away from transit, or the area's designation. (SP10)

The company was bootstrapped by the founder and their family and friends. The founder also received support from local incubators and accelerators to build the company and the solution. In addition, the informant revealed that the developed product/service received the approval and backing of a large Canadian real estate company and is marketed as information as a service, giving potential users/customers the ability to access the information online from its data portal. Nevertheless, the company has yet to generate the necessary revenue to grow its workforce; therefore, it is considering accepting seed capital to launch it in the US.

SP11 focused on enhancing people and government communications. The firm identified a trust issue between the government and the people as an obstacle to effective and efficient consultation. It developed a location-based citizen engagement platform to solve the issues by consulting people geo-facially. The motivation for such a project came about when the founder, a geographer with an ambition to solve problems in the public sector, was involved in a project requiring public support: "When I could not find a solution, I began analyzing how to create one."

The company did not start with open data because open datasets did not exist. However, open data usage evolved as more governments adopted the solution. Financing to establish the

company and create the product/service came from grants and funds from a government agency. The informant explained that external investors were not welcomed because they contradicted the founder's interests: "Investors were more interested in mining the database of people rather than promoting the social good the company is aiming at."

The company's business model is based on SaaS and a direct approach to potential customers. Although the products/services have received customer approval, they have not yet generated enough revenue to reach the break-even point, affecting growth and limiting the company's workforce to 6 people. The slow growth has been attributed to the company's focus on the public sector, which is considered slow in accepting and applying innovation. The informant also expressed frustration with governments, noting that, on the one hand, they provided financial support to innovators to develop and create innovative products/services and, on the other hand, that they failed to buy the products/services innovators made.

SP12 was created by two co-founders who saw potential in the availability and use of open data. The company's main objective was helping organizations communicate better with visualized data, as the founders believed that many organizations do not have the in-house capabilities and capacities to build the application they need: "Companies lack expertise either with sourcing the data or finding, cleaning, analyzing, or communicating the data." Therefore, the firm aimed to fill the gap by offering its data science competencies to assist potential customers in extracting insights from their data, optimizing and customizing services, and enhancing decision-making. "We help clients with data communication in general, but specifically in most cases, it is a

dynamic, custom-built web-based visualization. The data sources we use are often open data,” the informant declared.

The company has been entirely bootstrapped. Its primary markets are the private sector and NFP organizations. After losing the co-founder, the ODI continued its operation from the home of the remaining founder using a business model based on face-to-face business development with targeted clients to discover opportunities, propose services, and sign contracts. Although the revenue generated by this ODI is satisfactory to its founder, who indicated it has been profitable for six years, the potential to grow beyond the one-person operation is constricted by the founder’s philosophy and perspectives on the future.

SP13 was started in 2014 by two global travellers facing visa issues travelling from one country to another. Their personal experience influenced their decision to create the company. The founders’ main objective was to build a product/service to help people understand visa requirements and provide them with the knowledge and tools to apply for a visa. The informant indicated that the product/service was created using data scraped from governments’ websites. The captured data were considered open data: “In terms of open data, we would not be able to do what we did without openness that most governments have had with information about that entry requirements.”

The product/service developed is a tool that uses advanced technology to automatically checks thousands of different sources hourly, daily, and weekly, collecting the latest information and data on visa requirements and updating the visa database. The development of the

product/service was supported by grants from an incubator and accelerator that helped sustain its growth. The informant described the support received from the incubator and accelerator as essential to the company's success and expressed appreciation for the government's support of startups in Ontario: "I am more than happy with government financial support. When you are at the idea stage, it is tough to get outside funding; that is where the government helped us."

The company's target market covers the travel industry, including individuals, airlines, and travel agencies. The company uses a data portal that provides travelers with information about visa requirements and its visa services. The informant described the product/service as follows:

It is a set of application software that enables companies to provide basic non-personal information about a passenger and receive information about the documents the passenger requires for his/her trip. Once they know what documents a person needs for his/her trip, some additional personal information about the traveller is collected, and the system returns an electronic travel visa or travel authorization reference number. (SP13)

The company is now profitable and experiencing a steady revenue increase from services provided to travellers and airlines. Its success has attracted venture and angel capital, providing financial support to grow the business and expand its workforce, now 20 people. The company uses two business models: information as a service and product as a service provided over the Internet. A third model is being considered for providing access to the insights from the data services and the data portal to airlines and travel operators.

SP14 was established in 2011 to inform travellers about health issues and problems worldwide. The organization was started by an ex-government employee who was a team member that developed the world's first digital disease surveillance system for the Canadian government. The founder's motivation was to contribute to safe travel by informing travellers about health conditions, problems, and worldwide issues.

The company created the original version of the system using machine learning algorithms and natural language processing to detect threats or disruptions reported in the news or shared on social media. Data were collected, stored in a database, and analyzed to extract information about health conditions in the corresponding country. For example, the COVID-19 pandemic and related travel restrictions provided open datasets the company used to inform travellers and airlines about regulations and requirements in a specific country.

Initially, the company's target customers were travellers. However, as the concept was tested and its usefulness confirmed, the company also focused on travel insurance companies, airlines, human resource departments, and security departments at international corporations. The business model to promote and sell the service is based on information as a service through online access to the company's data portal. Individual travellers are offered a subscription-based service that allows them to download the software on their mobile devices and access the necessary data. In contrast, business organizations are offered customized services that respond to their needs and requirements.

The company was bootstrapped by its founders, who managed to sustain the operation, grow it, and record profit. Access to an accelerator service recently provided the company with the needed support to enhance its profitability and expand its workforce to 13 people. Additionally, through accelerator support, the company increased its network in the travel industry and explored additional services to international travellers.

According to the informant, **SP15** started ten years ago with the objective of “putting maps on the web, not on paper.” The founder, an experienced coder, software developer, and product creator from the resource sector, saw the opportunity in the geospatial market. As a result, he started his own business to provide geospatial expertise to technology companies in and outside Canada. Although open data was not the original resource to develop the initial product/service, it became part of the core ecosystem. The company pulled data from different sources and curated, aggregated, and analyzed it to provide answers to questions asked by its clients. The informant described the work the company does:

One of our key jobs these days is data aggregation and curation to ensure that the data we capture is valuable. First, it is the best we can capture for a particular client’s question. Then, we augment that data with other sources and create an opinion on some phenomenon. (SP15)

The business model used by the company was based on the concept of “build against revenue.” Thus, the company identified potential customers first, initiated, and established contact with them, and negotiated contracts based on custom-made services in the geospatial space. The

informant described the geospatial business sector as “a very hot sector; people want to build maps, understand geography, want answers on locations, and understand the relationship between location and whatever their business does.”

In addition to the contractual service, the company developed a standard analytical product marketed and sold on a subscription basis using the SaaS business model. It now also provides consulting in the geospatial arena, where its expertise in communication, needs assessments, project management, and technology guidance/selection is used. The company’s original target market was mainly technology companies; however, it has provided services to NFP organizations and large enterprises over the years.

The company is self-financed and has been profitable since the beginning, enabling it to expand geographically and grow its workforce to 25 employees. In addition, employees work remotely from locations around the country, eliminating the need for a central headquarters.

SP16 was an open data repository startup created by an open data enthusiast who had experience working for the federal government on the open data initiative. The founder’s motivation came from the explosion of open data published by governments and the lack of a central repository to hold all the diverse datasets. With the support of a co-founder, the company built a portal to fill the void in the open data space. The developed repository aggregates open data from various open data portals and offer a one-stop location to people and organizations looking for comprehensive datasets. Besides the open data, the company expanded its services to respond to customers’ requests, including proprietary data. The company website describes the company’s activities as follows:

We are creating the fastest, most efficient solutions to the problems of data variety. Our platform allows our clients to access and collaborate on standardized data, blend, and enrich their data, and integrate the finished product into any analytics or business intelligence solution. We are the bridge between people with data and those who can use it to power decision science—a standard gateway for organizations to power their decision science with clean, accessible data. (SP16)

Access to the company's open datasets is free of charge to everyone. However, revenue is generated from custom-made services provided to global customers on top of open data. The informant described how the income is generated: "Most of our revenue comes from the data products we build by combining open data into much more valuable products. It combines multiple datasets, custom fields added, transformation and building a pipeline to keep it updated for the client."

The company has been supported by venture capital, which has provided the funds for growth and expansion. Recently, the company announced 8 million dollars in series A funding, which indicates capital market approval of the business model behind its success and growth, as demonstrated by its profitability and the expansion of its workforce to 40 people operating from offices in Canada and the UK.

SP17 was started due to the local government's inability to communicate efficiently with the public. In response, two individuals founded the company to develop a software tool that

visualizes and disseminates data. The informant described the motivation behind starting the company:

We started the business because we noticed an issue with local governments providing information, whether the demographics such as population, a report such as labour force dataset, employment rates, or information on the housing market that is useful to all stakeholders. (SP17)

The two founders bootstrapped the company while building on the intellectual resources, business, data, and software development skills they had to create the product and services based on dashboard software that communicated information related to local governments. For example, the dashboard included demographics and labour force data, housing data, utility rate data, weather data, and many others. The product was promoted and sold directly to municipalities in Canada and the US and proved its viability through its acceptance by 500 local governments. Building on the success of its flagship product, the company is expanding into more advanced solutions customized to municipalities' needs and requirements.

Governments' open data is the primary resource for the startup's products/services. The informant described this data as essential to the company's existence and performance: "The role open data played in starting our business is essentially everything. We would not have started the company without open data. It would not be possible." The company created more than 38,000 premade community profiles for municipalities in Canada and the US using open data. The solution is now sold using the SaaS business model and is available for two tiers of subscriptions. One

is targeted at cities, and the other is made for regions. The product includes many options customers can select from using a pricing model built to satisfy their requirements and the service provided. In addition to the standard product, the company offers customizations, data analysis, optimization, and predictive analytics consulting services.

The informant acknowledged receiving financial support on several occasions that helped sustain and grow the company's workforce and product. However, its profitability increased its hesitancy to accept external investment due to the founders' fear of losing control over the company. The informant justified the company's hesitancy to accept external investment from investors who lack knowledge of government technology, where the sales cycle is longer than in the private sector and noted the company's preference for organic growth.

4.3 Open Data Intermediaries and Open Government Data

Analysis of the interviews indicated that OGD was not on the radar of early data intermediaries that began their operations in the 2010s. Instead, these data intermediaries used available data from government agencies, private organizations and/or the web. They considered the data collected or captured as open data because it was given to them by the government or was freely available on the Internet:

- In the early days, it was not as structured. There were many patchworks, duct tape, and things to try and get together. There was no concept of open government data, open datasets, or anything like this today. (SP14)

- When I started in 2012, we were underwhelmed with open data even though there was a lot of momentum and positive energy in the space. We were underwhelmed because people were still fighting around usability and making the data machine-readable. So, it was like this is an inspiring open data movement with lots of energy, and we sure like what people are trying to do. (SP04)
- We would consider the data we are getting as open data because the city will give us the information. (SP07)

As the federal government began to release government data in an open, machine-readable format and provincial and local governments followed suit, OGD use increased. OGD played an essential role in the activities of early and new open data intermediaries. Established ODIs integrated the data into their platforms or initially relied on it entirely for their products/services. Informants acknowledged the importance of open data in their offerings. For example, the SP01 founder described the influence of open government data on the company business: “It played a huge factor in our ability to do many projects. Open data allows a specific set of opportunities to be pursued because many things need enough data to answer questions.” Informant SP04 stated:

Thank goodness for open data and the evolution of that whole space because then we can find our data-poor customers some relevant data to use. I am a bigger believer today in open data than ever, but I know it is not the only data source.
(SP04)

Data captured by ODIs were classified as either structured or unstructured. Structured data were captured from governments' open data portals²⁷, and unstructured data were provided by private customers, city departments, and/or agencies that do not have open data portals. Analysis of the interviews revealed that ODIs' use of OGD varied depending on their solutions. ODIs, which worked with local governments, relied on data from government departments and agencies, which did not meet the definition of open data. ODIs that worked with provincial governments relied on open and non-open data captured from regional data portals or received from government departments and agencies. Other ODIs developed solutions addressing a national or global issue using open data from the federal open data portals and other open and non-open data portals worldwide. For example, one company working in the mapping and 3D domain discovered that OGD was insufficient to develop the solution. Therefore, it decided to create the needed datasets from independent fieldwork.

All informants reported different issues and problems working with governments' open data. One of the critical issues identified was access to the data. When the data was found and accessed, ODIs discovered various problems in the quality of the data, the format used, the versioning, the timeliness of the data, and the lack of standards among different levels of government and between various departments or municipalities. These issues and problems hampered ODIs' efforts and affected their performances (see Appendix E for a list of challenges and problems encountered by ODI startups). For example, the informant at SP15 described the data as follows:

²⁷ See <https://open.canada.ca/en/open-data>; <https://data.ontario.ca/>; <https://open.toronto.ca/>; <https://opendata.vancouver.ca/pages/home/>; <https://www.donneesquebec.ca/>; <https://open.alberta.ca/opendata>.

It is almost nearly useless unless you have a team of data scientists to harness the data meaningfully and identify what things are comparable. Data is never good; it is various shades of wrong. It is always a little bit not correct, the quality is poor, the format is not suitable, or the scale is a little bit off, which is the nature of using open data products. (SP15)

SP09 described its experience working with open data as follows:

There are many excellent open data online and various open data products. However, there is no centralization or coordinated efforts to standardize across all platforms. Therefore, it requires much work to go and find the information we need to start any project. (SP09)

ODIs considered OGD essential resources for developing their products/services but found it insufficient to depend entirely on them. Instead, these resources were used jointly with proprietary and web-based open data to create the final products/services and respond to customers' requests. The availability and access to OGD helped several ODIs establish their operations and develop their businesses and customer bases. The roles OGD played in the creation and development of ODIs' operations are reflected in the following statements:

- It was incredibly critical during the first few years; it is an important tool and a necessary ingredient. (SP01)

- It would have been difficult, and we would not be nearly as successful without access to open data. It is important to our business. (SP03)
- We would not exist without this data. (SP05)
- It was a huge factor in the beginning. (SP01)

Three ODIs pointed out the unavailability of specific sets of geospatial and mapping data from the Ontario and Quebec, and federal data portals, which affected their performance and forced them to look at other sources to capture the required data. However, ODIs expressed appreciation for the availability of OGD:

- The government data has been the easiest way to demonstrate why our platform provides value. So, I think opening the data was a significant and excellent first step. (SP09)
- We would not be able to do what we do without the kind of openness that most governments have had with information about those entry requirements. (SP13)
- As an ecosystem, we could accomplish much more if there was more open data across all industries. (SP14)

The use of OGD varied among open data intermediaries, who indicated that the selection and use of OGD were influenced by the gaps, opportunities, and issues they identified and not by attraction to open data. They knew what they wanted; they focused their attention and established their objectives first, then went out and captured the required data. However, there was an agreement among ODIs that data alone does not have value or generate any return in its raw format. For

data to be valuable, it must be exploited in product/service development before being offered to consumers and generating economic value.

Different objectives and visions were found among ODIs. Consequently, their approaches to open data, the business models used, and the markets they targeted defined their potential and levels of success. Table 7 below lists ODIs' products/services, target markets, and business models used to promote and sell their products/services.

Table 7: *Target Markets, Products/Services, and Business Models*

ODI	Target Market	Product/Services	Business Model	Availability
SP01	NFP & Commercial	Data services & consulting	Information as a service	Canada
SP02	Municipal government	Data visualization & communication	Direct sales/project oriented	Canada & US
SP03	Municipal government & the private sector	Data visualization & communication	Direct sales/project oriented	Canada & US
SP04	NGO & NFP	Data services	Freemium	Canada
SP05	Energy consumers & providers	Predictive analytics services	Software as a service (SaaS)	Canada & US
SP06	Information seekers (globally)	Media intelligence	Information as a service	Worldwide
SP07	Parking operators (public & private)	Ready-made parking management system	SaaS	Canada
SP08	Municipal government	3D mapping services	SaaS/project-based	Canada
SP09	Architectural, engineering, & construction industries	Mapping software services	SaaS	Canada
SP10	Real estate industry	Map intelligence platform	SaaS	Canada & US
SP11	Public sector	Civic engagement service	SaaS	Canada

ODI	Target Market	Product/Services	Business Model	Availability
SP12	No specific industry	Custom-made web-based visualization services	Direct sales/project oriented	British Columbia
SP13	Airline & travel space	Visa information and processing services	Information as a service	Worldwide
SP14	Travel industry	Travel protection and insurance	Information as a service	Worldwide
SP15	Commercial & NFP sectors	Geospatial data services	Direct sales/project oriented	Canada & US
SP16	All industries	Open data aggregator	Data as a service	Worldwide
SP17	Municipal government	Ready-made municipality profiles	SaaS	Canada & US

As reported in the interviews, analysis of the objectives and motivations of the 17 ODIs revealed three types of ODIs. Each type approached its markets with different perspectives and different products/services:

1. Custom-Made Solution Builder

Five ODIs, equipped with knowledge and skills in data science, provided customized solutions and services to their customers. Four of the five ODIs' business models were based on direct sales through interest creation, identification of issues, and contract negotiation based on clients' needs and requirements. The fifth ODI was a volunteer-based organization that offered free-of-charge services to NFPs and NGOs. Data for the projects combine customers' provided data and open and non-open data captured from available data portals or the web. The outcomes of ODI interventions were a mix of software solutions, answers to questions, insights and recommendations prepared per customers' requirements. The solutions vary by the customer; however, they can be characterized as optimization services to enhance decision-making, improve

communications, save time, and reduce cost. Three ODIs focused on the public/private sectors, the fourth targeted the commercial industry, and the fifth focused entirely on NFPs and NGOs.

2. Data/Information Service Provider

Nine ODIs captured open and non-open data and created data portals with specific datasets to cater to the needs of their targeted customers. Moreover, they developed web-based application software and tools to enable customers to access and use the information provided on their portals. The business models used by this category of ODIs are based on information as a service and data as a service. To promote their products and services, ODIs used multiple channels, including online marketing, direct sales, and partnering with industry consultants.

3. Independent Software Developer

Three ODIs focused on building application software that addressed specific issues or opportunities they uncovered and found to be of interest to local governments and the public. They developed the software application before involving their target clients. Data needed to populate the applications were captured from local government departments and agencies, open data portals, and other public sources. ODIs used a direct sales approach and SaaS business model to promote and sell their products/services.

In addition to the products offered, 12 ODIs provided consulting services to their clients based on their knowledge and competencies in data and data science. As a result, the revenue of these ODIs was found to be generated from a mix of products/services sales and consulting services provided. Figure 5 shows the distribution of ODIs per category.

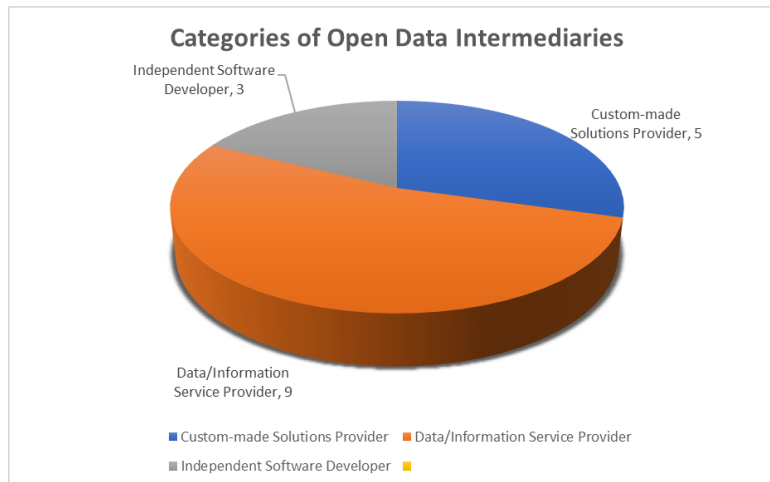


Figure 5: *ODI Distribution by Category*

4.4 How Are Open Government Data Used?

The use of OGD by ODIs varied according to the business objective, the targeted industry, and the users. However, analysis of the interviews revealed similarities and differences in how OGD was used to develop the products/services intended. Figure 6 presents a view of the products/services produced by ODIs.

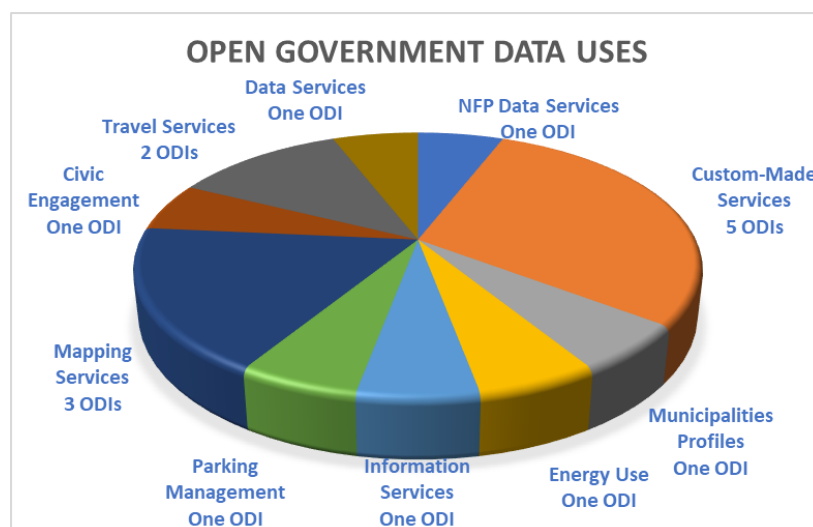


Figure 6: *Open Government Data-based Products/Services*

4.4.1 Custom-Made Solutions Builder

The custom-made solutions builder category consisted of five ODIs working directly with organizations in the public, private, and NFP sectors. They offered their knowledge and technical skills in data science as resources to help customers use their data. The ODIs approached OGD from their clients' perspectives and focused on their objectives and needs. They provided the know-how and technical skills, used clients' proprietary data, and captured additional data from open data portals to develop the solution requested by their clients. Creating solutions was a multistage process involving data capturing, then data curation and structuring before incorporating any captured external open and non-open data to produce the final product.

The informant at SP02 described its products and services' value to its customers as "breaking down barriers around communicating complex ideas to a wider audience." The one at SP03 explained how governments' departments and agencies use the company's services: "We help our clients communicate more effectively. We help them find insights from their data and turn them into impact."

SP04, a volunteer-driven national organization, focused on optimizing its partners' operations, marketing, donor, and financial strategies. The organization described its services: "finding cost and donor efficiencies, identifying optimization gaps, analyzing sponsorship dollars' flow to find cost gaps that could be leveraged, determining donor trend analysis, and discovering campaign optimization." Similarly, the informant at SP12 characterized the services provided in terms of knowledge and skills missing from customers' organizations. The informant described customers as "lacking expertise with sourcing the data, like scraping or finding, cleaning, analyzing or

communicating it.” They noted that the company helped build custom-made applications to visualize the data.

SP15’s extensive expertise in the geospatial domain was aimed at the private technology sector interested in location intelligence and geospatial data, mapping, and cartography. The ODI was established to provide services clients cannot obtain with their resources, including sourcing and analysis of geospatial data, location intelligence, and building maps to answer questions and requests by clients. In their interview, the informant summarized the work performed by the company as “data sourcing, data aggregation, curation and creation of opinion on some kind of a phenomenon.”

4.4.2 Data/Information Service Provider

This category consisted of nine ODIs that sourced and collected open and non-open data from various and diverse sites, including governments’ open data portals, government departments and agencies, governments’ websites, the Internet, and other online media. A common feature among members of this category was their use of advanced technologies, such as machine learning and artificial intelligence, to search, find and capture the data needed. The data was then curated, consolidated, and integrated with open datasets to form the core of the data portals on which they built the products/services offered to individuals and public and private organizations.

SP01 captured its data from the federal data portal to populate a centralized database targeted at NFPs. The service provided allowed NFPs and charities to access a comprehensive database of funders that improved their performance and increased the efficiencies of their operations.

In addition, the ODI built on its knowledge and experience with open data and data, in general, to offer data consulting services to clients inside and outside the NFP sector.

SP05 collected its open data from an independent government organization operating in the energy sector. It combined the energy data with weather data to build an interactive predictive analytic software that forecasts energy use, allowing industrial users to plan optimum energy use and reduce expenses. In addition, insights captured from the data were made available to large energy users through a simple application programming interface (API), allowing them to make informed decisions to lower energy costs and decrease emissions.

SP06 built its products/services on open and non-open data collected worldwide on business and non-business-related subjects. The company used advanced technologies to monitor media and government and gather business intelligence on specific topics worldwide. It captured and curated the data to populate a portal that generates insights and reports per clients' requests. In addition, it scraped non-open data from websites worldwide to build databases of specific topics of interest to its customers.

Four ODIs in this category, SP08, SP09, SP10, and SP15, focused on capturing geospatial data from governments' open data portals, websites, and the Internet. They also captured non-open data related to targeted industries and curated and integrated them with open datasets to create industry-specific data portals.

SP08, working in the geospatial and mapping space, developed its own sets of 3D geospatial data through groundwork like Google's data capturing method. It then integrated the data with other open geospatial data captured from government websites and the Internet. The final product/service, a high 3D mapping service, is available to AEC experts, municipalities, and provincial and federal governments. It permits them to geo-reference real estate on a map. In addition, the company's knowledge and experience with geospatial data and mapping enabled it to benefit from contracts with local government departments and agencies.

SP09 and SP10 provided similar services but to different industries. The two companies collected open geospatial data from governments' portals, the Internet, and private providers, integrated the data with industry-specific data and built software tools to enable customers to access the database to get insights and answers to their questions.

SP09 focused on the industrial engineering and construction industries, collected and consolidated data required by the industrial engineering and construction professionals in a central data portal, and developed an easy-to-use software tool to access the data. The informant described the objective of the company and the product they built as follows:

We built an online mapping platform to help professionals solve two problems: the inexistence of search engines for maps and enabling people who do not have training in GIS or CAD software to interact with that data and leverage it for communication. (SP09)

The service's primary purpose was to enable professionals to access relevant data in a centralized data portal. As a result, they can use it to find insights, facilitate their work, save time, and increase their productivity.

SP10 focused on the real estate industry, which was found to have efficiency problems in obtaining complete information about real estate properties. Therefore, the company built a real estate data intelligence platform that provides users with information on the risks and opportunities for a given parcel. Non-open and open data were used to populate the portal, providing clients with fast access to the regulations and planning requirements. In addition to the portal, the company developed application software that visualizes data on a Google-like platform showing all the regulations and planning requirements for a specific property, thus exposing all risks associated with the property. This approach enhances customers' knowledge and enables efficient decision-making.

Two ODIs, SP13 and SP14, focused on the travel industry and developed solutions that used open and non-open data to support travellers. The target customers were individuals and organizations in the travel industry worldwide. Data was captured from governments' websites, the Internet, and online media to create a central repository targeted at specific needs. In addition, the two ODIs developed software tools to provide online services to clients worldwide, including information about visa requirements, travel restrictions, threats and health conditions, insurance protection, and local health facilities and services in the countries clients are planning to visit.

SP13 offered its online clients, individuals, and organizations the ability to subscribe to its services on visa requirements in all countries. In addition, it used the knowledge captured from the data to help individuals and travel organizations learn about visa conditions and requirements worldwide and assist them in applying for electronic visas using an online software tool developed for this purpose. However, the data captured from governments' websites were not always open, so it had to be curated before being used and integrated into the data portal database.

SP14 focused on travellers' health. The founders were motivated by the absence of a global system that notified travellers about risks and dangers existing in the world. They, therefore, decided to build a "warning system" data portal using advanced technology and algorithms to collect specific health data from countries worldwide. The company captured data from governments' websites, scraped them from the Internet and various online media, and prepared them for analysis and extraction of information. During the COVID-19 pandemic, the ODI began to capture open data and integrate it into the database. The insights and information obtained from the database alerted travellers to health issues and conditions in countries worldwide. Later, the service was expanded to include information on local health facilities and providers.

SP16 focused solely on open datasets in the starting phase. The company consolidated open datasets from federal, provincial, and local government data portals, locally and internationally, in a central portal to offer its clients one-stop services that provided them with sought-after data. Access to the data portal was offered free of charge to all people. The company also decided to provide fee-paid customized services to organizations worldwide to generate revenue. Upon client requests, the company began to source and capture non-open data and blend them with its open

datasets to answer customers' requests. It also assisted customers in combining their proprietary data with open datasets to add value.

4.4.3 Independent Software Developer

This category consisted of three ODIs that developed solutions to specific opportunities in the market using governments' open and non-open data. Their initial target markets were municipalities that provided the data needed to populate their applications.

SP07 discovered an opportunity in local governments' inability to manage their parking spaces and payment methods efficiently. The company built a standard software to manage parking spaces and payments by mobile phone. The product/service was sold to local governments. In addition, customization and optimization services were offered to meet clients' needs. The software was populated with the municipalities' data and data from provincial government agencies. Following success in the public sector, the software was made available to private operators, including universities, hospitals, airports, and private parking operators. Access to the software product is offered based on an online subscription or through application licensing.

SP11, a socially motivated ODI, built a location-based citizen engagement platform to enable efficient communication between governments and citizens. Data used to populate the platform was captured from multiple sources, curated, structured, organized, and geocoded to serve as the core of the data portals. The data was later segmented geographically and indexed by areas. In addition to the data portal, the company developed a software tool that enables government

departments and agencies to have fast and efficient communication and consultation with their constituents.

SP17 identified a weakness in local governments' communications with the public as a significant problem facing municipalities:

We started the business because we noticed an issue with local governments providing information, whether the demographics such as population, a report such as labor force dataset, employment rates, or information on the housing market that is useful to all stakeholders. (SP17)

The company built visualization software that broadcasts various municipal's information to the public online. The software is a dashboard communicating different datasets representing communities' profiles. It was populated with OGD provided by the municipalities and other data sourced from government departments and agencies such as Statistics Canada. The product is now offered as-is or customized to meet the requirements of each city.

In summary, OGD contributed to the products/services developed by ODIs but was not solely responsible for the value generated. ODIs used whatever government data they found, open and non-open, to solve the issues they identified. Those that started operations in early 2010 first worked with non-open data, as OGD was still debated. However, as governments began opening and publishing their data in an open format, that data was integrated into the companies' resources. Conversely, those ODI firms established after 2014 began with OGD, initially their sole data

resource. However, their dependence on OGD alone was short-lived, as they discovered the need to integrate non-open data to answer customers' requests and develop solutions.

4.5 What Are the Resources and Capabilities Needed to Create Value from OGD?

This section reviews the resources used by ODIs to build their operations, create products/services, develop their markets, sell their products/services, and generate revenue. Data analysis revealed three critical resources needed to build operations and conduct business: data, human and financial. Additionally, the researcher discovered a valuable resource that facilitated the development of startups, one represented by incubators and accelerators.

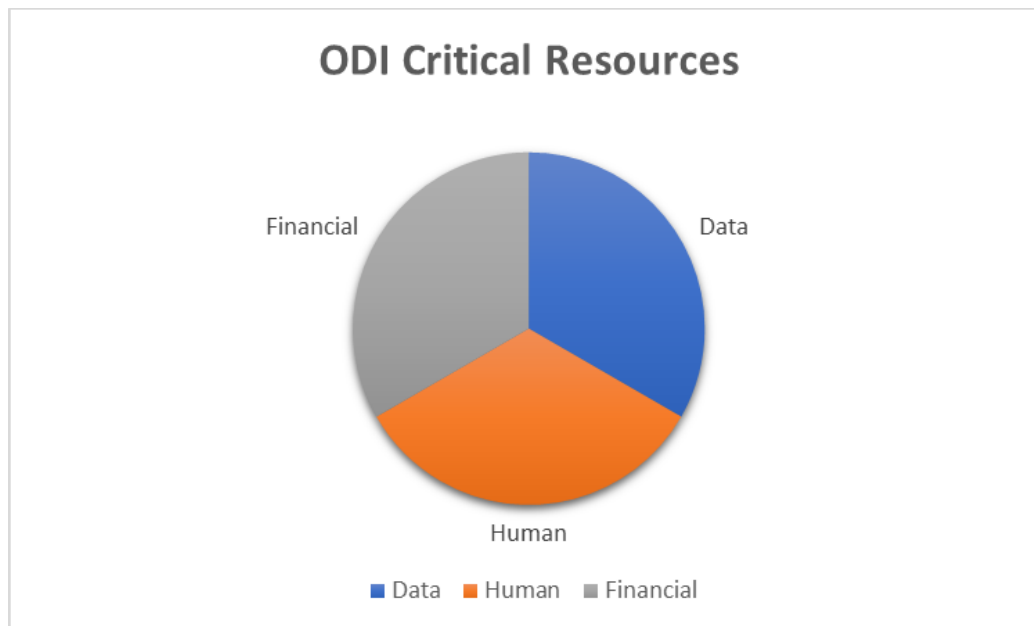


Figure 7: *Critical ODI Resources*

The following sections discuss the identified critical resources from the perspective of each ODI.

4.5.1 Data Resources

Finding data to build the products/services were not a significant problem faced by ODI firms. On the contrary, it was the motivator behind the appearance of this new type of intermediary, the ODI. The explosion in the volume of data made available by the pervasiveness of the Internet and Web 2.0 technologies and the emergence of social media applications in the early 21st century produced the perspective that data was a valuable resource with huge potential (Hanna, 2010; Magalhaes & Roseira, 2017). Furthermore, the drive to digitize and automate public and private organizations' operations accelerated in the mid-90s with the commercialization of the Internet. It led to new business practices based on electronic business and electronic commerce. These new approaches to conducting online business created a large volume of digital data in computer storage devices waiting to be used again. Moreover, adopting and implementing OGD initiatives by democratic governments worldwide led to increased digital data available for public use.

The adoption and implementation of local OGD by the 78 democratic countries, and members of the Open Government Partnership (OGP), led to enormous growth in the availability of open government datasets. Governments' open datasets were published according to a standard defined by the Open Knowledge Foundation (OKF) and adopted by the OGP. Starting in 2011, Canada commenced publishing its data in machine-readable formats for use by the public²⁸. The Canada open data portal²⁹ listed more than 35,000 open datasets available to the public for use and reuse as of January 2023.

²⁸ <http://www.tbs-sct.gc.ca/media/nr-cp/2011/0317a-eng.asp>

²⁹ <https://open.canada.ca/en/open-data>

Before the government release of open datasets, Canadian data intermediaries were active data users, as indicated by this study. However, they used non-open data furnished by their clients and collected from websites, including newspapers, blogs, press releases, and governments' PDF documents, to extract insights and information and respond to customers' requests. As open datasets were made available, they became part of the resources used by ODIs, which integrated these sets into their products/services. Commenting on how they started their operations and the role OGD played in their decisions, informants from Canadian ODIs offered the following statements:

The company got started ten years ago. It was not explicitly because of open data but certainly was related. The basic information we were using at the time was not open data. We lobbied to make it open two years later. It was public information, publicly available, but we had to get it from the government using other sources because there was no open data policy at that point. (SP01)

We started with the issues and worked on the solution. We could get data from different departments for some of the projects with the city, but for one project, we got almost all the data we needed from their open data portal. It was surprising how much we could do with the freely available data. (SP02)

I created the original evolution of the system using machine learning algorithms and natural language processing to detect threats or disruptions from news and

social media. If we can include access to news information, Twitter, and social media in the open data concept, that is our company's foundation. (SP14)

Most of the data we needed for the type of work we were doing was proprietary and owned by companies or governments. (SP03)

Data collected directly from government departments and agencies were non-open, which needed cleaning, linking, aggregating, and structuring. Statistics Canada and the Canada Revenue Agency (CRA) were among the many agencies named as data sources. However, the ODI firms in this study that worked in the geospatial and GIS domain depended on municipal and open data found on the Internet or scraped from governments' websites.

The use of open datasets by the data intermediary firms that participated in this study increased as the federal government instituted the policy of open-by-default and as provinces, regions, and municipalities created data portals and published their open datasets. Therefore, new ODIs entering the market after the open government initiative was put into practice initially depended on governments' structured open datasets, as SP10 and SP16 informants noted:

The company started in 2014. It was essentially begun because more and more governments were publishing open data, but they were publishing it on their public or private portals. That posed challenges to those trying to plug into open data portals. So, the co-founder created a data management platform that could act as a single repository to connect to all these dispersed open data. (SP16)

We mainly pull from open data websites from all five levels of government. Primarily we draw from the lower-tier cities, Toronto, the upper tier, like the York region, the conservation authority and ministry of natural resources, and the Ontario government. We can also pull from PDF, vector line, or tracer. (SP10)

The informant from SP09 also described the data they use and where their company gets it: “We pull from Ontario, where massive sets of LIDAR information like 3D topographic data are highly accurate.”

Data sources were not limited to Canada. ODIs whose products/services are available to customers worldwide cited the US as having more high-quality open data than Canada and captured open and non-open government data from governments’ websites and portals. For example, the informant from SP13 described the source of the data their company uses:

Regarding open data, we would not be able to do what we do without most governments’ openness with information about entry requirements. We check thousands of sources hourly, daily, and weekly, whatever makes the most sense, to ensure we have up-to-date information. (SP13)

The SP08 informant, citing the poor quality of available government open data, indicated that the company had established its Google-like data collection arm and gathered the 3D data needed for its target market:

The company's main objective was to offer AEC experts, municipalities, and provincial and federal governments the highest quality data from LIDAR data in high 3D data and point clouds. Consequently, we decided to capture the data ourselves because the available data were of minimum quality, and what we wanted was unavailable. (SP08)

The availability of and access to open data, specifically those from smaller municipalities, were not straightforward. ODI firms participating in this study indicated they faced many challenges and barriers in finding and accessing needed data. Analysis of the interviews revealed that smaller municipalities lacked the human and financial resources to implement the open government data initiative. ODIs had to request the data from municipal departments and agencies and wait for it to be delivered. Moreover, the provided data were mostly unstructured and had to be cleaned, linked, and structured before use. It should be noted that OGD is more readily available from larger municipalities and federal data portals.

The findings also revealed a severe misunderstanding of OGD and where to find it among ODIs (13 out of the 17 interviewed), which perceived data on the internet or government websites or received directly from government departments as open data. This inaccurate perception made their claim that their products/services were based on open data. The lack or shortage of knowledge of what constitutes "open data" exposes a problem in the communication between governments and the public, which will impact the use of open data if it is not addressed.

4.5.2 Human Resources (HR)

Skilled human resources (HR) professionals are critical to new firms' development and growth; firms with limited HR capacities may not develop their products/services and expand their market presences efficiently (Cardon et al., 2004; Stevens & Campion, 1994). Analysis of the interviews revealed that each of the 17 ODIs approached the subject of human resources from a different perspective. Several recognized the importance of skilled HR personnel as a prerequisite for growth. Others managed to hire as needed when they required specific technical or managerial expertise. For a few, the issue was not necessary because their business models centred around the knowledge and skills of their founders. Eight of the ODIs had workforces of between one and six employees. Four had a workforce of between 10 and 18 employees, and five had between 20 and 40 employees. The workforce size does not accurately indicate the situation of smaller ODIs that could grow their workforces and markets but have elected to limit their exposure for various reasons, including financial and philosophical perceptions.

The researcher's analysis uncovered two skills needed by the sampled ODIs: data science and marketing. Data science skills are essential to work with data. Hence, ODIs who were short of this capability recruited data scientists for their organizations. At a particular time, three ODIs used outsourced data science services on a project-by-project basis. It should be noted that many ODIs had personnel with the technical skills and proficiencies to work with data. The data scientist profession covers a large spectrum of skills. IBM (2021) has defined the data scientist skill set as follows:

- Apply mathematics, statistics, and the scientific method.
- Use a wide range of tools and techniques for evaluating and preparing data.

- Extract insights from data using predictive analytics and artificial intelligence (AI), including machine learning and deep learning models.
- Write applications that automate data processing and calculations.
- Tell—and illustrate—stories that convey the meaning of results to decision-makers and stakeholders at every level of technical knowledge and understanding.
- Explain how these results can be used to solve business problems. (IBM, 2021)

The ODI firms in this study engaged in many activities that required different skills depending on the products/services they aimed to develop. The skills needed were application software development, application design, knowledge and experience in computer languages and platforms, data analytics, and artificial intelligence and algorithm development skills.

The second skill set identified by the researcher was marketing and sales. This skill was missing from nine ODI firms due to the founders' technical backgrounds. The SP01 informant provided a common view of the problems faced by technically competent data intermediaries:

We had many implementation questions where we did not know what we were doing with marketing and sales. We are good technical and product builders but were terrible at marketing and sales. We did not know anything about that or branding. However, unfortunately, that is the situation when technical people are leading companies. (SP01)

The SP03 informant summarized the company's experience with deficient management skills:

We started as a fairly technical group and did not know how to sell what we had very well. As a result, it took a long time to build a business. However, as we have grown, we have hired more experienced people who helped. (SP03)

Eight ODI firms were founded by business-minded people who recruited technical resources or outsourced technical services from outside firms. In contrast, those established by technical people complimented their competencies by hiring experienced, business-minded employees.

SP07 acknowledged that the hard part of the business was not technology but finding the right people to promote the business. SP02 attributed the success of their company to the business skills and experience of the people.

I go out, and I find challenges, I find capabilities, I find capital, and I bring it together. We are a small core team but have access to many capabilities and know-how. Our core team has just moved to five, but if we count the people working to create our offering, the number can increase to 18. (SP02)

The success/failure and growth of ODIs depended on their recognizing that their competencies and limitations impacted their actions and acquiring any missing capabilities. Analysis of

the interviews revealed that ODIs that relied only on the founders' knowledge and skills experienced limited success and growth. In contrast, the ODI firms that saw the need for technical and/or business skills were more successful. The number of employees hired, and the reported revenue indicated their success, suggesting that relying only on one's technical skills alone would not be sufficient to grow the business and expand its market presence. A good product/service known only to its creator would have a limited chance of success.

4.5.3 Financial Resources

Financial resources are the third pillar of a viable ODI (Dunn & Holtz-Eakin, 1996; Okrah et al., 2018). The interviews uncovered three sources of capital used to start and develop a business: (a) pre-seed investment, (b) government funds, and (c) seed investment. It also revealed incubators and accelerators' vital role in supporting organizations during the startup phase.

4.5.3.1 Pre-seed Investment. The sampled ODIs' efforts to explore the world of open data and create innovative products/services required funds to form and build the organization's infrastructure. Initially, all 17 ODIs relied on the founders' funds and support from families and friends. For example, one ODI founder started the business using his student loan. Another was built on a significant investment made by the founder. However, 15 ODIs built their operations on sweat equity and personal savings ranging from a few hundred to a few thousand dollars.

The need for financial resources to establish the businesses was directly related to ODIs' objectives and visions of the products/services founders intended to develop. All of them relied on the founders' or co-founders' knowledge and experiences, and when the need arose, they hired

external resources to build the products/services. Operating from home offices, university-provided spaces, and accelerators/incubators removed the need to rent expensive office space before commercializing their products/services. For example, one ODI built its operation in a university lab space. Another used the building space and services provided by an incubator, while a third built a virtual organization with no head office.

4.5.3.2 Government Grants and Funds. The second source of capital came from local, provincial, and federal governments in grants and funds. Thirteen ODIs looked to governments for financial support at one time or another. However, the results were not the same for all. There was success coupled with satisfaction and failure mixed with disappointment. The suitability of the governments' grants and funds programs for startups was questioned, as these did not consider whether startups had the resources to fulfill the stated requirements. In addition, the ODIs that received government grants complained about governments' hesitation to work with startup companies. The informant from SP09 described the federal funding programs as follows:

Federal funding is like a black box. We spent weeks on some of those applications, only to be rejected without feedback. One of the biggest things for a startup such as an open data intermediary is efficiently accessing funding. Unfortunately, the system intended to help by providing the financing is sucking away the applicant's time. (SP09)

SP04 discussed the issues and benefits:

If there were an easier path to grant writing and funding for capacity building, it would benefit our resilience because we are pretty sure lean. We do not have the skills or experience to be grant researchers and writers. Grant writers expect compensation, and we cannot compensate them on speculation. (SP04)

SP13 informant reflected on the needs of a new startup: “I think that governments can give. However, what would be most empowering for startups would be a simplification, making it simple for people to start a business and receive financing.” Meanwhile, the SP02 informant recognized the company’s funding successes, noting, “we have been successful with some federally funded programs.”

Analysis of the interviews revealed that government grants and funds helped ODIs start and sustain their operations at the early development stage of their products/services. However, applying for government grants was complex and required skills and resources that ODI startups lacked. The complexity of the granting system and its requirements created a feeling among ODIs that considered the system favoured large organizations with substantial resources and the luxury of time. The interviews also revealed that ODI founders with good business experience and large networks had a better chance to apply successfully and receive grants and funds than those new to the business environment.

4.5.3.3 Seed Capital. The third source of capital data intermediaries access is seed capital, Series A funds, and private investors. Seed capital refers to the first official money that a business venture or enterprise raises. Series A financing refers to an investment in a privately held startup

company after it has progressed in building its business model and demonstrated the potential to grow and generate revenue³⁰. Several ODIs in this study received seed funds from different sources; five ODIs managed to raise series A funds or are in the process of doing it. One ODI started its operation with an investment from venture capital and recently announced the successful completion of an 8 million series A capital investment. In addition, two ODIs attracted contributions from angel and private investors. However, the COVID-19 pandemic affected ODIs' efforts to access the capital market and hampered development efforts.

It should be noted that ODIs with social objectives and those working with governments indicated that they do not see or anticipate investors' interest in their types of business; hence they do not expect to seek financial support. The informant from SP01 reflected on the issue, saying, "The venture capital model fundamentally does not work with social investments because even though one out of 10 projects has enormous success, it does not give you a colossal payoff." The one from SP11 indicated that "being focused on solving a large social problem in our democracy that's not something is going to get the money market excited because their primary motive is profit." Likewise, the SP04 informant compared the situation to the chances firms have in the commercial sector:

If you build something in the tech sector, you get attention, money, and things keep moving. It does not seem that way in the social sector, where many are under-resourced and underfunded, and there is much competition for limited

³⁰ <https://www.investopedia.com/terms/s/seriesa.asp>

funding. I did not feel the same camaraderie and healthy competition in the social sector. (SP04)

While investors' interest is concentrated in the commercial industry, analysis of the interviews found that four startups opted not to pursue external investment for different reasons:

- If you take investors' money, they want something back. I do not want to give anything. I am not in a rush, and investors will rush me. (SP07)
- We are a 20-person organization and continue to grow using our resources, and our operation is profitable. We are presently self-funding our growth. (SP03)
- I have no intention of seeking external investment. I would be nervous about too much micromanagement and influence in the day-to-day operation of the business, pulling me in a direction that I do not think is necessary to the best of my interest. I do not want to be beholden to investors primarily. (SP12)
- A few investors understand how we operate in the big tech space. We are a government technology company just as much as a data company. Govtech is not a Netflix subscription model where you punch in your credit card and have monthly recurring revenue. (SP17)

Financial resources are critical to startups' long-term survival and development (Dunn & Holtz-Eakin, 1996). Analysis of the interviews showed that all ODIs approached the financial issue from different perspectives, espoused different policies, and selected methods aligned with their adopted business models. Two ODIs that developed ready-made application products successfully

generated the revenue needed to finance the growth and expansion of their operations. ODIs that provided custom-made services did not see a need for external financing. They were satisfied with the revenue generated from the one-customer-at-a-time business approach. The third type of ODIs, which created industry-focused solutions, sought external finance to grow and expand their business to territories beyond their local market. Five ODIs in this category obtained private and series A investments, while others are still trying to find their way forward. The last three socially motivated ODIs were interested in private investment but found it challenging to attract investors to their business sector. One informant (SP05) summed up the feeling of many open data intermediaries regarding access to investors by stating, “Access to cheap capital is critically important in the long run to build a rapidly scaling and successful company in any industry, an area in which Canada is not succeeding in compared to the US.”

4.5.3.4 Incubators and Accelerators. Analysis of the interviews uncovered incubators and accelerators’ vital role in developing and enabling new startup success. These two types of organizations provided seed money, mentoring, and office spaces to several ODIs, among other services. An accelerator is “a program that gives developing companies access to mentorship, investors and other support that help them become stable self-sufficient businesses” (BDC, n.d.). The Entrepreneur website³¹ described an *incubator* as “an organization designed to accelerate the growth and success of entrepreneurial companies through an array of business support resources and services that could include physical space, capital, coaching, shared services, and networking connections.” Leitao et al. (2022) defined a business incubator as:

³¹ <https://www.entrepreneur.com/encyclopedia/business-incubator>

Organizations that provide support to new businesses during a period of time, either through tangible assets (for example, space and administrative services) or through intangible ones (for example, knowledge and access to networks), while BA is defined as organizations that help nascent ventures get a jumpstart by providing them with funds and resources, seed capital and sometimes even working space; helping ventures to define and build their products/services; arranging for customers, and securing capital, investors, and employees. (Leitão et al., 2022)

Five ODIs, mainly in Ontario and Quebec, received support from local incubators/accelerators. In contrast, 12 ODIs did not. Nevertheless, the services provided by incubators/accelerators were highly appreciated by those ODIs that received them:

- We used an accelerator service in Waterloo; they gave us advice and a place to work. Also, the campus-based accelerator funding was beneficial for us as well. (SP13)
- We have been in many startup communities and entrepreneurial communities in the Waterloo region, where there are many support systems in place that we have harvested and taken advantage of, including the Wilfrid Laurier University Launchpad program, the University of Waterloo Velocity program, the Communitech Accelerator Center, and another accelerator in Toronto. All these programs are a combination of training, such as workshops and mentorship from experts. Some of them provided financial resources on top of the relationship. (SP10)

- The company is receiving support from Communitech in Waterloo, where we are located. Communitech offers us office space and some mentoring from other established entrepreneurs. We can leverage the community's hub and interact with many people with lots of experience in building companies, and some of them mentor us in many of our disciplines. (SP05)

The five ODIs that received support from incubator and accelerator organizations recognized the organizations' valuable services and support in putting their operations on the road to success. However, the absence of support organizations in British Columbia prompted the following comment from the SP12 informant:

There is not that much funding around here; Vancouver is not a great funding place, which has been frustrating. There is this notion that many eggs are in one basket, especially in the Waterloo region, where engineering and technology are happening in Canada. That is crazy. (SP12)

4.6 Open Government Data and Value Creation

Proponents of OGD have noted three kinds of economic benefits created by opening the data and making it available to the public for reuse: increased innovation, job creation, and economic growth (Carrara et al., 2015; Manyika et al., 2013). The following section discusses how a sample of Canadian open data intermediaries are reusing OGD, what innovations in products/services are produced, what jobs are being created, and what economic growth is occurring.

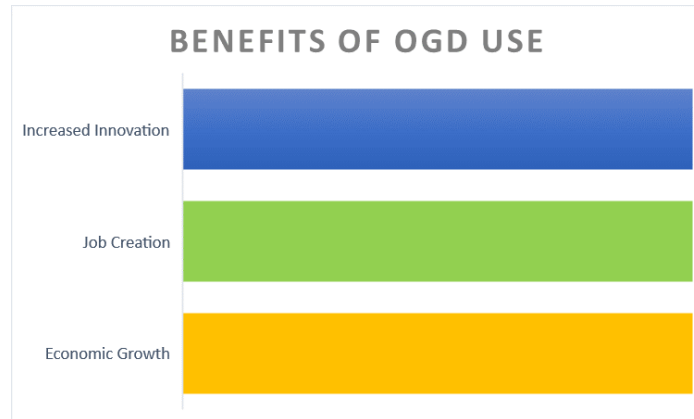


Figure 8: *Benefits of Open Government Data Use*

4.6.1 Open Data Intermediaries and Value Creation

The primary objectives of the ODIs in the sample are to create economic and social value for themselves, their customers, and their stakeholders. Therefore, they have different visions, motivations, and strategies to conduct their business, build and sell innovative products/services, solve problems, and create value.

Analysis of the interviews revealed two types of ODIs: One ODI is a fully dedicated volunteer-based social organization supporting NFP and NGO organizations to increase their efficiency and performance. Even though this organization does not intend to generate a financial return, its founder believes it creates social and economic value for society. The other 16 startups are commercially motivated, building products/services using open and non-open data to generate revenues that sustain their operations, finance the development and growth of their firms, and generate profits. Of these, three emphasize their social and philosophical principles.

SP01, a social enterprise, has built products/services to enhance charities and not-for-profit organizations' efficiencies and productivity by facilitating their efforts to find funders and secure funds. The informant described the company's work and the value of its products:

Our work is focused on ensuring that the not-for-profit sector generates sufficient evidence for decision-making. Our tools offer value in the order of magnitude, improving the efficacy of the information they can find so the fundraising departments can target their efforts. (SP01)

The informant revealed that organizations that use the startup's products/services tend to have a 30% increase in their funding. As a result, more efficient operations have saved them time and money and generated more funds for their missions³². Regarding job creation, the informant claimed that the financial return from selling SP01's products/services had permitted the creation of 10 jobs that have contributed to expanding the company's services to include data consulting. This additional service has generated critical revenue that has sustained the company and rendered it profitable.

SP02 considers itself a social enterprise with a mission to contribute to society's goods. It has built solutions to facilitate local governments' communication with the public and enable citizens' understanding of government decisions through interactive predictive analytics and data visualization tools. The informant described the solution developed and the value it has created:

³² Claim not verified.

Our software flattens the issue so everyone can view it on one platform and understand the correlated matters. As a result, we help government and citizens collectively plan a more viable future. It is our way of helping shape how the world develops: street by street, block by block, building by building. (SP02)

Sales of the products/services have generated constant revenue that contributed to the company's profitability and created a core workforce of five people to manage operations and client relationships. Additionally, the company has used the services of external technical workers on a project-by-project basis. As a result, the informant believed the solution has increased citizens' participation and contributed to efficient operations and cost reductions for participating municipalities.

SP03 was established to provide customized data services to public and private organizations. The informant described the services offered to clients: "data cleaning, data modelling, and help in making better decisions from the data which translate into either reduced cost, increased revenues, or better service to their customers." Sales of the products/services have generated constant revenue, enabling the company to build a strong workforce of 20 people and expand its products/services to the private sector. In addition, the informant expressed satisfaction that the company's products/services have contributed to social value creation:

Our product generates social value; more lives are saved, and ambulance and fire services are improved. There is also a fiscal value benefit from using the product to avoid the costs they would otherwise have from poor decisions. (SP03)

SP04 is an NFP social purpose organization focusing on improving partners' capacities and strengthening their ability to influence society. The organization, built on the work of approximately 8,500 volunteers, provides free data services to NFP and NGO organizations nationwide to enhance their efficiencies, productivity, and efficient use of funds. The informants expressed satisfaction with the free services the organization can offer its partners, which directly affect partners' efficient use of funds to support other causes and increase the organization's social footprint:

Our services impact our partners' efficiencies, operations, fundraisings, and the effectiveness of their campaigns. Our services are more significant than just creating economic value for not-for-profit or charity organizations because what we do for them advances their greater good and missions. (SP04)

SP05 developed a software tool enabling industrial energy users to save on energy bills. In addition, the software tool provides valuable data that enhance users' capacity to make educated decisions that reduce the company's expenses and increase its efficiency. The informant described the direct and indirect impact of using the software:

Our predictions fundamentally help the energy market operate more efficiently. Our end customers save millions of dollars by reducing electricity usage when prices spike suddenly. By reducing use, the entire energy grid is more efficient and can, in the long run, save billions by building more efficient grids. Consequently, we also significantly reduce CO₂ emissions by reducing the need for natural gas peaking facilities. (SP05)

Sales of the products/services have been growing steadily and generating rising streams of revenue that have contributed to the creation of 22 jobs and placed the company on the road to profitability.

SP06 is a data science company that provides insights and knowledge from open and non-open data collected worldwide. The company's services are available to all individuals and organizations in the public and private sectors. The informant described the value of its services as being located in insights and knowledge it shares with its clients and in its saving customers the efforts to recruit and build their resources to do the work: "Our services relieve customers from building dedicated internal resources to do the same job, reducing cost and increasing efficiency." In addition, the informant expressed satisfaction with the revenue the products/services generate, enabling the company to expand its technology-based infrastructure from one physical server to more than 70 cloud-based servers and a team of three people.

SP07 built a digital tool to help small municipalities efficiently manage parking spaces and facilitate citizens' mobile phone access and payments. The market for the products/services was later extended to private parking operators and other public organizations. The informant described the products/services and their value to potential customers:

Our solution is about better infrastructure and money management employing the same assets. We optimize the software tool for private parking, transit, and taxis. Customers do not have to manage the services anymore. They do not have to

collect parking fees through meters; they can now use the phone and receive the money automatically. (SP07)

Sales of the products/services to 75 customers have generated good returns and enabled the company to expand its workforce to 22 people and its market to other provinces in Canada and cities in the US.

SP08 saw an opportunity in the 3D mapping industry and set to build the infrastructure to make 3D data available to the market. Its products/services are still under development and targeted at municipalities, the AEC industry, and future autonomous driving and smart cities industries. The informant described the purpose and value of the developed product/service: “The product aims to reduce the cost associated with the customer developing a similar product internally. In addition, it enables time-saving and provides capabilities that permit users to offer better services to their internal and external customers.” The company currently provides data services to municipalities project-by-project and expects its 3D-based products/services to be available soon. Consequently, the revenue generated so far is limited, and the company still depends on government funds and grants to sustain its operation. Nevertheless, the company has created 15 jobs and plans to expand its products/services to the US market.

SP09 focused on simplifying access to open government 3D data and making it available to professionals in the AEC industries. The company built a portal that aggregated related 3D data from multiple sources and made them available to AEC professionals. The informant described the products/services this way:

We built an online mapping platform that attempts to solve two problems: search engine for maps—we are pulling from open data portals worldwide where we can access them—and interaction with data. We developed a tool for people who do not have GIS or CAD software training to interact with that data and leverage it for communication. (SP09)

The informant described the products/services' value as "efficient access to a massive amount of open geospatial data that would have otherwise been difficult for end users to access." The company received support from a local accelerator to build its customer base. However, it has not yet generated sufficient revenue to sustain its operations and grow its limited workforce.

The **SP10** data intermediary built a solution targeted at the real estate industry, offering access to data on planning and construction regulations that facilitate decision-making. The company captured and aggregated data on planning and construction regulations, created a data portal, developed software tools, and extended its services to real estate developers and AEC professionals in the real estate industry. The informant described the products/services offered and the value created:

We pull development application and planning data and visualize them onto a Google Maps platform. For example, it takes a few seconds for a developer who wants to build a condo to understand the conditions of the real estate or location intelligence: how far away from transit, or the designation of the area, mixed-use

or not, and allows them to know who is building around that area, and whether it is a hot neighborhood or not. (SP10)

The company has been in business for almost five years and is actively pursuing business development efforts to recruit customers and partners. As a result, it has received support from an essential player in the real estate industry. It has started to seek seed investment to sustain its operations, finance expansion into the US market, and develop its limited five-person workforce.

SP11 is a social enterprise driven by its founder's ambition to solve problems in the public sector, build a solution to facilitate governments' and public agencies' interactions with citizens, and conduct consultations with geo-verified individuals. The platform is based on a database of geo-coded citizens segmented by provinces, regions, municipalities, and neighborhoods. The informant described the critical value created by the platform in one word: "trust."

Trust is the most critical problem to solve. We solved it through a combination of digital identity authentication, protecting individual privacy on the one hand and the ability to stand behind the quality and integrity of the data produced from the process. (SP11)

The company's sales activities have not yet generated the anticipated revenue, forcing it to consolidate its services, concentrating on supporting its existing clients and maintaining the products/services through a team of six people.

SP12 is a one-person, project-oriented company providing public and private organizations with custom solutions. The value offered to its potential clients is based on its knowledge and skills in data science and visualization. The informant described the value delivered to clients as “insights that they would not be able to create themselves without hiring skilled resources, which saved them time, reduced their cost exposure, and increased their decision-making and operations efficiencies.” The informant expressed satisfaction with the financial results obtained since the company’s inception and revealed no plan to grow or expand the operation.

SP13 created a data portal for visa requirements in all countries. It used advanced technologies to build the software tool to capture worldwide data from and curate them for ultimate use by travellers, airlines, and travel agencies. It also offered visa application processing services to those who requested assistance. The informant described the value of the services to travellers as “comfort,” “peace of mind,” and “time-saving.” The value to airlines is that the service ensures passengers have proper travel papers and avoids disruption.

The sales of the products/services generate revenue streams that have contributed to the company’s profitability and led to expanding the workforce to 20 people. Moreover, its success has attracted venture capital investors who have provided new funds to finance the growth and expansion of the company.

SP14 built a data portal populated with data on diseases and health conditions worldwide. In addition, it created a system that alerts travellers to health hazards in the countries they intend to visit or are currently visiting. The informant described the value of the services offered to

travelers in terms of “lowering anxiety”, “comfort”, and “trust.” As evidenced by new customers, the products/services have proved their viability and market acceptance, and higher sales have generated sufficient funds to expand the company workforce to 13 people. The informant indicated that the company had attained profitability; however, this is not its top priority. Instead, it is focused on building a stable workforce and valuable products/services to ensure the company’s sustainability and future profitability.

SP15 is a custom-made solutions provider offering knowledge and expertise in geospatial data to build solutions for clients based on their needs and requirements. The informant described the company’s offering: “We provide geospatial expertise and our ability to back up our advice with code, if necessary.” The informant also indicated that its services give its customers “comfort” about what they do with geospatial data and save them “lots of money.” In addition, the company’s business model has ensured continuous growth, allowing it to reach the 5 million dollar mark and build a strong workforce of 25 people working virtually from locations around the county. The company sees value in its business model and method of operation, which have been proven to deliver the anticipated revenue without the associated cost of being a national service provider.

SP16 was built on open data to facilitate customers’ access to a central open data portal where open datasets from different sources and locations worldwide are aggregated or linked and made available to the public free of charge. On top of the data portal, the company offers data services to its clients that generate the revenue necessary for its sustainability and growth. The informant described the services, and the value customers get from them:

We help customers connect to external data from a single place where the data is always up to date at the desired service level in the schema they want, allowing easy integration. They could connect to as many datasets as they wished to, blend them, perform the analysis, and get the answers they were looking for to a problem. We freed up that extra data cleansing and finding and connecting work and let them work with the data they needed. (SP16)

The services provided by the company have received acceptance from customers, generating streams of revenue that have enabled the company to increase its workforce to 40 people and open an international office outside Canada. The company's success has also raised public awareness, generated attention from the financial market, and led to a significant external investment to support its growth and expansion.

SP17 built a software tool to help municipalities communicate with constituents and the public. The product has enabled cities to communicate with the public, creating business opportunities and visually increasing citizens' participation and trust. The software product is a dashboard that facilitates local governments' publication of their data on demographics such as population, labor force, employment rates, or information on the housing market, which is useful to all stakeholders. The informant described the purpose of the application as being to inform and engage citizens and the business community, build trust, increase community services, and stimulate economic development. The potential value enabled by the product was noticed by more than 500 municipalities that purchased and used the tool. This sales success translated into higher revenue, contributing to company profitability, and creating 30 jobs. The company's workforce was later

reduced to a small team of five people responsible for supporting customers and maintaining the product.

Analysis of the interviews showed that the sampled Canadian ODIs using open and non-open government data have been busy creating economic and social value for the last decade. Their economic value is created directly from the innovative products/services they have introduced to the market and is represented by revenue and job creation. Input by all 17 ODIs confirmed revenue generation, creation of jobs, and contribution to economic development. The impact of the income generated is evident in the investments made in expanding ODIs' workforces to build their products/services, satisfy customers' requirements and needs, and expand their respective market presences. The 17 ODIs reported having created more than 217 jobs (see Figure 9) and experiencing different levels of success with expansion. Sixteen said they were adding employees to their operations, while one ODI had opted not to add any. The number of internal jobs created ranged from none to 40. No figure is available for the indirect jobs created; many ODIs acknowledged recruiting part-time employees or contracting services from other service providers.

Revenue is the second indicator of value creation. Nine ODIs described their operations as profitable. Three reported business growth and are expected to reach profitability quickly. Four are still at the development stage regarding their market and customers; however, their revenue has not reached the break-even point. As previously noted, the volunteer-based ODI is not seeking any direct revenue and believes the services it provides to its partners contribute to economic development.

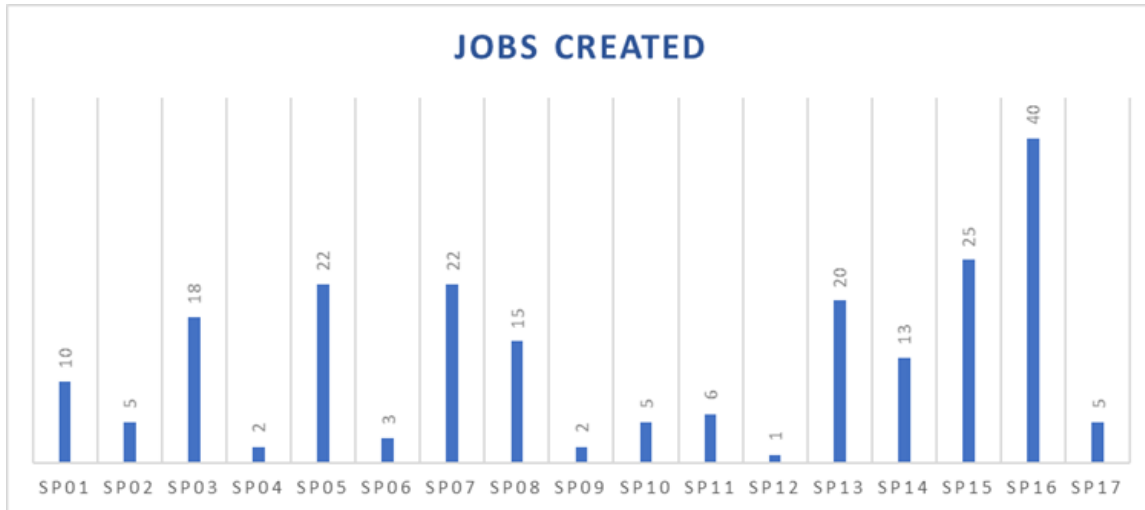


Figure 9: *Jobs Created by ODIs*

ODIs also indicated that their products/services had created indirect value by increasing customer operations efficiency, time savings, and cost savings. The increased efficiency is attributed to better decision-making resulting from insights and knowledge provided by the ODIs and the products/services their customers have purchased. The cost savings are linked to the outcomes of efficient operations and saving from outsourcing the service of ODIs instead of developing internal capabilities. The time savings are associated with cost savings and are attributed to organizations' choice between building internal expertise or outsourcing the services of knowledgeable and skilled professionals. ODIs recognized that their companies' knowledge and skills are in short supply and expensive. Therefore, those companies' services are in demand by small and medium-sized organizations, municipalities, and NFP organizations with limited financial resources. The informants considered the products/services provided, insights from the data, consulting services, or a software product to enable customers to create additional value beyond efficiencies, time, and cost savings.

Many open data intermediaries expressed that in addition to creating economic value, their products/services have contributed to social value creation: building trust, increasing citizens' participation, expanding the footprints of NFP and NGO missions, efficient decision-making, and efficient energy grids and environmental benefits. For example, one ODI described the social benefits of its products/services as: "better ambulance and fire services, lifesaving, which saved governments "millions of dollars." Moreover, the evidence discovered in the analysis of the interviews confirms that innovative products/services created by reusing OGD contribute to economic value creation.

4.7 Challenges and Barriers

The researcher's analysis exposed multiple challenges open data intermediaries faced while establishing their operations, building their infrastructure, developing their products and services, finding customers, and selling their products/services. These challenges and barriers are classified into three categories: (a) organization development, (b) financial, and (c) data sourcing.

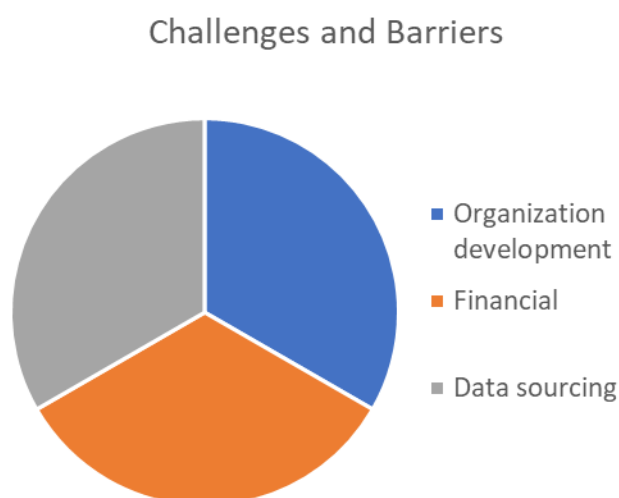


Figure 10: *ODIs Challenges and Barriers*

The Canadian ODIs that participated in this research are startup firms under ten years old. They were all founded by one or two people with specific knowledge and skills, visions, and a limited amount of money. The founders' backgrounds varied: six came from the private sector, four were government employees, four transitioned directly from academia to business, and three were previous entrepreneurs. Their diverse backgrounds and experiences produced different challenges; however, they all engaged in building their organizations' human resources, securing funds, and sourcing data.

The lack of business knowledge and experience among founders affected the performance and development of their firms. Founders with technical backgrounds had problems recognizing the importance of the soft skills needed to manage, promote and sell the products. Those transitioning from public and private organizations experienced difficulties building and promoting their initial products/services due to a lack of knowledge of customers' expectations and requirements. Socially motivated data intermediaries overestimated the market and the impact of their products/services on revenue generation. Thus, almost all ODIs reported having had to take corrective measures to refocus their efforts, refine their products/services, and build soft and hard skills to ensure successful operations.

4.7.1 Organization Development Challenges

The challenges reported under this category were attributed to a lack of business experience, staff shortage and affordability of skilled and experienced people. One ODI informant summarized the feelings of many others by describing the most significant challenge as having to be "comfortable with uncertainty." Others admitted that while their technical skills were excellent,

their ability to manage the business and develop sales was inadequate. The SP03 informant summarized their company's challenges: "We started as a fairly technical group and did not know how to sell what we had. Early on, we did not have a lot of mentorship and relationships with other startups or executive coaching." The SP09 informant considered the lack of staff to do administrative work a challenge that affected the startup's innovativeness: "one of the hardest things is having to do the administrative work running the company while also trying to be innovative."

Constrained by their limited financial and technical resources and business experience, several ODIs took different measures to mitigate staff shortages. For example, SP01 hired technical people in return for a stake in the company, SP05 recruited students, and a few other ODIs subcontracted services on a project-by-project basis. In addition, five ODIs established relationships with local incubators and accelerators to build their business knowledge. They received mentoring, training, and opportunities to connect with other executives that augmented their business knowledge and perceptions. One additional challenge related to organizational development that two ODIs faced was the development of their products/services. As a result, they had to drop their "presumed knowledge" of what customers wanted and work with potential customers to learn and understand the latter's actual needs and requirements before developing the final products/services.

4.7.2 Financial Challenges

Funding a startup operation is a critical challenge many data intermediaries reported having faced. Access to public and private funds is not simple. For example, the SP04 informant described finding and applying for government grants and funds as time-consuming. In addition, it requires

a full-time job requiring specific skills and expertise, noting, “We do not have the skills or experience to be grant researchers, writers, or the time or ability to develop that skill set.”

ODIs that applied for government grants and funds reported mixed results. Some were satisfied, and some were not. Analysis of the interviews revealed many obstacles when applying for government grants and funds. The governments programs required some knowledge and experience of the application process and dedicated resources to research, use, and follow up on the applications, which startup companies did not have and could not afford. Several data intermediaries complained about the application process for federal grants and funds, considering it to be aligned with corporate capabilities and not those of new startups. The SP09 informant described the federal funding program as a black box: “We spent weeks on some of those applications only to have been rejected with no feedback.”

Those who were unsuccessful in accessing government grants and funds turned toward the private market using different approaches, including crowdfunding. However, the four socially motivated ODIs were not enthusiastic about getting private investment. For example, the SP01 informant described the social organization as less attractive to venture capital (VC): “The VC model fundamentally does not work with social investments. Even though one out of ten projects has enormous success, it does not give an enormous payoff.”

The findings also revealed that ODI founders with prior business experience and networks were more successful in securing grants and funds than newcomers. Moreover, the ODIs’ business models affected their revenue generation capabilities and external investment needs. For example,

those that selected the direct sales approach and operated based on one customer at a time ensured adequate financial returns, eliminating the need to seek outside investors. Incubators and accelerators' financial and business support programs helped five startups overcome hurdles and improve their chances of success. However, it should be noted that not all ODI startups had access to incubators and accelerators in the provinces where they operated. Consequently, they had to look for other financial options to support their organizations.

4.7.3 Data Challenges

Data are the raw materials open data intermediaries depend on to build their products/services. Sourcing these critical materials was not as easy as some ODIs expected. The process varied between one ODI and another depending on the intended use of the data. Before implementing the open government data initiative, data intermediaries working with local governments reported many problems accessing governments' data. They cited a lack of policies and regulations, insufficient knowledge of public employees, and finding departments and people with the appropriate authority. As the informant from SP09 noted, "It was a convoluted process to access the data and second to have any insight from it." Data intermediaries had to approach local government departments or agencies to request the data, negotiate conditions of release, and wait until their requests were accepted and fulfilled. When the data was received, it was not structured, requiring multiple tasks to make it usable; data had to be cleaned, structured, and linked to form the basis for software applications and tools to be developed and used. Early data intermediaries expressed their frustrations with the process of accessing the data they needed:

- We had to spend lots of energy and build many technical infrastructures just to access the data appropriately because it was coming off websites and had to go into a database where the information is linked. (SP01)
- It has been difficult to understand this complicated market, as data is very difficult to access. (SP05)
- We do not have a problem with data structure or data cleaning issues. The main issue we often face is access to data. (SP07)

Following the adoption and implementation of OGD initiatives by governments at all levels, data intermediaries had problems finding and accessing governments' data on data portals. When the data was located and accessed, ODIs discovered different standards used by other departments and agencies and local governments reporting on the same issue or topic (budgets). The inconsistencies cost ODIs time, energy, and money to restructure, link, and prepare the data for use. The informant from SP09 described the problems and solutions needed for open data: "One extremely frustrating thing is the difference in standards of what municipalities do. I think, as a country, some more standardization is needed. We should standardize how data is shared."

ODIs reported many problems with captured OGD. For example, data was found to have quality issues; to be "incomplete," "broken in smaller datasets," "not updated," "inaccurate," and "highly aggregated,"; and to lack appropriate "versioning," among other problems. Another common issue reported by many ODIs was difficulty finding the needed datasets. This caused them to contact the individual/department responsible for the datasets to get the required data. Table 8 lists the primary challenges reported by ODI startups related to data.

Table 8: *List of Open Data Challenges Reported by Open Data Intermediaries*

OGD Challenges	
Data access	Data validity
Data update	Data format
Data Quality	Data Search
Data Accuracy	Data Finding
Data Availability	Metadata
Data Structure	Data Aggregation
Data Version	

The following statements by ODIs describe the various issues with accessing and using open data from governments' portals:

- The online data is so aggregated or summarized that it is not very useful. (SP03)
- Publicly available data, especially weather information, can be a little bit less accurate. (SP05)
- Quality varies significantly from dataset to dataset. (SP09)
- They only update every quarter or every year. (SP10)
- It has gotten a bit better in some ways, but it is the same old challenges that I think have been with us for a long time: Various data formats and data quality issues. The ability to search for data has not been great, and it is still not. The versioning of data, finding somebody to contact in case they have a question or need a correction. (SP12)
- There is usually a lag between when the data is available and when the actual thing ends. There is no common way to access the data. (SP16)

The objectives the Canadian open data intermediaries participated in this study set out to accomplish were not easily achieved due to the startups' lack of knowledge, business, or technical experience, or human or financial resources. However, those that recognized their limitations and weaknesses and addressed them through mitigation strategies paved the way to build stable operations that delivered valuable products/services and created economic value for themselves and their customers. For a complete list of the challenges reported, please refer to Appendix E.

4.8 Chapter Summary

This chapter presented the findings from the 20 interviews conducted with members of 17 open data intermediary firms. In addition, it offered an overview of the open data intermediaries who participated in this study, their profiles, and their motivation for working with OGD. It also described how the ODIs used OGD, what products and services they created, and the resources and capabilities needed to achieve their tasks. The final section illustrated the challenges facing ODIs from organizational, financial, and data resource perspectives.

Chapter 5. Discussion

Open government data is valuable material projected to create social and economic benefits. Studies of this new phenomenon have estimated the benefits generated in billions of US dollars (Manyika, Chui, Groves, Farrell, Van Kuiken, & Doshi, 2013). However, few empirical studies have attempted to validate this claim and provide an evidence-based description of how and what value is generated from OGD.

Canada was among the leaders in adopting and implementing an open data initiative. It opened government data and made it publicly available for reuse in 2011. However, after a decade of data availability, there are no empirical studies on how Canadian organizations use this new material, what products/services are created, and what value is generated.

This study explored how a sample of 17 Canadian open data intermediaries used OGD to innovate and create products/services that generate value. To accomplish the objective of this study, the researcher raised the following question and sub-questions to guide the research and focus the data collection activities:

- Q. How do data intermediaries create value using open government data?
 - a. How are open data intermediaries using open government data? (Answers can be found in Section 5.2).
 - b. What are the resources and capabilities vital to open data intermediaries? (Answers can be found in Section 5.3).

- c. What are the challenges and barriers faced by open data intermediaries? (Answers can be found in Sections 5.1, 5.4 and 5.5).
- d. What value is created, and how? (Answers can be found in Section 5.6).

5.1 Open Government Data

The concept of OGD, its meaning, and the conditions that define it are not yet well understood by all participating Canadian ODIs. For example, several ODIs incorrectly viewed the unstructured data given to them by government departments and agencies as open data, while others wrongly considered the unstructured data and information available on the Internet to be open data. Only a few ODIs understood the concept, structure, and implications of open data and open government partnership. Those few were found to be individuals who transitioned from governmental jobs to the private sector. The Treasury Board of Canada's evaluation report on the open government program recognized the lack of awareness of the open government data initiative. The report found that only 8% of the public participated in the open government learning program, with the majority, 74%, coming from the federal government (Treasury Board of Canada, 2021). This observation partially answers research sub-question *c* on the challenges and barriers faced by open data intermediaries (Section 1.5).

5.2 ODIs' Use of Open Government Data

This section answers research sub-question *a* on ODIs using OGD (Section 1.5). Analysis of the sampled ODIs' uses of OGD found that their primary objectives were to find opportunities to exploit or issues to solve. Upon deciding on the opportunity/issue, they selected the related data. The analysis identified four archetypes of ODIs that used OGD to create products/services aligned

with their objectives and philosophies: (a) aggregators, (b) developers, (c) enablers, and (d) suppliers.

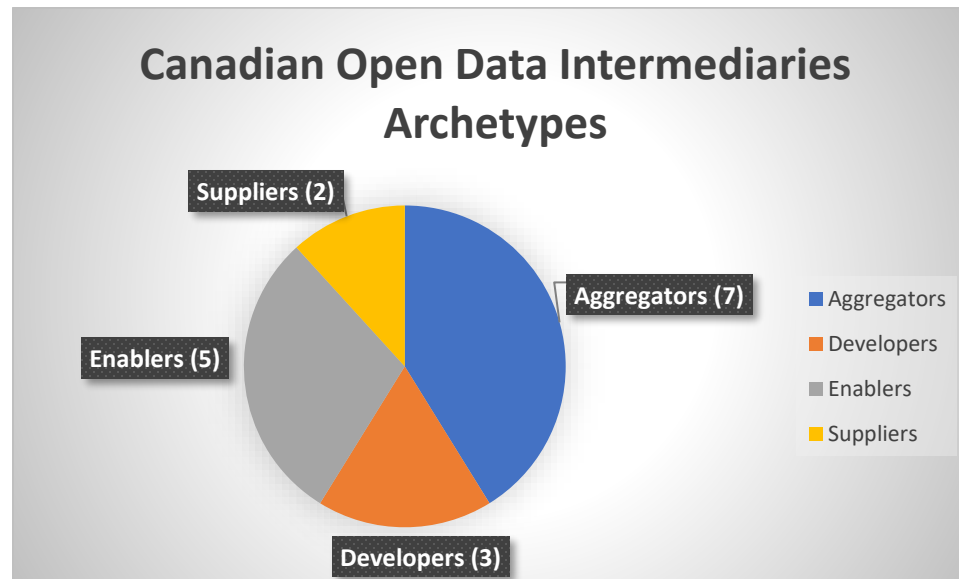


Figure 11: *Canadian ODI Archetypes*

The *aggregators* collected data from different sources, added value through data curation, and created data repositories to provide information and knowledge to their target customers. Loshin defined an aggregator as “an organization that collects data from one or more sources, provides some value-added processing, and repackages the result in a usable form” (Loshin, 2013, p. 295). Seven of the sampled Canadian ODIs are classified as aggregators and represent 41% of the 17 ODIs in the sample.

The *developers* are professional organizations designing, creating, deploying, and updating programs for a particular device, the web, or an operating system (Indeed, 2022). They have expert knowledge of information and communication technologies used to develop software and create a

computer program or a set of programs to perform the tasks required by a business. Three ODIs, or 18% of the 17 sampled ODI firms, are classified as developers.

The *enablers* are organizations that work directly with their customers. They provide knowledge and technical skills to analyze and extract knowledge and information from data that enhances decision-making, reduces cost, and saves time. Knowledge enablers are mechanisms for intentionally developing knowledge in an organization (Ichijo, Von Krogh, & Nonaka, 1998). Someh et al. (2017) described business analytics (BA) as a system that analyzes organizational data and enables informed and insightful decision-making. Lee and Choi (2003) described knowledge management enablers as “organizational mechanisms for fostering knowledge consistently; they can stimulate knowledge creation, protect knowledge, and facilitate the sharing of knowledge in an organization” (p. 181). Five ODIs, representing 29% of the sampled firms, are classified as enablers.

The *suppliers* are organizations that capture open government and non-open data, create data repositories, and supply datasets or insight and information derived from the captured data. Two Canadian ODIs have chosen to develop data repositories and provide insights or datasets per their customers’ requirements. They represent 12% of the sample population. The Macmillan Dictionary has defined a supplier as “a company, organization, or country that supplies or sells a product or a service” (Macmillan Dictionary, n.d.). Similarly, the data firm Snowflake has described a data supplier as follows:

A data supplier is an organization or business that provides data for use or consumption by third parties. Some data suppliers provide access to datasets for free; some sell data for a cost, and some offer free and fee-based data services.

(Snowflake, 2022)

The European Commission study of open data reuse in Europe reported five types of re-users of open data, namely the four identified in this study and a fifth described as *enrichers* (Berends et al., 2020). Data enrichment is about merging third-party data from an external authoritative source with an existing database of first-party customer data (Zisk, 2018). The analysis of the findings in the current study showed that the functions of enrichers were performed by all the sampled Canadian ODIs, which used external data to enrich the products and services they provided to their customers. Therefore, enrichers are not included as a separate archetype in the classification of sampled Canadian ODIs. Table 9 compares the European and Canadian types of users of open data and the percentages of that population they represent.

Table 9: *European vs. Canadian ODI Archetypes*

Archetypes	Canadian	%	European	%
Aggregators	√	41	√	38
Developers	√	18	√	17
Enablers	√	29	√	17
Enrichers		0	√	14
Suppliers	√	12	√	15

Note: European data retrieved from the European Commission study of open data reuse in Europe (Berends et al., 2020).

The roles played by Canadian ODIs led to the creation of four different categories of products/services that used OGD to create value: (a) industry-focused data portals, (b) insight and knowledge extraction services, (c) custom-made solutions/services, and (d) software solutions for general application (see Figure 12).

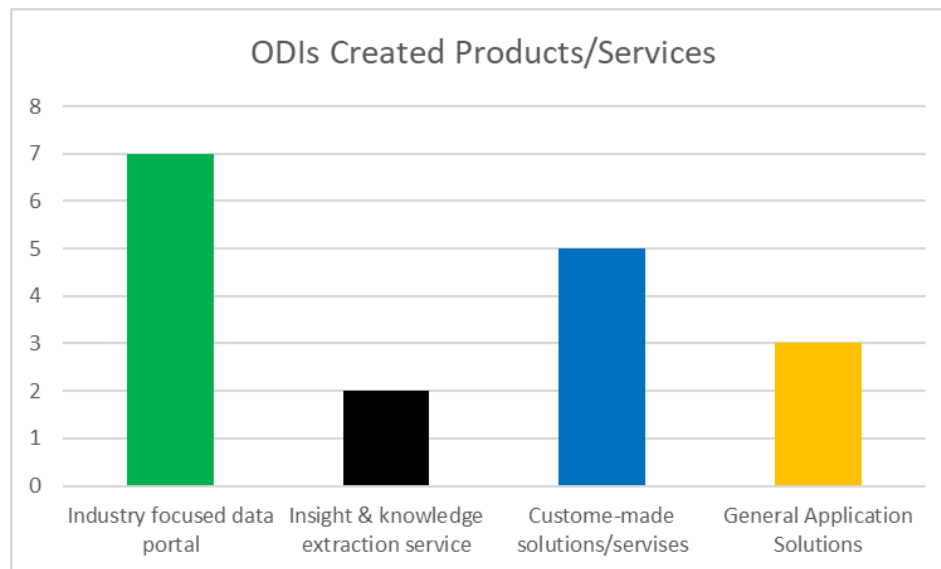


Figure 12: *ODI-created Products/Services*

5.2.1 Data Portal Creation

Seven ODIs focused on specific industries, captured open and non-open government data, and created data portals to meet the needs and requirements of a particular business sector. A data portal is a “private or publicly facing application built to allow end-users to explore, interact with, and derive intelligent insights from data to perform a task or answer a question” (ZUAR, 2022).

The capture and maintenance of the data needed to populate the portals required industry knowledge of its needs and requirements. Technical skills, advanced information technologies, and software development expertise were also necessary to create the service. The analysis of the interviews showed that seven ODIs had previous knowledge and experience in the targeted

business sectors, used their internal technical resources, hired skilled engineers, or outsourced the service from specialized firms to develop the data portals. Creating the data portals involved multiple activities, including capturing and curating data, preparing data, designing user interfaces, and hosting the data portals in the cloud. In addition, those ODIs that pursued this business model invested in computer hardware and software development to enable users to use their data portals, subscribe to the services, pay, and access the relevant information. Table 10 lists the ODIs' data services and target customers.

Table 10: *ODIs' Data Services and Target Customers*

ODI	Data services	Targeted customers
SP01	Funders' data	NFP organizations
SP08	3D mapping data	3D industry, AEC industries
SP09	3D mapping data	AEC industries
SP10	Real estate data	Real estate industry
SP13	Global visa requirements data	Travellers and the travel industry
SP14	Global health data	Travellers and the travel industry
SP16	Open data	Open data seekers

The data came from open government portals and government websites, including the Internet, private organizations, and online media. They ranged from data on NFP, funders, geospatial, open data, visa requirements, and health-related issues. The data capturing was not bounded by the Canadian border; it went beyond that when the conceived services required so.

5.2.2 Insights and Knowledge Extraction Services

Two ODIs captured and used open and non-open government data to extract insights and knowledge from the data. To accomplish their tasks, they used machine learning (ML) and artificial intelligence (AI) technologies to automatically scan local and global websites, capture the

data, perform analysis, and extract knowledge and insights. AI is a machine technology that simulates human intelligence. McCarthy (2004) offered the following interpretation of AI: “It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable” (p. 2). The Oracle Corporation (2021c) has similarly defined AI as “systems or machines that mimic human intelligence to perform tasks and can iteratively improve themselves based on the information they collect” (Oracle, 2021c, n.d.). ML is a branch of AI and computer science that uses data and algorithms to imitate how humans learn, gradually improving accuracy (Bell, 2015; IBM, 2021c).

5.2.3 Custom-Made Solutions/Services

Five of the participating ODIs relying on their information technology skills and experience, chose to work with private, public, and not-for-profit organizations on a project-by-project basis. Using data provided by their customers and/or captured from the Internet or governments’ data portals, they offered insights, answered questions, or developed custom-made solutions to issues as requested by their customers. For example, IBM (2021a) defined custom-made software development as “the process of designing, creating, deploying and maintaining software for a specific set of users, functions or organizations” (IBM, 2021a, n.d.). The services provided by this category of ODIs are dependent on customers’ needs and requirements and follow the specifications set in contractual agreements.

5.2.4 General Application Software Solution

The last group consisted of three ODIs that developed solutions to opportunities identified in the market before performing marketing and sales activities. They focused on facilitating local

governments' use of resources to communicate with their citizens through innovative products and services. As a result, three standard software applications were created.

The researcher's findings show that open data sources were not the only ones used by ODIs to innovate. Almost all ODIs used open and proprietary data to develop their offerings and solutions. Early ODIs started working with publicly available and proprietary data and later integrated open data when open datasets became available. In contrast, ODIs that began their business operations after governments implemented the Open Data Initiative discovered that the open public datasets were insufficient to satisfy their and customers' needs. Consequently, they incorporated proprietary and publicly available data to deliver the products/services requested.

This finding concurs with the European Commission study on the reuse of open data (2020), which found that European startup companies reusing open data do not solely rely on that data. Instead, businesses often combine open data with private data they hold or their clients provide (Berends et al., 2020). Table 11 lists the data types and sources the sampled ODI firms used.

Table 11: *Data Types and Sources*

Data Type	Open	Proprietary/ Private	Public/ Web
Mapping and geospatial	√		√
Health	√		√
Energy	√		
Weather	√	√	√
Statistics	√		
Population		√	
Visa requirements	√		√
Motor vehicle		√	

Data Type	Open	Proprietary/ Private	Public/ Web
Public/private parking spaces		√	
Funding organizations	√		√
Regional and urban	√	√	
Federal government	√		√

5.3 ODIs' Resources and Capabilities

This section answers research question *b* on resources and capabilities vital to ODIs (Section 1.5). The 17 ODIs that participated in this research are startup firms that entered the business arena of open data 10 years ago or fewer. They had diverse visions and purposes. Some had substantial experience, and others had limited or no experience. However, all of them aimed to contribute to the common good. Eleven ODIs were founded by pairs of partners, while six had one founder each. The founders were individuals with technical and/or business knowledge and experience who transitioned from public organizations, the private sector, academic research, universities, and entrepreneurial activities to build their startups using open and non-open data resources. All but one ODI started business operations with limited financial and human resources.

The findings reveal that human and financial resources were the critical elements that ODIs needed and sought to acquire. However, many challenges forced them to take different approaches and actions to address them, resulting in mixed outcomes.

5.3.1 Human Resources

The findings show that the ODIs in this study needed two types of human resources to build their operations and develop their products/services: business and technical resources. Those with in-house concentrations of technical skills discovered they could not market and sell the

products/services they built. They were forced to acquire business knowledge and skills by hiring experienced professionals or learning the trade. Unfortunately, neither option was easily achieved; hiring professionals was financially tricky in the early stage of product/service development, and there were short timelines for understanding marketing and sales. As a result, five data intermediaries turned to accelerators and incubators to get the support they needed in business mentoring and relationship building; a few ended up hiring the required resources, while a smaller number learned to live with their limited resources at the expense of business development. This finding concurs with the European Commission study of open data reuse in Europe, which found that 8% of European re-users of open data lacked marketing resources, and 13% lacked salespeople (Berends et al., 2020).

People with experience, knowledge, and technology and data science skills were the second resource needed. This group, described as data scientists, is critical for every organization working with data. People in this group are skilled in computer hardware and software, competent in machine language and/or artificial intelligence, good at designing and building application software and tools, and proficient in statistics and analysis. Data scientists are responsible for scanning, capturing, cleaning, organizing, combining, and preparing data. They are also involved in analysis, programming, and application development. Several ODI founders possessed some of the competencies and skills of a data scientist.

In contrast, ODIs with data science skills shortages turned to the market to recruit resources or contract services from professional organizations. The European Commission report listed data scientists as the human resource profile most demanded by European organizations, representing

49% of their human resource needs. The report described the implications of working with open data and proprietary data and the skills and knowledge needed:

Working with open data brings certain particularities not experienced in working with proprietary data, which have to be considered by a potential employee. The employee needs to understand, for example, the concept of open data as a shared resource but also requires the skills to contextualize or refine the data to his needs. (Berends et al., 2020, p. 63)

The availability and affordability of data scientists in Canada were reported as a significant challenge by several ODIs, which resorted to different recruitment options, such as learning some skills or hiring students to develop products/services. The European Commission study also recognized this challenge and considered it a barrier to the future reuse of open data (Berends et al., 2020).

5.3.2 Financial Resources

Access to financial resources influences a new startup's potential success and survival. However, accessing external funds is not easy or at times even feasible during the early days of operations. The ODIs in this study were all bootstrapped by their founders, except for one, where the investment was significant. The other 16 open data intermediary founders started their businesses with limited funds, using their savings and support from their friends and families. Recognizing that their success and long-term survival depended on generating revenue to sustain the operations early, the ODIs pursued different avenues to secure the funds needed.

5.3.2.1. Revenue. Income generation is the first objective pursued by ODIs selling their products/services to potential customers using a direct sales approach or other business models. The revenue generated by the ODIs in this study from sales of their products/services varied among them. Those that experienced success opted not to pursue external investment. Instead, they concentrated on growing their companies and developing their markets using the revenue generated from the sales. It should be noted that many of the ODIs generated revenue from data consulting services in addition to the products/services they sold, with a small number reporting income from the consulting services that exceeded the revenue generated from the products/services.

5.3.2.2 Government Grants and Funds. Government grant and fund programs were the second type of financial resources most ODIs sought to secure at one time or another, with mixed results. Some were successful, and others were not. The federal programs were considered a nightmare, and the requirements of the applications and the resources needed to complete them were described as prohibitive for new startups and only made sense for large organizations. Thus, a startup's chance of winning funding was slim. However, the disappointment with the procedures to apply for government grants and funds was not shared among all ODIs; several expressed satisfactions with the money they could receive and put into development. This contradiction in the views of governments' funds and grants exposed the critical role experience and human networks play in applying for and securing government funding.

5.3.2.3 Business Incubators and Accelerators (BI and BA). Incubator and accelerator support programs were critical for five ODIs that received financial support, training, mentoring, and office space that helped them sustain and consolidate their operations. However, the existence

of these organizations in all provinces and territories was not evident. Incubators and accelerators are organizations dedicated to supporting new ventures in establishing themselves during the early years of their operations.

Susan Cohen (2013) compared both programs and drew the following distinction: “incubators tend to nurture nascent ventures by buffering them from the environment to give them room to grow. In contrast, accelerators speed up market interactions to help nascent ventures adapt quickly and learn” (p. 21).

5.3.2.4 Angel Investors and Venture Capital (VC). Angel investors and venture capital are sources of funds that ODIs used to acquire the capital needed to develop their organizations further. Angel investors are “wealthy individuals who provide startup companies capital and are considered the first source of capital” (Giaquinto & Bruscatto Bortoluzzo, 2020; Morrissette, 2007). They are described as “informal venture capital-equity investments and non-collateral forms of lending made by private individuals, using their own money, directly in unquoted companies in which they have no family connection” (Edelman et al., 2017, p. 4).

Venture capital is a “specialist form of professionally managed finance designed to meet the financing needs of emergent firms, particularly in technology sectors, which are pursuing significant growth opportunities and which offer the prospects of high return” (Mason, 2020, p. 1).

While both sources of financing have some similarities, they have different motivations and approaches, which are reflected in the timing of investments and the level of risk considered

acceptable: “Venture capitalists invest in securing a financial return, while angel investors may have different motivations, such as coaching and mentoring entrepreneurs (Edelman et al., 2017, p. 9).

Several ODIs in this study secured angel and venture capital (VC) financing. However, three ODIs refused to consider this financing option, fearing a loss of control over their operations. In addition, ODIs operating in the NFP sector were not enthusiastic about the potential to secure any investments. Informants from these ODIs expressed frustration with the funding system in Canada, describing it as oblivious to companies operating in the sector.

5.4 Challenges and Barriers

This section further helps answer research sub-question *c* on challenges and barriers faced by ODIs (Section 1.5). The analysis of the interviews uncovered many challenges faced by ODIs during the early years of their operations. Some of the challenges were found to be shared among several intermediaries. Other challenges were unique and related to the business model adopted in a given case or specific business sectors. However, the data analysis revealed some critical challenges that affected the development and growth of all ODI startups. These were (a) organizational development, (b) fundraising, (c) product development, and (d) data acquisition.

5.4.1 Organizational Development

All participating ODIs began their operations with limited human resources centred on the founder(s); therefore, their first task was to build the teams responsible for developing their

products/services. However, recruiting qualified human resources was not an easy task due to two constraints: affordability and availability.

The informants cited data scientists, marketing, and sales professionals as expensive and in short supply. Due to the difficulty of hiring these professionals, ODI startups resorted to other options to fill their knowledge and experience gaps. They recruited students and independent developers, contracted specialized companies, and sought support from incubators and accelerators.

5.4.2 Fundraising

The second major challenge was fundraising, critical to the startups' survival and sustainability. Knowledge of where to go and how to apply for grants and funds was not something most ODIs possessed. They had to learn their way and gain the necessary experience and knowledge of the financial market and available instruments before they saw success. ODIs' constrained capacities and limited human resources at the early stage of their development made finding and applying for government grants and funds a challenging task and one that was, as noted above, already complicated by funding program requirements. Federal programs' application requirements, in particular, were found to favour large corporations. Likewise, when applying for accelerators and incubators programs, ODIs competed against many other businesses for the limited available funds, and their success was not guaranteed.

The third source of funds that ODIs could tap into was the capital market, which is considered the most challenging source. It demanded successful operations, proof of concept, and high

revenues, a set of requirements that ODIs could not immediately meet. Additionally, the capital market was unfavourable to startups operating in the NFP and government sectors.

5.4.3 Product Development

The third challenge found was building a viable product/service. Founders of ODI startups who transitioned from private- and public-sector jobs to their current roles discovered that their visions and plans for targeted products/services did not fit the actual needs of potential customers. In addition, their technical knowledge and skills proved insufficient to build viable products/services that received customers' acceptance. Consequently, they redesigned and reconfigured their products/services and built what was referred to as minimum viable products (MVPs) that received customers' acceptance and purchase approval.

5.4.4 Data Acquisition

The last and most crucial challenge ODI startups faced was acquiring the raw materials for their products/services. Finding, accessing, capturing, and preparing open and non-open data for reuse were the most challenging tasks ODIs performed. The raw data were captured from multiple sources, including government agencies and departments, private companies, the Internet, global governments' websites, and open data portals. The challenges varied from one source to another, requiring different actions to structure the data in a standard format that facilitated analysis. Over 30 challenges were faced by ODIs seeking to capture the needed data for their products/services. Many of the difficulties cited were related to data quality and metadata issues.

The challenges and barriers experienced by the 17 sampled ODI startups were not unique, though the startups differed in the raw materials they used. Scholars and professional organizations

studying startups' obstacles and barriers reported similar challenges to those uncovered in this study. For example, Giardino et al.'s (2015) analysis of the difficulties in early-stage software startups identified four key areas where challenges are reported: teamwork, product, finance, and market (pp. 57-59). Salamzadeh and Kesim's (2015) study on the startup life cycle and challenges reported four similar challenges: financial, human resources, support mechanism, and environmental elements (p. 7). Konsek-Ciechońska (2019), studying startup companies in Poland, identified three key challenges: financial, human resources, and support mechanisms (p. 1625).

Canadian scholars and professional organizations studying startups in Canada reported similar results. Rowe et al. (2019) found that startups experienced difficulties in four areas: accessing funding and talent, navigating the system, dealing with a lack of support for long-term development, and dealing with inefficient and ineffective government support. The authors indicated that Canada's risk-averse business culture is one of the challenges entrepreneurs face at all stages of development (p. 9). The Brookfield Institute for Innovation + Entrepreneurship (2019)³³ Identified four challenges that hindered the growth and expansion of Canadian startups: talent acquisition, access to capital, market access, and government incentives and support.

In a European Startup Monitor (ESM) report, Kollman et al. (2016) reported ten challenges European startups faced using OGD. The challenges covered team development, product development, fundraising, staff recruitment, sales and customer acquisitions, and internal organization processes (Kollmann et al., 2016). However, the challenges reported are found in the Canadian ODI startups participating in this study to different degrees and levels.

³³ <https://brookfieldinstitute.ca/>

While the challenges reported by the sampled Canadian ODIs were found in many studies, only the European Commission study on the reuse of open data provided empirical evidence of the difficulties of working with data. The study identified four key barriers to reusing open data (Berends et al., 2020): quality of open data, lack of standardization, and availability of open data. The report’s authors summarized its finding by stating “external, as well as internal barriers, do remain and hinder re-users from standardizing or automating the collection and processing of Open Data” (Berends et al., 2020, p. 2).

In conclusion, many of the challenges and barriers faced by the 17 sampled Canadian open data intermediaries are common among data intermediaries worldwide. However, what differentiates the situation of Canadian firms from those of their counterparts in Europe, the US, and the rest of the world are the business environment in Canada, the market size, governments’ programs and policies, and the capital market. Although starting a new business in Canada is easy, the growth and sustainability of the firm are determined by other internal and external factors. For example, Canada’s small and fragmented market size, modest capital market, and limited skilled human resources impact new startups’ survival ability and growth. Furthermore, while governments have created multiple support programs to assist entrepreneurs and small businesses with grants and funds, access to these programs is a complex process that new startups cannot accomplish easily. Thus, the success of Canadian ODI startups depends on overcoming the challenges and barriers reported in the study and summarized in Table 12:

Table 12: *Factors Affecting ODIs’ Success*

1	Organization and team building
2	Access to human and financial resources

3	Product/service development
4	Access to data (raw materials)
5	Marketing and sales

5.5 Data Quality and Metadata Issues and Challenges

This section further expands on the answer to research sub-question *c* on challenges and barriers faced by ODIs. The following sections discuss the data quality and metadata issues reported by informants and their impact on OGD use.

5.5.1 Data Quality. Data quality and one or more of its attributes, namely accessibility, accuracy, availability, structure, format, update, and aggregation, were reported by all the participating data intermediaries as impediments to using the datasets needed. This finding is consistent with the conclusion of many scholars (Beno et al., 2017; Immonen et al., 2014; Vetro et al., 2016; Zuiderwijk et al., 2012).

Data quality is critical for accurate results, as poor data quality might cause irreparable effects on the economy and society (Masip-Bruin et al., 2013). Therefore, ensuring data quality is of the utmost concern to people working with data. Consequently, extensive research has been conducted on this issue, with many definitions proposed, and many methodologies developed to assess the data quality and confirm its readiness. One formal definition has come from the International Organization for Standardization (ISO/IEC 25012), which has defined data quality as “the degree to which the characteristics of data satisfy stated and implied needs when used under specific conditions” (4.3). Wang and Strong (1996) defined data quality as “data that are fit for use by data consumers.” They extended the definition to data’s dimensions, described as a “set of data

quality attributes that represent a single aspect or construct of data quality” (p. 6). The UK government explained that “fit for use” meant “the data needs to be good enough to support the outcomes it is being used for. Data values should be right, but there are other factors that help ensure data meets the needs of its users” (Government of the UK, 2021). Redman (2012), writing for the *Harvard Business Review*, reflected on the importance of data quality as well:

There are two interesting moments in the lifetime of a piece of data: the moment it is created, and the moment it is used. Quality, the degree to which the data is fit-for-use, is judged at use. If it meets the needs at that moment, it is judged “high-quality,” and conversely. The whole point of data quality management is to connect those moments in time — to ensure that the moment of creation is designed and managed to create data correctly so everything goes well at the moment of use. (P.1)

With so many definitions of data quality that revolve around the users and usability aspect of the data, Batini et al. (2009) argued that although the literature provided a comprehensive classification of data quality dimensions, “discrepancies in the definition of most dimensions exist due to the contextual nature of quality” (18.6). The observation of Batini et al. resonates with the writing of Vetrò et al. (2016), who confirmed the absence of a comprehensive theoretical framework that can be used to assess data quality.

The need for a framework to assess open data quality was recognized by Berners-Lees (2006), who published the 5-star scheme for deploying open datasets. Vetrò et al. (2016) proposed

an OGD assessment framework consisting of six attributes: understandability, time aspects, accuracy, completeness, traceability, and compliance. Wang & Strong (1996) also addressed the data quality issue. They proposed a conceptual framework based on four categories and their associated dimensions: intrinsic data quality, contextual data quality, presentational data quality, and accessibility data quality (p. 18).

In an analysis of 13 data quality assessment and improvement methodologies by Batini et al. (2009), the authors concluded that the data quality research field is evolving and cannot be considered mature. Furthermore, the authors indicated that no general agreement exists on which set of dimensions defines the quality of data or the exact meaning of each dimension.

In a blog³⁴ post published on May 31, 2017, The Open Knowledge Foundation explained that there are many different interpretations and ways of measuring data quality. The foundation listed nine attributes that should be considered when assessing the data quality: data interpretability, data accuracy, timeliness of publication, reliability, trustworthiness, accessibility, discoverability, processability and completeness. All these impacts the use of OGD.

A comprehensive quality framework encompassing all the categories, attributes and dimensions associated with data has yet to be developed. However, combining all these elements in one framework that meets the needs of diverse users is challenging. Therefore, the negative effect of uncertain data quality will continue to impact users, society, and the economy (Beno et al., 2017a; Immonen et al., 2014; Sadiq & Indulska, 2017; Zuiderwijk & Janssen, 2014). Sadiq &

³⁴ <https://blog.okfn.org/2017/05/31/open-data-quality-the-next-shift-in-open-data/>

Indulska (2017) considered the quality problems with the published open datasets to threaten the value generated from the data.

5.5.2 Metadata. One of the challenges reported by the ODIs that participated in this study was with the metadata of the datasets available on open data portals. Absent, incomplete, or incorrect metadata impacts users' performance and reduces efficiency (Neumaier et al., 2016). Therefore, metadata is crucial for the description and classification of data and for ensuring their long-term value (Schauppenlehner & Muhar, 2018).

Metadata is “data about data” (Loshin, 2013b; Riley, 2017). Riley (2017) has stated:

Metadata is pervasive in information systems and comes in many forms. The core features of most software packages we use every day are metadata driven. People listen to music through Spotify; post photos on Instagram; locate a video on YouTube; manage finances through Quicken; connect with others via email, text, and social media; and store lengthy contact lists on their mobile devices. (p. 2)

There are many other definitions of metadata as well. For example, from the perspective of business intelligence (BI), Loshin described metadata as a “catalogue of the intellectual capital surrounding the creation, management, and use of a dataset” (p. 120). From a data management view, metadata describes a dataset: how it was collected, when it was collected; what assumptions were made in its methodology; its geographic scope; if there are multiple files, how they relate to one another; the definitions of individual variables and, if applicable, what possible answers were;

the calibration of any equipment used in data collection; and the version of the software used for analysis³⁵.

Metadata is critical in finding, selecting, and understanding the datasets in open data portals. Access to data requires comprehensive metadata structures, high-quality metadata and user-centred interfaces for search operations and data downloads (Schauppenlehner & Muhar, 2018). Neumaier et al. (2016) noted that “for the data to be discoverable by consumers, the publishers need to describe their data accurately and comprehensively” (p. 2:2). Conversely, if metadata is not defined correctly with enough quality, it hampers the discoverability and accessibility of resources through open data portals (Nogueras-Iso et al., 2021, p. 60365). Lisowska (2016) described the challenges facing open data users:

Data scientists, journalists and analysts are often left with the daunting and time-consuming task of finding relevant data, discovering new datasets, and, most importantly, understanding it before any analysis can be done. Such information should be found in the metadata, which should pair with the published data. (p. 1)

There are many standards to ensure that quality metadata is published. Lisowska (2016) explained that metadata generation is dictated by the type of data published, such as statistical, geographical, or financial, linked to the metadata standards they employ. The author considered that “These standards are a direct response to the need for descriptive, structured information on the data published on any given subject” (p. 4).

³⁵ <https://guides.lib.unc.edu/metadata/definition>

The availability of many metadata standards poses an interoperability problem, impacting data use across platforms. Because of their importance in facilitating the finding, accessing, understanding, and use of datasets, Lisowska (2016) called for a standardized way of publishing metadata through open data portals.

5.6 OGD, ODIs, and Value Creation

This section answers research sub-question *d* on what value is created, and how (Section 1.5). “The creation of value is the core purpose and central process of economic exchange” (Vargo et al., 2008).

The purpose of a business is to create customers and generate revenue. Therefore, it must offer something beneficial, a product or service that captures customers’ attention and results in a transaction that creates value for all parties. The 17 ODIs that participated in this study demonstrated different levels of success in creating value for their customers, societies, and themselves since 2010. The value created is evidenced by their survival for more than four years of existence. That survival resulted from their income and growth, as demonstrated by the number of jobs created.

The findings revealed that ODI participants in the study generated direct and indirect benefits. The direct benefits were monetary, exemplified by the revenue generated and the jobs created. These benefits aligned with those identified in the European Commission report on reusing open data (2020). The report also indicated cost savings as an indirect benefit ODIs reported

creating for the users of their products/services, along with time savings, increased efficiency, enhanced trust, and life-saving information.

5.6.1 Direct Benefits

Ten ODIs indicated that their business operations had surpassed the break-even point and were generating profits. Four others showed their operations were on the verge of recording profits; however, the COVID-19 pandemic affected their performance and growth. Nevertheless, the revenues realized enabled ODIs to invest in their operations and hire administrative, technical and sales personnel to develop their products/services, manage operations, and capture new customers. As a result, the number of jobs they created reached 217 by 2021, evidence that open and non-open government data contributes to job creation.

5.6.2 Indirect Benefits

All ODIs reported various benefits realized by their customers and society, the most common of which was increased efficiency, time, and cost savings. Additional indirect benefits included strengthened trust between government and citizens, increased public participation, environmental benefits, and stimulation of new businesses. Such benefits, which cannot be monetized and are beyond the scope of this paper, were also confirmed by the European Commission's report *Creating Value Through Open Data*. The report described the indirect benefits generated from the reuse of open data: "economic benefits, new goods and services, time savings for users of applications using Open Data, knowledge economy growth, increased efficiency in public services and growth of related markets" (p. 8).

5.6.3 Value Creation Process

This study revealed that OGD is essential to the value created by ODIs products/services but is not its sole source. The value creation process requires more than open data. It needs users with the right competencies to work with the data and transform it into products and services that create value when customers accept, purchase, and use it. Therefore, the data intermediary that uses OGD, among other data types, to create/develop products/services and generate value is essential.

Creating value from OGD includes four stages that must be completed before reaching the value that can be generated. Each stage consists of multiple activities, and when complete, it becomes an input for the next one. Figure 13 illustrates the process used by 12 open data intermediaries.



Figure 13: *The Process of Value Creation from Data*

Twelve of the data intermediaries in this study followed the process presented above. Another five, described as custom builders of solutions, first started the value creation process through marketing and sales activities, then, after securing contracts with customers, proceeded through the other three categories/stages, data preparation, data transformation, and product/service development, before reaching the value creation stage.

The data preparation stage includes finding the needed datasets and accessing, downloading, cleaning, structuring, and formatting the data in a standard way. Finally, it is ready to be used as input to the transformation stage. Gartner (2021) defined data preparation as “an iterative-agile process for exploring, combining, cleaning and transforming raw data into curated datasets for self-service data integration, data science, data discovery, and BI/analytics.” Each activity in the data preparation stage feeds to the next, ending in ready-to-use datasets in a common standard. Finding datasets is crucial to meeting the use objectives of ODIs. Accessing and downloading are additional activities performed sequentially at the sites where the datasets are found. This process begins with searching the web for a specific dataset. Then, it requires software tools to automatically scan multiple, sometimes hundreds, websites and data portals to find the needed datasets. Chapman et al. (2020) described dataset searching as “the discovery, exploration, and return of the dataset to an end-user” (p. 252).

There are legal rights and permissions required to access data. These include the ability to examine data stored within a database or repository. Those granted access privileges can retrieve, store, move, or manipulate data across various systems (*Law Insider*, 2021; OSTHUS, 2021). Downloading is transmitting a file or data from one computer to another over a network. All these activities require users to have data science knowledge and skills to perform the tasks successfully.

Cleaning, formatting, and structuring are performed on the datasets captured from the web. Data cleaning is “fixing or removing incorrect, corrupted, incorrectly formatted, duplicate, or incomplete data within a dataset” (Tableau Software, 2021). Data formatting is organizing data according to pre-set specifications for computer processing. The format is defined as “the structure

of data within a database or file system that gives the information its meaning” (C3 AI, 2021). Finally, data structuring is the process of “storing data in a standardized format, with a well-defined structure that complies with a data model, follows a persistent order, and is easily accessed by humans and programs” (TIBCO Software, 2021b).

The transformation stage involves linking open datasets or records and combining them with proprietary and other data, when needed, to form a database of information that will be made available as a source for the product/service development stage. Stitch Software (2021) has defined data transformation as:

The process of changing the format, structure, or values of data; maybe constructive (adding, copying, and replicating data), destructive (deleting fields and records), aesthetic (standardizing salutations or street names), or structural (renaming, moving, and combining columns in a database. (Stitch Software, 2021)

Data linkage is “a method of bringing information from different sources about the same person or entity to create a new, richer dataset” (University of Tasmania, 2022). In contrast, data merging is “combining two or more similar records into one. For example, merging is done to add variables to a dataset, append or add cases or observations, or remove duplicates and other incorrect information” (TIBCO Software, 2021).

Database and data warehouses are the repositories where data is stored for further processing. A database is “an organized collection of structured information, or data, typically stored

electronically in a computer system” (Oracle, 2021b). A data warehouse, on the other hand, is defined as “a type of data management system that is designed to enable and support business intelligence (BI) activities, especially analytics” (Oracle, 2021a). The database holds structured data, while the data in a data warehouse can be both, structured and unstructured.

The product/service development stage is a process that takes a product from concept to market availability. The product is a software application tool that transforms OGD into valuable products/services in the open data ecosystem. IBM has defined software development as “a set of computer science activities dedicated to creating, designing, deploying, and supporting software. Software development is primarily conducted by programmers, software engineers and software developers. However, these roles interact and overlap, and their dynamics vary greatly across development departments and communities” (IBM, 2021d).

The product/service development stage is vital to meeting an ODI’s final value creation objective. The study identified four products/services developed by the participating Canadian data intermediaries: industry-specific data portals, custom-made solutions, insights and knowledge, and standard application software. Developing the products/services requires data science expertise, experience, and profound knowledge of target customers’ needs and requirements. Failure to understand the customer and the market needs and requirements result in product/service rejection and loss.

The marketing and sales stage is critical to the success of ODIs. Without it, potential customers would not know about the products/services; hence, no sales, revenue, or value would be

created. Marketing creates awareness, generates leads, and prepares the market for efforts to transform leads into sales that generate revenue. ODIs with only technical expertise that recognize the importance of this stage and take the necessary measures to ensure customers/markets' awareness of the products/services and their value have a higher chance of achieving their objectives. Conversely, ODIs that fail to recognize the importance of marketing and sales and depend on their limited knowledge and skills to sell the product/service have difficulty realizing the benefits they could have achieved otherwise.

Finally, the value creation stage is the stage that reveals evidence of the value created by the data products/services sales. Two value types are purportedly created using OGD: economic and social value. Economic value is evidenced by the number of jobs created and the revenue generated by ODIs. It is also found in customers' organizations through increased efficiencies and cost reduction; however, it is difficult to measure. There are also suggestions that OGD-based products/services sales have contributed to social benefits. For example, trust between the government and the public, leading to increased citizen participation in local government affairs, was identified in this study as one result of OGD-based product/service use. Saving lives was another. The social and economic benefits customers generate from data products/services are numerous but difficult to identify and measure efficiently. Figure 14 depicts the value creation process from open data and the activities performed at each stage by 12 of the data intermediaries.

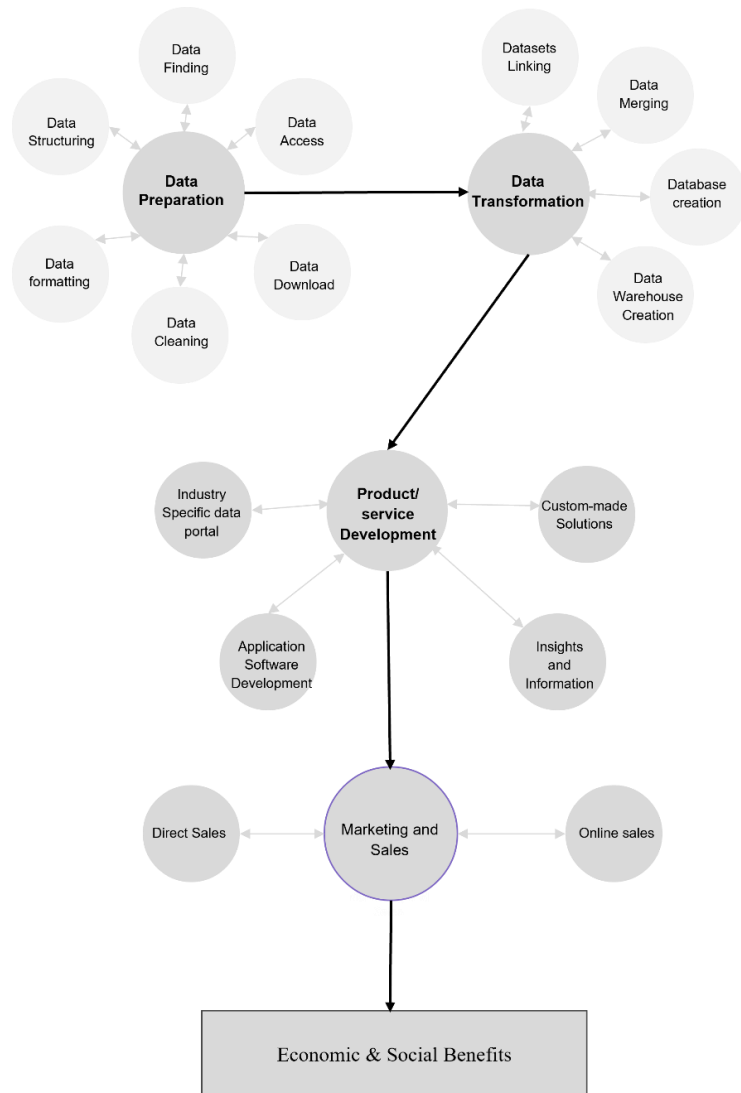


Figure 14: *Expanded Value Creation Process*

The evidence discovered in this study confirms that OGD when used by competent users with access to capital, leads to economic value creation. Accordingly, it can be established that economic value creation from OGD is contingent on open data intermediaries possessing the business experience, data science competencies, marketing and sales competencies, and access to the capital market. When these four elements are found in an ODI organization, the chances of its creating economic and social value from OGD increase.

Chapter 6. Conclusions, Implications, Limitations, and Future Work

6.1 Conclusions and Insights

This research was an explorative, inductive study to discover how data intermediaries use OGD to create products/services that contribute to economic value creation. A qualitative research approach was used to design the data collection, analysis, and reporting. With this approach, rooted in grounded theory, 20 interviews with 17 open data intermediary firms were coded and analyzed in a way that allowed the researcher to document patterns emerging from the data.

A qualitative inquiry was used because it seeks to discover rather than test variables and allows researchers to get at the inner experience of participants to determine how meanings are formed through and in culture (Corbin & Strauss, 2012).

The research commenced with a theoretical sampling of potential users of OGD, followed by a purposive selection of data intermediary representatives for an invitation to participate in the study. When an acceptance was received, an interview followed, and data was collected. Following data collection and transcription of the interview, the information coding started immediately, and a list of initial categories was developed. For every successive interview, the data coding went through a constant comparative analysis with the previously developed categories to avoid repetitiveness and ensure proper representation. Following coding the 20 interviews and developing initial categories, categories were compared and merged into higher-level ones through inductive and deductive thinking. The constant comparative analysis continued with the categories evaluation to find consistencies and differences, enabling the researcher to generate abstract concepts. According to Corbin and Strauss (1990), concepts are the basic units of analysis as theories cannot

be built with actual incidents or activities reported (Corbin & Strauss, 1990). The coding of the interviews went through the three stages advocated by Corbin and Strauss (1990): initial coding, axial coding, and selective coding. In the selective coding stage, all categories are unified around central core categories representing the study's significant phenomenon. The core categories/concepts identified in the selective coding process were then theorized as insights that required further investigations in future research.

This research aimed to explore and evaluate the use of OGD in Canada and discover its implications and contribution to value creation. Open data intermediaries were identified as the primary users of OGD; therefore, the research focused on them and their activities and uses of the data to create products/services that have the potential to generate economic value.

The thesis began by introducing the key elements that are the focus of this research: OGD, ODIs, and value creation. A literature review followed to discover how these topics and their relationships were discussed in the literature. The aim was to understand the concept of OGD, the objectives of open data intermediaries, and the processes they used to create products/services that generate economic value. Chapter 3 outlined the research methodology and design and the process used to identify case studies and collect and analyze data.

The study found that the ODI firms were purposefully founded by individuals who considered data a valuable raw material with the potential to create economic and social value. Armed with their experience, knowledge, and skills, ODI founders built their organizations to provide

products/services based on data to the public, private, and NFP sectors. They used a mix of open and non-open data selected based on the need of their customers.

In interviews, founders noted that building new organizations requires, among other things, experienced staff, and sufficient funds to support and grow the organizations during their early years. Having one of these elements without the other may be insufficient to fulfill an organization's objectives. Experienced staff in the open data ecosystem include people with business and technical knowledge and skills in information and communication technology critical to identifying, collecting, curating, and processing open data and creating products/services. However, technical knowledge and skills are insufficient to build successful businesses. The technical skills must be accompanied by business knowledge, marketing, and sales skills that ensure the products/services capture potential customers' attention and generate the expected revenue.

To sustain their operations and build their products/services, ODI organizations explored all available financial avenues, including support from families and friends, government funds and grants, incubators and accelerators, and angel and venture capital investment. As a result, not all ODIs experienced the same level of success; however, after ten years or less of operations, all participating ODIs established their operations, built valuable products/services, and generated revenue. Some grow their operations using income from their products/services; others still depend on government funds and grants. Unfortunately, few received investments from the capital market.

This study is the first empirically based research on Canada's OGD value creation potential. Users of OGD were studied to determine their capabilities and the competencies required to

enable value creation from OGD. The study discovered issues and problems plaguing the published OGD and the limited support available to ODI startups from government agencies and departments. However, it also confirmed that economic value creation from OGD is possible if and when the issues and challenges plaguing this new industry are mitigated. Several patterns and themes emerged that shaped the conclusions of this study. They led to the following theoretical insights, which are expected to advance theories building on value creation from OGD uses from the perspective of Canadian ODIs.

1. Most of the participating Canadian ODIs did not sufficiently understand the concept of OGD. Some viewed the unstructured data made available by government departments and agencies and data found on governments' websites and the Internet as open data, which is false. This alarming lack of knowledge of what constitutes OGD led to the following insight:

Knowledge of open government data and governments' data portals among participating Canadian users is inadequate.

2. The quality of available OGD affected intermediaries' use and dependence on the data. Hence, data intermediaries looked beyond governments' data portals to acquire additional data and combined it with OGD to produce the final products/services. Thus, the following insight emerged:

Open government data is an enabling resource. However, its quality and availability impact users' ability to depend on the data as a sole source for value creation.

3. ODIs working with OGD indicated that published open datasets are not easy to find or use; hence, it is difficult for the public to explore and benefit from them. Furthermore, exploring open data and using it would be impossible without formal training and experience in computer and information technology. This finding generated the third insight:

Competencies in data science are essential to work with open government data efficiently.

4. The published OGDs were found to have quality issues. ODIs reported many challenges and barriers arising from the datasets, including different standards, different/wrong metadata, structures, and high aggregations that make open data challenging to use and explore. In addition, ODIs reported encountering various formats and designs for the same datasets used by local, provincial, and federal governments between provinces, departments, and agencies. They also discovered that some datasets have validity issues, and others are unavailable on the portals. These difficulties hampered ODIs' productivity requiring further interventions to correct and use. This information produced the fourth insight:

A common standard for structuring and publishing open datasets among Canadian government departments and agencies is missing.

5. Open government data is valuable raw material but does not produce value alone. It needs to be used to generate the value anticipated. The efficient use of OGD is due to the resources and capabilities that ODIs have managed to acquire or develop and use

to create economic value. The value created is evident in the growth experienced by many participating ODI firms, the jobs created, and the revenue generated. The business activities of ODIs also have an indirect economic impact through increased efficiency in their clients' operations, cost savings, and higher productivity. Recognition of these factors led to the fifth insight:

Participating Canadian ODIs using open government data exhibited the potential to contribute to economic growth and job creation.

These insights are critical but preliminary and cannot be generalized to the Canadian space without further empirical research and analysis. However, two insights, the first and the fifth are crucial to advance theories on ODIs' roles in creating value from OGD. First, awareness of the open government data initiative among citizens and its effect on citizens' participation is essential to increase their involvement in the open data ecosystem and, thus, in value creation. Further research into and investigation of this challenge should lead to a theory on the impact of citizens' awareness on OGD value-creation potential. Increasing citizens' participation in the open data ecosystem should significantly impact the use of OGD. Evans et al. (2014) argued that "the data-driven focus had not been proven to significantly increase citizen understanding of the complexities of issues and policies or their participation" (p. 172). They called for a re-evaluation of the initiative and recommended that governments explore new approaches if the primary goal of open government is to engage citizens (Evans et al., 2014).

Insight five corroborates the view that value creation from using OGD is feasible. However, value creation from OGD depends on ODI organizations' resources and capabilities. Thus,

further research on ODIs' resources and capabilities should advance theories of their roles and contribution to value creation from OGD. The resources and capabilities of ODIs are critical enablers in working with data; they are essential for value creation and increasing a firm's performance (Barney, 1991).

The insights into data qualities and publishing standards presented here are also critical, and further research into the impact of these factors on value creation is needed. However, it should be noted that the open government phenomenon is nascent, and the work on releasing government data according to the standard is still in its early years and needs more time to develop and reach a satisfactory level. In its global report (2017), the Open Data Barometer found that only 7% of the data is fully open, one of every two datasets is machine-readable, and only one in four datasets has an open licence (p. 12). Additionally, the report pointed out incomplete and low-quality datasets and raised concerns that governments are not publishing the data needed to restore citizens' trust. The report concluded with the statement:

Most governments are not meeting the basic Open Data Charter principles. In most cases, the right policies are not in place, nor is the breadth and quality of the datasets released sufficient. Thus, this means we cannot collectively use open data to change people's lives for the better truly. (p. 11)

This thesis contributes to the knowledge of the state of OGD and open data intermediaries in Canada. In addition, it highlights the existing challenges that impact ODIs' performance and use of OGD, and the issues associated with that data. Furthermore, the findings provide essential

knowledge about the problems and challenges of working with OGD to new data intermediary firms entering that domain. Finally, this thesis will also be of value to future researchers studying the use of OGD.

6.2 Implications

The insights presented above are critical, and the issues they highlight would have significant negative implications for the use of and value creation from OGD if left unaddressed. Therefore, concerted effort by all levels of government in Canada is needed to mitigate these problems.

Citizens' awareness of the open government data initiative, what it constitutes, and its benefits and objectives is critical to augment the use of such data. If no efforts are made to increase citizens' knowledge of the purpose of the government's initiatives to release public data in an open format and the benefits expected from using the open datasets released on data portals, the use of the open data will not increase. Failure to entice citizens to participate in the open government data program will defeat the purpose of opening government data to increase public participation, government transparency, and accountability and stimulate economic growth.

Additionally, the limited use of OGD is an issue its providers must address. Open government data intermediaries reported many problems with the available datasets. These issues must be addressed to ensure that the published datasets are of good quality and can be used. The current availability, accessibility, and quality issues hinder users from using the available open datasets and drive them away from the government's data portals. Consequently, more efforts should be made to ensure data quality, leading to greater use of OGD.

Data competencies are critical to the use of OGD. Thus, such data is not for all, even when claimed to be for all citizens. Even when the data is in an open format, accessible, and license-free, most citizens would have difficulty accessing and using the data. Furthermore, the availability of competent users was low, affecting firms' ability to find and recruit people with the necessary skills and knowledge. This shortage of people with data science knowledge and skills directly impacts the use of OGD and, subsequently, the creation of value from the data. Hence, government programs to increase citizens' digital literacy and efforts by educational institutions to create programs that educate students in data science would enhance the use of OGD.

The absence of a common standard for structuring and reporting open government datasets among all levels of government renders the work of the users of OGD complex, inefficient, and costly and impacts their use of that data. When trying to link multiple datasets from different and/or the same level of government, users face the challenge of restructuring and rebuilding the datasets and finding the owners who can provide information on the structure of the datasets. Again, governments at all levels must collaborate to have their data published in a standard, structured format to enable efficient use of the datasets.

The value creation potential of OGD use by open data intermediaries is evident in the study. However, the number of identified open data intermediary startup firms in Canada is small, which is alarming. All levels of government should address this issue by creating programs to entice individuals and organizations to enter this field and by ensuring adequate support is available for startup firms. Only when the number of users of OGD grows will a more significant impact on the economy be realized.

Increasing the use of OGD and generating substantial economic value from its usage will require greater efforts by all levels of government to mitigate the issues and problems facing its use and development.

6.3 Limitations

This study has several limitations, which are attributable to multiple factors. First, OGD and ODIs are nascent phenomena, about ten years old, and the concept of OGD is still under development. Hence, the number of Canadian OGD users is limited and inadequately researched and reported. Second, the participants were selected on the assumption that they were homogeneous users of OGD. However, the analysis found that they are heterogeneous ones, with differences in their knowledge of OGD, business approaches, objectives, and uses of OGD. Third, the COVID-19 pandemic's arrival coincided with the start of the recruitment stage, affected firms' willingness to participate in the study, and limited the researcher's data collection efforts. Fourth, the closure of offices and businesses due to the pandemic changed the plan to conduct face-to-face interviews, which resulted in doing the interviews online over the Internet, which created problems of degradations and interruptions in communication that impacted the quality of the data collected. Fifth, theoretical saturation was arbitrarily considered reached due to the difficulty of recruiting new ODI users.

The final limitation concerns the generation of theory. The small sample of Canadian ODIs who accepted to participate in this research inhibited this thesis from generating a theory covering the use and value creation prospects of OGD in Canada. These limitations underscore the

possibility that further nuances could have shaped the discovery of patterns and insights from using OGD by data intermediaries.

6.4 Future Research

There are several opportunities for future research to address:

- Assess what the different levels of government in Canada are doing to promote awareness and use of open government data and explore their use of standards in that context.
- Expand the study of Canadian data intermediaries' use of open government data (and the underlying coding of concepts) to a larger sample to enable the generation of a theory covering economic value creation.
- Explore the social benefits of using open government data.

References

- Agogu , M., Berthet, E., Fredberg, T., Le Masson, P., Segrestin, B., Stoetzel, M., ... Ystr m, A. (2017). Explicating the role of innovation intermediaries in the “unknown”: a contingency approach. *Journal of Strategy and Management*, 10(1), 19–39. <https://doi.org/10.1108/JSMA-01-2015-0005>
- Al-Debei, M. M., & Avison, D. (2010). Developing a unified framework of the business model concept. *European Journal of Information Systems*, 19(3), 359–376. <https://doi.org/10.1057/ejis.2010.21>
- Alexopoulos, C., Loukis, E., & Charalabidis, Y. (2016). A Methodology for Determining the Value Generation Mechanism and the Improvement Priorities of Open Government Data Systems. *Computer Science and Information Systems*, 13(1), 237–258. <https://doi.org/10.2298/CSIS151210001A>
- Ambrosini, V., Bowman, C., & Collier, N. (2009). Dynamic capabilities: An exploration of how firms renew their resource base. *British Journal of Management*, 20 (SUPP. 1). <https://doi.org/10.1111/J.1467-8551.2008.00610.X>
- Amit, R., & Schoemaker, P. J. H. (1993). Strategic Assets and Organizational Rent. *Strategic Management Journal*, 14(1), 33–46.
- Attard, J., Orlandi, F., & Auer, S. (2016). Value Creation on Open Government Data. In *2016 49th Hawaii International Conference on System Sciences (HICSS)* (pp. 2605–2614). IEEE <https://doi.org/10.1109/HICSS.2016.326>
- Attard, J., Orlandi, F., Scerri, S., & Auer, S. (2015). A systematic review of open government data initiatives. *Government Information Quarterly*, 32(4), 399–418. <https://doi.org/10.1016/j.giq.2015.07.006>

- Augier, M., & Teece, D. J. (2007). Dynamic Capabilities and Multinational Enterprise: Penrosean Insights and Omissions. *Management International Review*, 47(2), 175–192.
- Baldrige, R., & Curry, B. (2021). What Is A Startup? How Do Startups Work? *Forbes Advisor*. Retrieved from <https://www.forbes.com/advisor/investing/what-is-a-startup/>
- Barney, J. (1991). Firm Resources and Sustained Competitive Advantage. *Journal of Management*, 17(1), 99–120. <https://doi.org/10.1177/014920639101700108>
- Barreto, I. (2010). Dynamic Capabilities: A review of past research and an agenda for the future. *Journal of Management*. <https://doi.org/10.1177/0149206309350776>
- Barry, E., & Bannister, F. (2014). Barriers to open data release: A view from the top. *Information Polity*, 19(1–2), 129–152. <https://doi.org/10.3233/IP-140327>
- Baškarada, S. (2014). Qualitative Case Study Guidelines. *The Qualitative Report*. <https://doi.org/10.46743/2160-3715/2014.1008>
- Batini, C., Cappiello, C., Francalanci, C., & Maurino, A. (2009). Methodologies for data quality assessment and improvement. *ACM Computing Surveys*, 41(3), 16. <https://doi.org/10.1145/1541880.1541883>
- Baxter, P., & Jack, S. (2008). Qualitative Case Study Methodology: Study Design and Implementation for Novice Researchers. *The Qualitative Report*, 13(4), 544–559. Retrieved from <https://nsuworks.nova.edu/tqr/vol13/iss4/2>
- BDC. (n.d.). What is a business accelerator? Retrieved September 6, 2021, from <https://www.bdc.ca/en/articles-tools/entrepreneur-toolkit/templates-business-guides/glossary/business-accelerator>

- Bell, J. (2015). What is Machine Learning? *Machine Learning* (pp. 01–16). John Wiley & Sons, Inc. Retrieved from <https://books.scholarsportal.info/uri/ebooks/>
- Beno, M., Figl, K., Umbrich, J., & Polleres, A. (2017a). Open data hopes and fears: Determining the barriers of open data. *In Proceedings of the 7th International Conference for E-Democracy and Open Government, CeDEM 2017* (pp. 69–81). <https://doi.org/10.1109/CeDEM.2017.22>
- Beno, M., Figl, K., Umbrich, J., & Polleres, A. (2017b). Perception of Key Barriers in Using and Publishing Open Data. *JeDEM - EJournal of EDemocracy and Open Government*, 9(2), 134–165. <https://doi.org/10.29379/jedem.v9i2.465>
- Berends, J., Carrara, W., Engbers, W., & Vollers, H. (2020). Re-using open data. Luxembourg. <https://doi.org/10.2830/876679>
- Berends, J., Carrara, W., & Radu, C. (2017). Analytical Report 9: The Economic Benefits of Open Data. Retrieved from <http://www.europeandataportal.eu/>
- Berners-Lees Tim. (2006). Linked Data - Design Issues. Retrieved November 8, 2022, from <https://www.w3.org/DesignIssues/LinkedData.html>
- Blank, S. G. (2007). The Four Steps to epiphany (Third Edition).
- Bowen, G. A. (2008). Naturalistic inquiry and the saturation concept: A research note. *Qualitative Research*, 8(1), 137–152. <https://doi.org/10.1177/1468794107085301>
- Bowman, C., & Ambrosini, V. (2000). Value Creation Versus Value Capture: Towards a Coherent Definition of Value in Strategy. *British Journal of Management*, 11(1), 1–15. <https://doi.org/10.1111/1467-8551.00147>

- Bowman, C., & Ambrosini, V. (2003). How the Resource-based and the Dynamic Capability Views of the Firm Inform Corporate-level Strategy. *British Journal of Management*, 14(4), 289–303.
<https://doi.org/10.1111/J.1467-8551.2003.00380.X>
- Brown, L. D., & Kalegaonkar, A. (2002). Support organizations and the evolution of the NGO sector. *Nonprofit and Voluntary Sector Quarterly*, 31(2), 231–258.
<https://doi.org/10.1177/08964002031002004>
- C3 AI. (2021). Data Format. Retrieved April 25, 2022, from <https://c3.ai/glossary/features/data-format/>
- Canada Treasury Board. (2014). Directive on open government. www.open.canada.ca. Treasury Board of Canada, Secretariat.
- Cardon, M., Cardon, M. S., & Stevens, C. E. (2004). Managing human resources in small organizations: What do we know? *Human Resource Management Review*, 14, 295–323.
<https://doi.org/10.1016/j.hrmr.2004.06.001>
- Carrara, W., San Chan, W., Fischer, S., & Van Steenberg, E. (2015). Creating Value through Open Data: Study on the Impact of Re-use of Public Data Resources. <https://doi.org/10.2759/328101>
- Chambers, E., & Patrocinio, M. (2011). Business Models and Value Creation: A case study of the New York economic development corporation. <https://doi.org/10.1007/s00373-002-0512-x>
- Chapman, A., Simperl, E., Koesten, L., Konstantinidis, G., Ibáñez, L. D., Kacprzak, E., & Groth, P. (2020). Dataset search: a survey. *VLDB Journal*, 29(1), 251–272.
<https://doi.org/10.1007/S00778-019-00564-X>

- Charalabidis, Y., Zuiderwijk, A., Alexopoulos, C., Janssen, M., Höchtl, J., & Ferro, E. (2018). The World of Open Data. *Springer (Vol. 28). Cham: Springer International Publishing.*
<https://doi.org/10.1007/978-3-319-90850-2>
- UK Government Cabinet Office. (2013). Open Data Charter and Technical Annex. Retrieved December 19, 2018, from <https://www.gov.uk/government/publications/open-data-charter/g8-open-data-charter-and-technical-annex>
- Chignard, S. (2013). A brief history of Open Data. Retrieved November 1, 2022, from <http://www.paris-techreview.com/2013/03/29/brief-history-open-data/>
- Cohen, S. L. (2013). What Do Accelerators Do? Insights from Incubators and Angels.
https://doi.org/10.1162/inov_a_00184
- Connor, T. (2002). The resource-based view of strategy and its value to practising managers. *Strategic Change*, 11(6), 307–316. <https://doi.org/10.1002/jsc.593>
- Corbin, J. M., & Strauss, A. (1990). Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative Sociology*, 13(1), 3–21. <https://doi.org/10.1007/BF00988593>
- Corbin, J., & Strauss, A. (2012). Basics of Qualitative Research (3rd ed.): Techniques and Procedures for Developing Grounded Theory. SAGE Publications, Inc.
<https://doi.org/10.4135/9781452230153>
- Creswell, J. W. (2007). Qualitative Inquiry and Research Design: Choosing Among Five Approaches. *Sage Publications, Inc. (2nd edition, Vol. 16)*. Sage. <https://doi.org/10.1177/1524839915580941>
- Creswell, J. W. (2009). Research Design: Qualitative, Quantitative and Mixed Methods Approaches. SAGE (3rd Edition). *SAGE Publications, Inc.* <https://doi.org/10.1080/14675980902922143>

- Corbin, J., & Strauss, A. (1990). Grounded Theory Research: Procedures, Canons, and Evaluative Criteria. *Zeitschrift Fur Soziologie*, 19(6), 418–427. Retrieved from <http://www.zfs-online.org/index.php/zfs/article/viewFile/2741/2278>
- Corbin, J., & Strauss, A. (2012). Basics of Qualitative Research (3rd ed.): Techniques and Procedures for Developing Grounded Theory. *SAGE Publications, Inc.*
<https://doi.org/10.4135/9781452230153>
- Creswell, J. W. (2013). Qualitative Inquiry & Research Design: Choosing Among. *SAGE Publications, Inc.* (3rd edition, Vol. 1).
- Creswell, J. W. (2014). Research Design (4th ed.). *SAGE Publications Ltd.*
- Crowe, S., Cresswell, K., Robertson, A., Huby, G., Avery, A., & Sheikh, A. (2011). The case study approach. *BMC Medical Research Methodology*, 11(1), 100. <https://doi.org/10.1186/1471-2288-11-100>
- Csorba, Z. (2020). Startup: Hype or Tendency? *Journal of Organizational Culture, Communications and Conflict*, 24(3). Retrieved from
<https://www.proquest.com/docview/2519439465/fulltextPDF/3B547A88AEC1455BPQ/1?accountid=14701>
- Curtin, G. G., Sommer, M. H., & Vis-Sommer, V. (2013). The World of E-Government. *The Hawthorth Press, Inc.* https://doi.org/10.1300/J199v02n03_01
- Davies, T. (2010). Open data, democracy, and public sector reform: A look at open government data use from data. gov. UK. Retrieved from <http://www.practicalparticipation.co.uk/odi/report/>

- Davies, T. (2014). Open Data in Developing Countries: Emerging Insights from Phase I. *The World Wide Web Foundation*, 1–38.
- Davies, T., & Perini, F. (2016). Researching the emerging impacts of open data ODDC conceptual framework. *Journal of Community Informatics*, 12(2), 148–178.
<https://doi.org/10.15353/joci.v12i2.3246>
- Davies, T., Walker, S. B., Rubinstein, M., & Perini, F. (2019). The State of Open Data - Histories and Horizons. African Minds. <https://doi.org/10.5281/ZENODO.2668475>
- Deloitte. (2012). Open Data, Driving Growth, Ingenuity, and Innovation. Retrieved from <https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/deloitte-analytics/open-data-driving-growth-ingenuity-and-innovation.pdf>
- Deloitte LLP. (2012). Open growth Stimulating demand for open data in the UK. A briefing note from Deloitte Analytics. Deloitte, 1–9. Retrieved from <https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/deloitte-analytics/open-growth.pdf>
- Denzin, N. K., & Lincoln, Y. S. (2018). The SAGE Handbook of Qualitative Research. (N. K. Denzin & Y. S. Lincoln, Eds.) (Fifth). *SAGE Publications, Inc.*
- Macmillan Dictionary. (n.d.). Supplier (noun) definition. Retrieved March 4, 2022, from <https://www.macmillandictionary.com/dictionary/british/supplier>
- Directive on Open Government- Canada.ca. (n.d.). Retrieved from <https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=28108>
- Dumpawar, S. (2015). Open Government Data Intermediaries: Mediating Data to Drive Changes in the Built Environment. PGCPD. Retrieved from

https://pdfs.semanticscholar.org/2bba/16f65301d13af906d25b1395623a87f77412.pdf?_ga=2.11798711.220178741.1565024436-1991950439.1552315267

Dunn, Thomas and Holtz-Eakin, Douglas, “Financial Capital, Human Capital, and the Transition to Self-Employment: Evidence from Intergenerational Links” (1998). *Center for Policy Research*. 443.
<https://surface.syr.edu/cpr/443>

Easterby-Smith, M., Lyles, M. A., & Peteraf, M. A. (2009). Dynamic capabilities: Current debates and future directions. *British Journal of Management*, 20(SUPP. 1). <https://doi.org/10.1111/J.1467-8551.2008.00609.X>

Edelman, L. F., Manolova, T. S. and Brush, C. G. (2017), “Angel Investing: A Literature Review,” *Foundations and Trends® in Entrepreneurship*: Vol. 13: No. 4-5, pp 265-439.
<http://dx.doi.org/10.1561/03000000051>

Eisenhardt, K. M. (1989). Building Theories from Case Study Research. *The Academy of Management Review* (Vol. 14). Retrieved from http://euroac.ffri.hr/wp-content/uploads/2010/06/Eisenhardt_1989_Building-Theories-from-Case.pdf

Eisenhardt, K. M., & Graebner, M. E. (2007). Theory Building from Cases: Opportunities and Challenges. *Academy of Management*, 50(1), 25–32. Retrieved from <https://aom.org/uploaded-Files/Publications/AMJ/Eisenhart.Graebner.2007.pdf>

Eisenhardt, K. M., & Martin, J. A. (2000). Dynamic Capabilities: What are They? *Strategic Management Journal Strat. Mgmt. J*, 21, 1105–1121.

- Elcock, G. (2021). Unlocking the value of data: Exploring the role of data intermediaries. *Center for Data Ethics and Innovation*, 1–29. Retrieved from <https://www.gov.uk/government/publications/unlocking-the-value-of-data-exploring-the-role-of-data-intermediaries/unlocking-the-value-of-data-exploring-the-role-of-data-intermediaries>
- EU Directive. (2003). Directive 2003/98/EC of the European Parliament and the Council. *Official Journal of the European Union*. Retrieved from <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32003L0098&from=en>
- European Commission. (n.d.). Open data portals. Retrieved November 19, 2022, from <https://digital-strategy.ec.europa.eu/en/policies/open-data-portals>
- European Commission. (2003). Directive 2003/98/EC of the European Parliament and of the Council of November 17, 2003, on the re-use of public sector information. *Official Journal of the European Communities*, L 269 (November 2003), 1–15.
- Evans, A. M., & Campos, A. (2014). Open Government Initiatives: Challenges of Citizen Participation. *Journal of Policy Analysis and Management*, 32(1), 172–203. <https://doi.org/10.1002/pam>
- Fernández, W. D. (2004). The grounded theory method and case study data in IS research: issues and design. *Information Systems Foundations: Constructing and Criticizing Workshop at The Australian National University*2, (July 16-17), 43–59. Retrieved from <http://www.ssnpstudents.com/wp/wp-content/uploads/2015/02/The-grounded-theory-method.pdf>
- Fernández, W., & Lehmann, H. (2005). Achieving Rigour and Relevance in Information Systems Studies: Using grounded theory to investigate organizational cases. *The Grounded Theory Review*, 5(1), 79–107. Retrieved from <http://groundedtheoryreview.com/wp-content/uploads/2012/05/GT-Review-vol5-no12.pdf>

- Fountain, J. E. (2001). Introduction Chapter - Building the Virtual State. Brookings Institution Press.
<https://doi.org/10.7864/j.ctvcb59n3.4>
- Gartner. (2021). Definition of Data Preparation. Retrieved April 25, 2022, from <https://www.gartner.com/en/information-technology/glossary/data-preparation>
- Germano, E. C., De Souza, C. A., & Sun, V. (2016). Business models adopted by intermediaries in the use of open government data. *REBRAE*, 9(1), 79. <https://doi.org/10.7213/REBRAE.09.001.AO05>
- Giaquinto, L. H., & Bruscatto Bortoluzzo, A. (2020). Angel investors, seed-stage investors and founders influence on FinTech funding: an emerging market context. *Macroeconomics and Finance in Emerging Market Economies*, 13. <https://doi.org/10.1080/17520843.2020.1737169>
- Giardino, C., Shahid Bajwa, S., Wang, X., & Abrahamsson, P. (2015). Key Challenges in Early-Stage Software Startups. *Pringer International*. <https://doi.org/10.1007/978-3-319-18612-2>
- Glaser, B. G., & Strauss, A. L. (1967). The Discovery of Grounded Theory - Strategy for Qualitative research. Retrieved from http://www.sxf.uevora.pt/wp-content/uploads/2013/03/Glaser_1967.pdf
- Glaser, B. G., & Strauss, A. L. (1999). The Discovery of Grounded Theory: Strategies for Qualitative Research. *Routledge*. <https://doi.org/10.4324/9780203793206>
- Gov.UK. (2021, May 6). What is data quality? Retrieved November 5, 2022, from <https://www.gov.uk/government/news/what-is-data-quality>
- Government of Canada. (2022). Open Government Analytics. Retrieved August 29, 2022, from <https://open.canada.ca/en/content/open-government-analytics>
- Government of Canada. (n.d.). About Open Government | Open Government, Government of Canada. Retrieved September 15, 2022, from <https://open.canada.ca/en/about-open-government>

- Government of Canada. (n.d.). Open Data 101. Retrieved September 13, 2022, from <https://open.canada.ca/en/open-data-principles>
- Grönroos, C. (2011). Value co-creation in service logic: A critical analysis. *Marketing Theory*, 11(3), 279–301. <https://doi.org/10.1177/1470593111408177>
- Gummerus, J. (2013). Value creation processes and value outcomes in marketing theory: Strangers or siblings? *Marketing Theory*, 13(1), 19–46. <https://doi.org/10.1177/1470593112467267>
- Gurin, J. (2014). Driving Innovation with Open Data. The future of data-driven innovation. Retrieved from [https://www.uschamberfoundation.org/sites/default/files/Gurin Article.pdf](https://www.uschamberfoundation.org/sites/default/files/Gurin%20Article.pdf)
- Gustavsen, N. (2017). Open Government Data and Value Creation - The Case of Nortel A/S. Retrieved from https://rucforsk.ruc.dk/ws/files/60142932/Ninni_Gustavsen_49883.pdf
- Halaweh, M., Fidler, C., & McRobb, S. (2008). Integrating the Grounded Theory Method and Case Study Research Methodology Within IS Research: A Possible “Road Map,” 165. Retrieved from <http://aisel.aisnet.org/icis2008/165>
- Halawi, L. A., Aronson, J. E., McCarthy, R. V, (2005). Resource-Based View of Knowledge Management for Competitive Advantage. *The Electronic Journal of Knowledge Management*, 3(2). Retrieved from <https://commons.erau.edu/publication/319>
- Halawi, L., & McCarthy, R. (2006). Which Theory Applies: An Analysis of Information Systems Research. *Issues in Information Systems (Vol. 7)*. Retrieved from <https://commons.erau.edu/cgi/viewcontent.cgi?article=1360&context=publication>
- Hanna, N. K. (2010). Implications of the ICT Revolution. *Transforming Government and Building the Information Society*, 27–65. https://doi.org/10.1007/978-1-4419-1506-1_2

- Hargadon, A., & Sutton, R. I. (1997). Technology brokering and innovation in a product development firm. *Administrative Science Quarterly*, 42(4), 716–749. <https://doi.org/10.2307/2393655>
- Harrison, T. M., Guerrero, S., Burke, G. B., Cook, M., Cresswell, A., Helbig, N., ... Pardo, T. (2012). Open government and e-government: Democratic challenges from a public value perspective. *The International Journal of Government & Democracy in the Information Age*, 17(2), 83–97. <https://doi.org/10.1145/2037556.2037597>
- Heath, T., & Bizer, C. (2011). Introduction. In *Linked Data Evolving the Web into Global Data Space* (pp. 1–6). https://doi.org/10.1007/978-3-031-79432-2_1
- Hedrick, C. W. (1999). Democracy and The Athenian Epigraphical Habit. *The Journal of the American School of Classical Studies at Athens*, 68(3), 387–439.
- Helfat, C. E., Finkelstein, S., Mitchell, W., Peteraf, M. A., Singh, H., Teece, D. J., & Winter, S. G. (2007). Dynamic capabilities and organizational process. *Dynamic Capabilities: Understanding Strategic Change in Organizations*. Blackwell Publishing Ltd.
- Hossain, M. A., Dwivedi, Y. K., & Rana, N. P. (2016). State-of-the-art in open data research: Insights from existing literature and a research agenda. *Journal of Organizational Computing and Electronic Commerce*, 26(1–2), 14–40. <https://doi.org/10.1080/10919392.2015.1124007>
- Howells, J. (2006). Intermediation and the role of intermediaries in innovation. *Research Policy*, 35(5), 715–728. <https://doi.org/10.1016/j.respol.2006.03.005>
- Huyer, E., & van Knippenberg, L. (2020). The Economic Impact of Open Data. Opportunities for value creation in Europe. Brussels: *European Commission*. Retrieved March 11, 2020. Retrieved from

<https://op.europa.eu/en/publication-detail/-/publication/755a3268-4832-11e8-be1d-01aa75ed71a1%0Ahttps://op.europa.eu/s/oomk>

Huyer, Esther, & Van Knippenberg, L. (2020). The Economic Impact of Open Data Opportunities for value creation in Europe. *Publications Office of the European Union*.
<https://doi.org/10.2830/63132>

IBM. (2021a). What is custom software development? | IBM. Retrieved March 18, 2022, from <https://www.ibm.com/topics/custom-software-development>

IBM. (2021b). What is Data Science? Retrieved August 24, 2021, from <https://www.ibm.com/cloud/learn/data-science-introduction>

IBM. (2021c). What is Machine Learning? | IBM. Retrieved March 18, 2022, from <https://www.ibm.com/cloud/learn/machine-learning>

IBM. (2021d). What is software development? Retrieved April 26, 2022, from <https://www.ibm.com/topics/software-development>

Ichijo, K., von Krogh, G., & Nonaka, I. (Eds.) (1998). Knowledge enablers. *SAGE Publications Ltd*,
<https://doi.org/10.4135/9781446280256>

Immonen, A., Palviainen, M., & Ovaska, E. (2014). Towards open data based business: Survey on usage of open data in digital services. *International Journal of Research in Business and Technology*, 4(1). <https://doi.org/10.17722/ijrbt.v4i1.197>

Indeed. (n.d.). What Is an Application Developer? | Indeed.com. Retrieved March 16, 2022, from <https://www.indeed.com/career-advice/finding-a-job/application-developer>

- ISO/IEC. (2008). ISO/IEC 25012 Data quality model. Retrieved November 7, 2022, from <https://www.iso.org/obp/ui/#iso:std:iso-iec:25012:ed-1:v1:en>
- James, L. (2013, October 3). Defining Open Data. Retrieved September 28, 2022, from <https://blog.okfn.org/2013/10/03/defining-open-data/>
- Janssen, H., & Singh, J. (2022). Data intermediary. *Internet Policy Review*, 11(1). <https://doi.org/10.14763/2022.1.1644>
- Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, Adoption Barriers and Myths of Open Data and Open Government. *Information Systems Management*, 29, 258–268. <https://doi.org/10.1080/10580530.2012.716740>
- Justice Law. (2014). Access to Information Act. Retrieved November 19, 2021, from <https://laws-lois.justice.gc.ca/eng/acts/a-1/>
- Kassen, M. (2018). Open data and its intermediaries: a cross-country perspective on participatory movement among independent developers. *Knowledge Management Research and Practice*, 16(3), 327–342. <https://doi.org/10.1080/14778238.2018.1471377>
- Kitchin, R. (2014). Conceptualizing data. SAGE Publications Ltd, <https://doi.org/10.4135/9781473909472>
- Kitchin, R. (2014). Ethical, political, social and legal concerns. SAGE Publications Ltd, <https://doi.org/10.4135/9781473909472>
- Kitsios, F., Papachristos, N., & Kamariotou, M. (2017). Business models for open data ecosystem: Challenges and motivations for entrepreneurship and innovation. *Proceedings - 2017 IEEE 19th Conference on Business Informatics, CBI 2017*, 1, 398–407. <https://doi.org/10.1109/CBI.2017.51>

- Kollmann, T., Stöckmann, C., Hensellek, S., & Kensbock, J. (2016). European Startup Monitor. *European Commission*. Retrieved from <https://www.araproject.de>
- Konsek-Ciechońska, J. (2019). Startup Companies, Challenges in Poland. *Knowledge International Journal*, 30(6), 1621–1626. <https://doi.org/10.35120/kij30061621k>
- Koski, H. (2015). The impact of open data-a preliminary study. Retrieved from <https://www.vm.fi/julkaisut>.
- Kraaijenbrink, J., Spender, J. C., & Groen, A. J. (2010). The Resource-based view: A review and assessment of its critiques. *Journal of Management*. <https://doi.org/10.1177/0149206309350775>
- Kučera, J., Chlapek, D., & Nečaský, M. (2013). Open government data catalogs: Current approaches and quality perspective. In *Technology-Enabled Innovation for Democracy, Government and Governance Author(s) (Vol. 8061 LNCS, pp. 152–166)*. https://doi.org/10.1007/978-3-642-40160-2_13
- Kundra, V. (2012). Digital Fuel of the 21st Century: Innovation through Open Data and the Network Effect. *Public Policy*, 21. Retrieved November 18, 2021, from <https://shorensteincenter.org/digital-fuel-of-the-21st-century-innovation-through-open-data-and-the-network-effect/>
- Lahti, T., Wincent, J., & Parida, V. (2018). A definition and theoretical review of the circular economy, value creation, and sustainable business models: Where are we now and where should research move in the future? *Sustainability (Switzerland)*, 10(8). <https://doi.org/10.3390/su10082799>
- Lakomaa, E., & Kallberg, J. (2013). Open data as a foundation for innovation: The enabling effect of free public sector information for entrepreneurs. *IEEE Access*, 1, 558–563. <https://doi.org/10.1109/ACCESS.2013.2279164>

Lathrop, D., & Ruma, L. (2010). Open Government. (First, Ed.). *O'Reilly*.

Law Insider. (2021). Data Access Definition | Law Insider. Retrieved April 25, 2022, from <https://www.lawinsider.com/dictionary/data-access>

Lee, H., & Choi, B. (2003). Knowledge Management Enablers, Processes, and Organizational Performance. *Journal of Management Information Systems*, 20(1), 179–228.
<https://doi.org/10.1080/07421222.2003.11045756>

Leitão, J., Pereira, D., & Gonçalves, Â. (2022). Business Incubators, Accelerators, and Performance of Technology-Based Ventures: A Systematic Literature Review. *Journal of Open Innovation*.
<https://doi.org/10.3390/joitmc8010046>

Lisowska, B. (2016). Metadata for the open data portals. Development Initiative. Retrieved from <https://devinit.org/wp-content/uploads/2018/01/Metadata-for-open-data-portals.pdf>

Lobo-Pulo, A. (2017). Open Data – The Promises, the Disillusion, and the Panacea | Part 1,2 and 3 — *Global Digital Foundation*. Retrieved July 22, 2021, from <http://www.globaldigitalfoundation.org/digital-consumer-hub/open-data-the-disillusion>

Loshin, D. (2013a). Business Intelligence. Business Intelligence. *Elsevier Inc*.
<https://doi.org/10.1016/C2010-0-67240-3>

Loshin, D. (2013b). Repurposing Publicly Available Data. In *Business Intelligence* (pp. 287–301). Elsevier. <https://doi.org/10.1016/B978-0-12-385889-4.00018-1>

Maccani, G. (2016). Exploring the factors that influence adoption of open government data for commercial service innovation in cities. *PQDT - Global*. Retrieved from <http://mural.maynoothuniversity.ie/8779/1/PhD>

- Magalhaes, G., & Roseira, C. (2017). Open government data and the private sector: An empirical view on business models and value creation. *Government Information Quarterly*.
<https://doi.org/10.1016/j.giq.2017.08.004>
- Magalhaes, G., Roseira, C., & Manley, L. (2014). Business models for open government data. *ACM*, 365–370. <https://doi.org/http://dx.doi.org/10.1145/2691195.2691273>
- Magalhaes, G., Roseira, C., & Strover, S. (2013). Open government data intermediaries: A terminology framework. *ACM International Conference Proceeding Series*, 330–333.
<https://doi.org/10.1145/2591888.2591947>
- Magretta, J. (2002). Why Business Models Matter. *Harvard Business Review*.
- Manyika, J., Chui, M., Groves, P., Farrell, D., Van Kuiken, S., & Almasi Doshi, E. (2013). Open data: Unlocking innovation and performance with liquid information. Retrieved from
<https://www.mckinsey.com/mgi>.
- Masip-Bruin, X., Ren, G. J., Serral-Gracia, R., & Yannuzzi, M. (2013). Unlocking the value of open data with a process-based information platform. *Proceedings - 2013 IEEE International Conference on Business Informatics, IEEE CBI 2013*, 331–337. <https://doi.org/10.1109/CBI.2013.54>
- Mason, C. (2020). Venture Capital. *International Encyclopedia of Human Geography*, 155–160.
<https://doi.org/10.1016/B978-0-08-102295-5.10153-2>
- Matheus, R., Vaz, J. C., & Ribeiro, M. M. (2014). Open government data and the data usage for improvement of public services in the Rio de Janeiro city. *ACM International Conference Proceeding Series, 2014-January*, 338–341. <https://doi.org/10.1145/2691195.2691240>

- Mccarthy, J. (2004). What is Artificial Intelligence? Retrieved from <http://www-formal.stanford.edu/jmc/>
- Meijer, A. J., Curtin, D., & Hillebrandt, M. (2012). Open government: Connecting vision and voice. *International Review of Administrative Sciences*, 78(1), 10–29.
<https://doi.org/10.1177/0020852311429533>
- Meyer, C. B. (2001). A Case in Case Study Methodology. *Case Study Methodology Field Methods*, 13(4), 329–352. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.921.9912&rep=rep1&type=pdf>
- Miles, M. B., & Huberman, A. M. (1994). Qualitative Data Analysis (Second). California: Sage. Retrieved from <https://vivauniversity.files.wordpress.com/2013/11/milesandhuberman1994.pdf>
- Morrisette, S. G. (2007). A profile of angel investors. *The Journal of Private Equity*, 10(3), 52-66,5. Retrieved from <https://www.jstor.org/stable/43503520>
- Neumaier, S., Umbrich, J., & Polleres, A. (2016). Automated Quality Assessment of Metadata across Open Data Portals. *Journal of Data and Information Quality (JDIQ)*, 8(1).
<https://doi.org/10.1145/2964909>
- Nogueras-Iso, J., Lacasta, J., Ureña-Cámara, M. A., & Ariza-Lopez, F. J. (2021). Quality of Metadata in Open Data Portals. *IEEE Access*. Retrieved from <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9405650>
- Newbert, S. L. (2008). Value Rareness, Competitive Advantage, and Performance: A Conceptual-Level Empirical Investigation of the Resource-Based View of the Firm. *Strategic Management Journal* *Strat. Mgmt. J*, 29, 745–768. <https://doi.org/10.1002/smj.686>

- Nielsen, C., Lund, M., & Bukh, P. N. (2014). Moving towards maturity in business model definitions. In *The Basics of Business Models* (pp. 28-54). Ventus. <http://bookboon.com/en/the-basics-of-business-models-ebook>. Retrieved from <https://www.bmdc.aau.dk>
- Obama, B. (2009). Transparency and Open Government | whitehouse.gov. Retrieved November 30, 2018, from <https://obamawhitehouse.archives.gov/the-press-office/transparency-and-open-government>
- Open Data Barometer. (2016). ODB Global Report Third Edition Open Data Barometer-Third edition. Retrieved from <https://opendatabarometer.org/doc/3rdEdition/ODB-3rdEdition-GlobalReport.pdf>
- OECD. (2002). E-government Definition. Retrieved November 19, 2022, from <https://stats.oecd.org/glossary/detail.asp?ID=4752>
- OECD. (2016). Open Government: The global context and the way forward, OEDC Publishing, Paris. <http://dx.doi.org/10.1787/9789264268104-en>
- OGP. (2011). Open Government Declaration | Open Government Partnership. Retrieved June 11, 2017, from <https://www.opengovpartnership.org/open-government-declaration>
- OGP. (2019). About | Open Government Partnership. Retrieved August 4, 2019, from <https://www.opengovpartnership.org/about/>
- OKF. (2019). The Open Definition - Open Definition - Defining Open in Open Data, Open Content and Open Knowledge. Retrieved August 4, 2019, from <http://opendefinition.org/>
- Okrah, J., Nepp, A., & Agbozo, E. (2018). Exploring the factors of startup success and growth. *The Business and Management Review*, 9.

OMB. (2009). Executive Office of The President Office of Management and Budget. Retrieved from <http://www.usdoj.gov/ag/foia-memo-march2009.pdf>

Open Data 500: Canada. (n.d.). Retrieved July 15, 2021, from <https://canada.opendata500.com/about.html>

Open Data Barometer. (2017). Open Data Barometer 4th Edition — Global Report. The World Wide Web Foundation, 1–34. Retrieved from <https://opendatabarometer.org/doc/4thEdition/ODB-4thEdition-GlobalReport.pdf>

Open Data Barometer. (2018). Leaders Edition: From Promise to Progress. Retrieved from <https://www.opendatabarometer.org>

Open Government Data - OECD. (n.d.). Retrieved August 2, 2021, from <https://www.oecd.org/gov/digital-government/open-government-data.htm>

Open Knowledge Foundation. (2017). The Open Definition. Retrieved December 7, 2017, from <http://opendefinition.org/od/2.1/en/>

Open Knowledge Foundation. (2019). Mission. Retrieved May 2, 2022, from <https://okfn.org/about/>

Open Knowledge Foundation. (2023). What is open? Retrieved February 3, 2023, from <https://okfn.org/opendata/>

Oracle. (2021a). What Is a Data Warehouse. Retrieved April 26, 2022, from <https://www.oracle.com/database/what-is-a-data-warehouse/>

Oracle. (2021b). What Is a Database. Retrieved April 26, 2022, from <https://www.oracle.com/database/what-is-database/>

- Oracle. (2021c). What is Artificial Intelligence (AI)? Retrieved March 18, 2022, from <https://www.oracle.com/ca-en/artificial-intelligence/what-is-ai/>
- Osterwalder, A., & Pigneur, Y. (2010). Business Model Generation. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.471.9178&rep=rep1&type=pdf>
- Osterwalder, A., Pigneur, Y., & Tucci, C. L. (2005). Clarifying Business Models: Origins, Present, and Future of the Concept. *Communications of the Association for Information Systems*, 16. <https://doi.org/10.17705/1cais.01601>
- OSTHUS. (2021). Definition of Data Access. Retrieved April 25, 2022, from <https://www.osthus.com/osthus-glossary/d/data-access>
- Magalhães, R. P. C. (2019). What is a Startup? A Scoping Review on How the Literature Defines Startup. Universidade Católica Portuguesa. Retrieved from https://repositorio.ucp.pt/bitstream/10400.14/28571/1/TFM_RuiMagalhães.pdf
- Peng, M. W., Lee, S. H., & Hong, S. J. (2014a). Entrepreneurs as intermediaries. *Journal of World Business*, 49(1), 21–31. <https://doi.org/10.1016/j.jwb.2013.04.003>
- Pira International Ltd. (2000). Commercial exploitation of Europe's public sector information. Retrieved from <https://op.europa.eu/en/publication-detail/-/publication/26571436-9097-447e-aed4-1b8522aaf153/language-en>
- Poppen, F., & Decker, R. (2018). The intermediary as an institutional entrepreneur: institutional change and stability in triple-helix cooperation. *Triple Helix*, 5(1), 2–22. <https://doi.org/10.1186/s40604-018-0063-7>

- Priem, R. L. (2007). A Consumer Perspective on Value Creation. *The Academy of Management Review*, 32(1), 219–235. Retrieved from <https://www.jstor.org/stable/pdf/20159289.pdf>
- Priem, R. L., & Butler, J. E. (2001). Tautology in the resource-based view and the implications of externally determined resource value: Further comments - ProQuest. *The Academy of Management Review*, 26(1), 57–66. Retrieved from <https://www.proquest.com/docview/210978661?OpenUrlRefId=info:xri/sid:primo&accountid=14701>
- Purwanto, A., Zuiderwijk, A., & Janssen, M. (2020). Citizens' trust in open government data. In *ACM International Conference Proceeding Series* (Vol. 20, pp. 310–318). <https://doi.org/10.1145/3396956.3396958>
- Ramon-Jeronimo, J. M., Florez-Lopez, R., & Araujo-Pinzon, P. (2019). Resource-Based View and SMEs Performance Exporting through Foreign Intermediaries: The Mediating Effect of Management Controls. *Sustainability 2019*, Vol. 11, Page 3241, 11(12), 3241. <https://doi.org/10.3390/SU11123241>
- Redman, T. C. (2012, May 30). Break the Bad Data Habit. Retrieved November 7, 2022, from <https://hbr.org/2012/05/break-the-bad-data-habit>
- Reggi, L., & Dawes, S. (2016). Open Government Data Ecosystems: Linking Transparency for Innovation with Transparency for Participation and Accountability. In IFIP International Federation for Information Processing. https://doi.org/10.1007/978-3-319-44421-5_6
- Richter, H., & Slowinski, P. R. (2019). The Data Sharing Economy: On the Emergence of New Intermediaries. *IIC International Review of Intellectual Property and Competition Law*, 50(1), 4–29. <https://doi.org/10.1007/S40319-018-00777-7>

- Riley, J. (2017). Understanding Metadata; What is Metadata, and what is it for? National Information Standards Organization. Retrieved from <https://www.niso.org>
- Rowe, A., Dong, L., Landon, J., & Rezkalla, E. (2019). Scaling Startups: Challenges in Canada Innovation Ecosystem. ISPIM Connects Ottawa, Canada, on 7-10 April 2019. <https://www.ispim.org>
- Rowley, J. (2002). Using Case Studies in Research. *Management Research News*, 25(1), 16–27. Retrieved from <https://pdfs.semanticscholar.org/4e18/426cc8767b4141c924236612aafaef75fa75.pdf>
- Ruijter, E., Grimmelikhuijsen, S., Van Den Berg, J., & Meijer, A. (2020). Open data work: understanding open data usage from a practice lens. *International Review of Administrative Sciences*, 86(1), 3–19. <https://doi.org/10.1177/0020852317753068>
- Sadiq, S., & Indulska, M. (2017). Open data: Quality over quantity. *International Journal of Information Management*, 37(3), 150–154. <https://doi.org/10.1016/j.ijinfomgt.2017.01.003>
- Safarov, I., Meijer, A., & Grimmelikhuijsen, S. (2017). Utilization of open government data: A systematic literature review of types, conditions, effects and users. *Information Polity*, 22(1), 1–24. <https://doi.org/10.3233/IP-160012>
- Salamzadeh, A., & Kawamorita Kesim, H. (2015). Startup Companies: Life Cycle and Challenges. *SSRN Electronic Journal*. <https://doi.org/10.2139/SSRN.2628861>
- Sandoval-Almazan, R., Gil-Garcia, J., Janssen, M., Hans, J., Scholl, J., Wimmer, M. A., & Bannis, F. (2014). Towards an Evaluation Model for Open Government: A Preliminary Proposal. In 13th International Conference on Electronic Government (EGOV), Sep 2014, Dublin, Ireland (pp. 47–58). Springer. https://doi.org/10.1007/978-3-662-44426-9_4

- Van Schalkwyk, F., Cañares, M., Chattapadhyay, S., & Andrason, A. (2015). Open Data Intermediaries in Developing Countries. <https://doi.org/10.6084/m9.figshare.1449222>
- Schauppenlehner, T., & Muhar, A. (2018). Theoretical availability versus practical accessibility: The critical role of metadata management in open data portals. *Sustainability (Switzerland)*, 10(2). <https://doi.org/10.3390/SU10020545>
- Reggi, L., & Dawes, S. (2016). Open government data ecosystems: Linking transparency for innovation with transparency for participation and accountability. In 15th IFIP WG 8.5 International Conference, EGOV 2016 (Vol. 9820 LNCS, pp. 74–86). https://doi.org/10.1007/978-3-319-44421-5_6
- Silcock, R. (2001). What is E-government? *Parliamentary Affairs*, 54(1), 88–101. <https://doi.org/10.1093/PA/54.1.88>
- Smith, G., Sandberg, J., Viktoria, R., & Gothenburg, S. (2018). Barriers to innovating with open government data: Exploring experiences across service phases and user types. *Information Polity*, 23, 249–265. <https://doi.org/10.3233/IP-170045>
- Snowflake. (2022). Data Suppliers. Retrieved March 17, 2022, from <https://www.snowflake.com/guides/data-suppliers>
- Someh, I. A., Davern, M., & Shanks, G. (2017). Enablers and Mechanisms: Practices for Achieving Synergy with Business Analytics. In 50th Hawaii International Conference on System Sciences. <https://doi.org/10.24251/HICSS.2017.649>
- Staab, S., & Bizer, C. (2009). The Emerging Web of Linked Data; The Emerging Web of Linked Data. *IEEE Intelligent Systems (Vol. 24)*. <https://doi.org/10.1109/MIS.2009.102>

- Stevens, M. J., & Campion, M. A. (1994). The Knowledge, Skill, and Ability Requirements for Team-work: Implications for Human Resource Management. *Journal of Management*, 20(2), 503–530. <https://doi.org/10.1177/014920639402000210>
- Stewart, J., & Hyysalo, S. (2008). Intermediaries, users, and social learning in technological innovation. *International Journal of Innovation Management*, 12(3), 295–325. <https://doi.org/10.1142/S1363919608002035>
- Stitch Software. (2021). What is data transformation: definition, benefits, and uses. Retrieved April 25, 2022, from <https://www.stitchdata.com/resources/data-transformation>
- Stott, A. (2014). Open data for economic growth. *World Bank*. <https://doi.org/10.1136/bmjopen-2018-024546>
- Strauss, A. L., & Corbin, J. (1998). Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory. (J. Strauss, A. Corbin, Ed.). *Sage Publications, Inc.* <https://doi.org/10.21225/d5g01t>
- Tableau Software. (2021). Data Cleaning: Definition, Benefits, And How-To | Tableau. Retrieved April 25, 2022, from <https://www.tableau.com/learn/articles/what-is-data-cleaning>
- Tauberer, J. (2007). 8 Principles of Open Government Data, 2009 (December), 4–8. Retrieved from <http://workspace.unpan.org/sites/internet/Documents/UNPAN042947.pdf>
- Teece, D. J. (2010). Business models, business strategy and innovation. *Long Range Planning*, 43(2–3), 172–194. <https://doi.org/10.1016/j.lrp.2009.07.003>
- Teece, D. J. (2018). Business models and dynamic capabilities | Elsevier Enhanced Reader. *Long Range Planning*, 51, 40–49. <https://doi.org/10.1016/j.lrp.2017.06.007>

- Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic Management, 18*(7), 509–533. [https://doi.org/10.1002/\(SICI\)1097-0266\(199708\)18:7%3C509::AID-SMJ882%3E3.0.CO;2-Z](https://doi.org/10.1002/(SICI)1097-0266(199708)18:7%3C509::AID-SMJ882%3E3.0.CO;2-Z)
- The Economist. (2017). The world's most valuable resource is no longer oil, but data | The Economist. Retrieved November 18, 2021, from <https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data>
- Thirani, V., & Gupta, A. (2017). The value of data. World Economic Forum. Retrieved August 5, 2021, from <https://www.weforum.org/agenda/2017/09/the-value-of-data/>
- TIBCO Software. (2021a). What is Data Merging? Retrieved April 26, 2022, from <https://www.tibco.com/reference-center/what-is-data-merging>
- TIBCO Software. (2021b). What is Structured Data? Retrieved April 25, 2022, from <https://www.tibco.com/reference-center/what-is-structured-data>
- Tinholt, D. (2013). The Open Data Economy: Unlocking Economic Value by Opening Government and Public Data. Retrieved from https://www.capgemini.com/consulting/wp-content/uploads/sites/30/2017/07/opendata_pov_6feb.pdf
- Treasury Board of Canada. (2021). Evaluation of the Open Government Program. Retrieved March 5, 2022, from <https://www.canada.ca/en/treasury-board-secretariat/corporate/reports/evaluation-open-government-program.html>
- Treasury Board Secretariat. (2012). Canada Action Plan 2012-2014. Retrieved from <http://www.open.gc.ca>

- Treuhaft, S. (2006). The Democratization of Data: How the Internet is Shaping the Work of Data Intermediaries. Institute of Urban & Regional Development - University of California at Berkeley. Retrieved from <https://cloudfront.escholarship.org/dist/prd/content/qt32961226/qt32961226.pdf?t=lnqzjk>
- Ubaldi, B. (2013). Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives. *OECD Working Papers on Public Governance, NO.22*, 61. <https://doi.org/10.1787/5k46bj4f03s7-en>
- UN. (n.d.). E-Government. Retrieved November 19, 2022, from <https://publicadministration.un.org/egovkb/en-us/Overview>
- UN Independent Expert Advisory Group. (2014). A World that Counts. Retrieved from <https://www.un-datarevolution.org>
- UNDESA. (2013). Guidelines on Open Government Data for Citizen Engagement. Open Government Data and Service. Retrieved from [http://workspace.unpan.org/sites/Internet/Documents/Guidelines on OGDCE May17 2013.pdf](http://workspace.unpan.org/sites/Internet/Documents/Guidelines%20on%20OGDCE%20May17%202013.pdf)
- University of Tasmania. (2022). What is Data Linkage? Retrieved April 25, 2022, from <https://www.menzies.utas.edu.au/research/research-centres/data-linkage-unit/what-is-data-linkage>
- Urquhart, C. (2017). Grounded Theory for Qualitative Research: A Practical Guide. *Grounded Theory for Qualitative Research: A Practical Guide*. <https://doi.org/10.4135/9781526402196>

- Urquhart, C., Lehmann, H., & Myers, M. D. (2010). Putting the “theory” back into grounded theory: guidelines for grounded theory studies in information systems Related papers. *Information System*, 20, 357–381. <https://doi.org/10.1111/j.1365-2575.2009.00328.x>
- US Department of Commerce Economics and Statistics Administration. (2014). Fostering Innovation, Creating Jobs, Driving Better decisions: The value of Government Data. Retrieved from <https://esa.gov/sites/default/files/revisedfosteringinnovationcreatingjobsdrivingbetterdecisions-thevalueofgovernmentdata.pdf>
- Van Schalkwyk, Francois & Willmers, Michelle & Czerniewicz, Laura. (2014). Open Data in the Governance of South African Higher Education. *SSRN Electronic Journal*.
<https://dx.doi.org/10.2139/ssrn.2557342>.
- Vargo, S. L., Maglio, P. P., & Akaka, M. A. (2008). On value and value co-creation: A service systems and service logic perspective. *European Management Journal*, 26(3), 145–152.
<https://doi.org/10.1016/j.emj.2008.04.003>
- Vetro, A., Canova, L., Torchiano, M., Minotas, C. O., Iemma, R., Morando, F., ... Morando, F. (2016). Open data quality measurement framework: Definition and application to Open Government Data. *Government Information Quarterly*, 33(2), 325–337.
<https://doi.org/10.1016/j.giq.2016.02.001>
- Walker, J., & Simperl, E. (2018). Analytical Report 10: Open Data and Entrepreneurship. Retrieved from <http://www.europeandataportal.eu/>
- Richard Y. Wang & Diane M. Strong (1996) Beyond Accuracy: What Data Quality Means to Data Consumers, *Journal of Management Information Systems*, 12:4, 5-33, DOI: 10.1080/07421222.1996.11518099

- Webfoundations.org. (2018). From promise to progress. Retrieved February 7, 2023, from <https://opendatabarometer.org/leadersedition/report/>
- Weerakkody, V., Irani, Z., Kapoor, K., Sivarajah, U., & Dwivedi, Y. K. (2017). Open data and its usability: an empirical view from the Citizen's perspective. *Information Systems Frontiers*, 19(2), 285–300. <https://doi.org/10.1007/s10796-016-9679-1>
- Wernerfelt, B. (1984). A resource-based view of the firm. *Strategic Management Journal*, 5(2), 171–180. <https://doi.org/10.1002/SMJ.4250050207>
- Winter, S. G. (2003). Understanding Dynamic Capabilities. *Strategic Management Journal*, 24(10), 991–995. Retrieved from http://repository.upenn.edu/mgmt_papers995. <http://dx.doi.org/10.1002/smj.318>
- Wirtz, B. W., & Birkmeyer, S. (2015). Open Government: Origin, Development, and Conceptual Perspectives. *International Journal of Public Administration*, 38(5), 381–396. <https://doi.org/10.1080/01900692.2014.942735>
- Ye, J. H., Kankanhalli, A., & Yang, Z. (2012). Knowledge brokering for open innovation: A case study of innovation intermediaries. In *International Conference on Information Systems, ICIS 2012* (Vol. 5, pp. 4306–4315). Retrieved from http://annualreport.pg.com/annualreport2010/downloads/PG_2010_AnnualReport.pdf
- Yin, R. K. (1994). *Case Study research - design and Methods*. Sage Publications, Inc. [https://doi.org/10.1016/0002-9149\(74\)90005-8](https://doi.org/10.1016/0002-9149(74)90005-8)
- Yin, R. K. (2003). *Case Study Research: Design and Methods*. Sage Publications, Inc. (Second edition). Sage Publications Inc.

- Yin, R. K. (2009). Six Sources of Evidence. In *Case Study Research: Design and Methods* (4th ed., pp. 99–125).
- Yoon, A., Copeland, A., & McNally, P. J. (2018). Empowering communities with data: Role of data intermediaries for communities' data utilization. *Proceedings of the Association for Information Science and Technology*, 55(1), 583–592. <https://doi.org/10.1002/pr2.2018.14505501063>
- Yu, C.-C. (2016). A value-centric business model framework for managing open data applications. *Journal of Organizational Computing and Electronic Commerce*, 26(1–2), 80–115. <https://doi.org/10.1080/10919392.2015.1125175>
- Yu, H., & Robinson, D. G. (2012). The New Ambiguity of “Open Government.” *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2012489>
- Zahra, S. A. (2021). The Resource-Based View, Resourcefulness, and Resource Management in Startup Firms: A Proposed Research Agenda. *Journal of Management*, 47(7). <https://doi.org/10.1177/01492063211018505>
- Zahra, S. A., Sapienza, H. J., & Davidsson, P. (2006). Entrepreneurship and Dynamic Capabilities: A Review, Model and Research Agenda*. *Journal of Management Studies*, 43(4), 917–955. <https://doi.org/10.1111/J.1467-6486.2006.00616.X>
- Zeleti, F. A., & Ojo, A. (2017). The 6-values open data business model framework. In *Public Administration and Information Technology* (Vol. 32, pp. 219–239). https://doi.org/10.1007/978-3-319-63743-3_9
- Zeleti, F.A., Ojo, A., & Curry, E. (2016). Exploring the economic value of open government data. *Government Information Quarterly*, 33(3), 535–551. <https://doi.org/10.1016/j.giq.2016.01.008>

- Zeleti, F. A., Ojo, A., & Curry, E. (2014). Emerging Business Models for the Open Data Industry: Characterization and Analysis. *Proceedings of the 15th Annual International Conference on Digital Government Research*, 215–226. <https://doi.org/10.1145/2612733.2612745>
- Zisk, S. (2018). What Is Data Enrichment? | Redpoint Global. Retrieved March 17, 2022, from <https://www.redpointglobal.com/blog/what-is-data-enrichment/>
- Zott, C., Amit, R. H., & Massa, L. (2011). The Business Model: Recent Developments and Future Research. *Journal of Management*. <https://doi.org/10.1177/0149206311406265>
- ZUAR. (2022). What Is a Data Portal? Retrieved March 2, 2022, from <https://www.zuar.com/blog/what-is-a-data-portal-everything-explained/>
- Zuiderwijk, A., & Cligge, M. (2016). The Acceptance and Use of Open Data Infrastructures – Drawing upon UTAUT and ECT. *15th IFIP Electronic Government and the 8th Electronic Participation Conference (EGOV EPart 2016)*, 23, 91–98. <https://doi.org/10.3233/978-1-61499-670-5-91>
- Zuiderwijk, A., & Janssen, M. (2014). The negative effects of open government data-Investigating the dark side of open data. *Proceedings of the 15th Annual International Conference on Digital Government Research*. <https://doi.org/10.1145/2612733.2612761>
- Zuiderwijk, A., Janssen, M., Choenni, S., Meijer, R., & Alibaks, R. S. (2012). Socio-technical Impediments of Open Data. *Electronic Journal of E-Government*, 10(2), 156–172.

Appendix A. Consent Form

Department of Communication Letterhead

Consent Form

Date:

Title of the Study: Open Government Data and Value Creation: Exploring the roles Canadian data intermediaries play in the value creation process.

Principal Investigator (Researcher)
Salah Merhi
Doctorate Student
Faculty of Engineering,
Electronic Business Program

University of Ottawa

Supervisor
Rocci Luppigini
Associate Professor,

Communication Department
University of Ottawa

Invitation to participate:

You are invited to participate in this doctoral thesis project undertaken by Salah Merhi, a Doctorate student supervised by Professor Rocci Luppigini from the University of Ottawa.

Purpose of the study:

The study aims to examine how value is created by Data Intermediaries using open government data in the development of digital products and services.

Participation:

You will participate in a face-to-face or video conference interview for about 60 minutes. The interview will take place at a convenient date, time, and location to be agreed upon. The researcher will ask questions about the organization's use of open government data and the essential resources and skills needed for the development of products and services that enable value creation. Your participation will be in the form of questions answering and providing supporting documentation if available. The interview will be audio recorded. Please indicate your preference below:

- ☐ I have no objection to an audio recording
- ☐ I object to an audio recording

Risks:

We do not foresee any risks from your participation in the research, nor expect the questions to cause you any discomfort.

Benefits:

Your participation in the research will enrich the study on how open government data enable value creation. In addition, the researcher expects the study to contribute to the literature on value creation from open government data and to serve as a reference to entrepreneurs exploring the use of open government data.

Compensation:

There will be no compensation for your participation in this study.

Confidentiality and anonymity:

The researcher will respect the confidentiality of the information and your privacy and confirm that no other persons will have access to the provided information. In addition, an identification code will be assigned to information received from your organization. The code will be stored offline, locked in a separate secure place. Documents received will be filed and stored in a locked cabinet. Electronic data will be encrypted and stored digitally under the unique identification code to prevent any accidental references. The result of the research will be presented in an aggregate format that will not identify any person or organization. All papers and electronic records will be destroyed five years after the completion of this research project.

Voluntary Participation:

Your participation in the study is entirely voluntary. You may choose to withdraw from the study at any time and refuse to answer any question without consequences of any kind. All data (electronic & physical) gathered until the withdrawal time will be securely destroyed, leaving no trace or reference to you or your organization.

Further questions & information:

You may contact the researcher or supervisor if you have any questions about the study. For any questions regarding the ethical conduct of this study, please get in touch with the Protocol Officer for Ethics in Research at the University of Ottawa,

Tabaret Hall,

550 Cumberland St, Room 154
Ottawa, ON, K1N 6N5
Tel.: (613) 562-5387
Email: ethics@uottawa.ca

Acceptance:

I, (name of participant), agree to participate in the above research study conducted by Salah Merhi of the School of Electrical Engineering and Computer Science (EECS), Engineering Faculty, the University of Ottawa, which is under the supervision of Professor Rocci Luppigini.

Participant signature: _____
Signature

Date: _____
DD/MM/YYYY

Researcher's signature: _____
Signature

Date: _____
DD/MM/YYYY

Copies of consent form: One copy for the participant
One copy for the researcher

Appendix B. Invitation Letter to Participate in a Doctoral Research Project

Date:

Mr. / Ms.

Position

Firm Name

Address 1

Address 2

Subject: Invitation to participate in a doctoral research project

Dear Mr./Ms.

This letter is to request (XYZ Company)'s assistance with the doctoral research project I am conducting as part of my Ph.D. degree in the Department of Electronic Business at the University of Ottawa, Ontario, under the supervision of Dr. Rocci Luppicini. The research project title is "Open Government Data and Value Creation: Exploring the roles Canadian data intermediaries play in the value creation process." Open Government Data is a new initiative and is considered to be an essential stimulator for more transparency, accountability, and people participation in the government's affairs. It is also credited with the power to stimulate economic growth and job creation. However, there is little knowledge of the value created from open government data used by data intermediaries.

This study aims to examine how entrepreneurs (data intermediaries) use open government and private open data to develop digital products and services that contribute to social and economic value creation and learn about the challenges and opportunities in this domain. Knowledge from this study may help the government's planners, policymakers, researchers, and entrepreneurs understand the potential of open data and its impacts.

I desire to connect with you and senior executives at (XYZ Company) who were and are still engaged in the decision to adopt and use open government data and are involved in selecting the target market and in the development of products and services. I will not be contacting any person directly. Instead, I include with this letter a recruitment letter to be distributed to senior executives at the helm of technology, operations, marketing, and finance departments and any other person you deem essential to the adoption and use of open government data. Contact

information for me and my supervisor is contained in this letter. If you and other senior executives are interested in participating, you are invited to contact me or contact my supervisor Professor Rocci Luppigini to discuss participation in this study in further detail.

The participation of any person is entirely voluntary. Each person will make his/her own independent decision as to whether or not they would like to be involved. All participants will be informed of their rights to participate or withdraw before the interview or in the study. Participants will receive an invitation letter including detailed information about this study and the informed consent form.

To support the findings of this study, quotations and excerpts from the information collected will be labelled with pseudonyms to protect the identity of the participants. Names of participants will not appear in the thesis resulting from this study. Similarly, [XYZ Company] name will remain confidential, and a pseudonym will be used to refer to the organization. All paper field notes collected will be locked in my office and a secure cabinet in the communication department at the University of Ottawa. All paper notes will be destroyed after the completion of the study.

Similarly, all electronic data will be encrypted and stored on an electronic storage device with no personal identifiers and will be destroyed after the project is completed. Finally, only I and my supervisor, Rocci Luppigini, in the Communication Department at the University of Ottawa, will have access to these materials. There are no known or anticipated risks to participants in this study.

I want to assure you that this study has been reviewed and received ethics clearance through the University of Ottawa Research Ethics Board. However, the final decision about participation belongs to the (XYZ company), you, and the firm's senior executives.

If you have any questions regarding this study or would like additional information to assist you in deciding on participation, please get in touch with me.

I look forward to receiving (XYZ Company) favourable decision to assist in my Ph.D. research project and your acceptance to be interviewed.

Salah Merhi
Doctorate Student
Digital Transformation and Innovation Program
Faculty of Engineering
University of Ottawa

Request for volunteers to participate in a research project

Research Project Title:

Open Government Data and Value Creation: Exploring the roles Canadian data intermediaries play in the value creation process.

Are you a senior executive from the technology, operation, marketing or financial or other departments involved in the decisions making process on the adoption and use of Open Government Data to develop digital products and services?

If you answer yes to the question noted above, you may volunteer in this study.

This study examines how entrepreneurs (data intermediaries) use open government and private open data to develop digital products and services that contribute to social and economic value creation and learn about the challenges and opportunities in this domain.

You will be asked to participate in an interview session, face-to-face or by video conference, which will take a maximum of 60 minutes. In addition, a follow-up session for about 30-60 minutes may be required to clarify specific points or issues. If you are interested in participating in this study or would like more information, please get in touch with the primary researcher, Salah Merhi or the supervisor, Professor Rocci Luppigini, at the address listed below. This research study has been reviewed and approved by the University of Ottawa Research Ethics Board, file number H-12-19-5124.

Contact persons:

Primary Researcher
Salah Merhi
University of Ottawa
Digital Transformation and Innovation Program

Supervisor
Rocci Luppigini
University of Ottawa
Department of Communication

Appendix C. Certificate of Ethics Approval

Université d'Ottawa
Bureau d'éthique et d'intégrité de la recherche

09/03/2020
University of Ottawa
Office of Research Ethics and Integrity

CERTIFICAT D'APPROBATION ÉTHIQUE | CERTIFICATE OF ETHICS APPROVAL

Numéro du dossier / Ethics File Number

H-12-19-5124

Titre du projet / Project Title

Open Government Data and
Value Creation: Exploring the
roles of data intermediaries in the
value creation process

Type de projet / Project Type

Thèse de doctorat / Doctoral
thesis

Statut du projet / Project Status

Approuvé / Approved

Date d'approbation (jj/mm/aaaa) / Approval Date (dd/mm/yyyy)

09/03/2020

Date d'expiration (jj/mm/aaaa) / Expiry Date (dd/mm/yyyy)

08/03/2021

Équipe de recherche / Research Team

**Chercheur /
Researcher**

Affiliation

Role

Salah MERHI

École de science informatique et de génie électrique / School of Electrical
Engineering and Computer Science

Chercheur Principal /
Principal Investigator

Rocci LUPPICINI

Département de communication / Department of Communication

Superviseur / Supervisor

Conditions spéciales ou commentaires / Special conditions or comments

Appendix D. Interview Protocol

Date of Interview: _____

Time: _____

Interview method: Zoom, Google, Cisco

Address _____

Interviewee: _____

Company Code: _____

The interview commences with a brief welcome by the interviewer and restating the objectives of the interview.

Questions

1. Please provide me with background information on the company, how and when it started.
 - a. What were your motivation and objectives behind the starting of the company?
 - b. What role does open data play in the decision process?
2. What resources were needed to start the company, and how did you acquire them?
 - a. People, Technology, Finance, and Data (what data, where did you get it)
3. How important is open data to the company?
 - a. Are you dependent on open data?
 - b. Would the company exist without open data?
 - c. What value do you attribute to open data?
4. What types of data are you capturing from government data portals?
 - a. What are the challenges faced you in accessing and using the data?
5. What benefits your product/service offers to your customers?
6. What value is generated by your products/services?
 - a. Social and/or economic
 - b. To the company, customers, society, and investors (if any)
7. How is the growth and development of the company financed?
 - a. Self-finance? Investors? Others?
 - b. How important is the availability of external finance?
8. How profitable is the company, and what is your future growth plan?

9. What challenges and problems faced you have in building your company and the products/services?
10. How important is government support to intermediary data startups, and what would you like to see from the government side (Local, provincial, and federal)?

Appendix E. Issues with Open Government Data

	Issues	Reported by
1	Internal structure and how quickly it is updated	SP01
2	Lots of the data available online is aggregated or summarized - not very useful	SP03
3	Lots of the available open data is the lower value data	SP03
4	Publicly available data, especially weather information, can be a little bit less accurate	SP05
5	Access the individual customers' data	SP05
6	The availability of data is still deficient in our industry	SP05
7	Problem with the validity of the data	SP06
8	The municipalities have a hard time creating open data portals,	SP07
9	The main issue we often face is access to data	SP07
10	What we wanted was not available	SP08
11	The available data were of minimum quality	SP08
12	quality varies significantly from dataset to dataset	SP09
13	Standardization is lacking	SP09
14	Limitation of data, parcel data, is still restricted	SP10
15	The biggest barrier is parcel data and postal code	SP10
16	Data update - they only update every quarter or every year	SP10
17	The quality for us is problematic	SP10
18	formats verification of data, and data quality issues	SP12
19	The ability to search for data has been, and it is still not great	SP12
20	The versioning of data	SP12
21	Finding somebody to contact when we have a question or need a correction about it	SP12
22	The regions aggregation for the data was completely different	SP12
23	The search and discovery component could be improved	SP13
24	Data is not formatted the way we want it to be	SP16
25	The problem is with the availability and timeliness of the data	SP16
26	There is usually a lag between when the data is made available and when the actual thing ends	SP16
27	They only make it available at national levels. They do not make a lot of data available on sub-national levels	SP16
28	The problem with the municipal level is the lack of standardization availability of that data	SP17

	Issues	Reported by
29	There is a significant gap in education about what to do with local data and how to get it out of the system	SP17
30	So many variances; province to province, municipality to municipality and then the attitudes towards open data in cities	SP17
31	There is no standard format to report a budget across the country	SP17
32	Data sharing among various departments in the same municipalities	SP17
33	Limited availability of data portal at the municipal level	SP17

Appendix F. List of Canadian Open Data Intermediaries found on the Web.

Company	Founded Year	Location	Category
Adante Consulting	2009	Toronto, Ontario	Data/Technology, Oil & Gas
Airo Health	2012	Waterloo, Ontario	Lifestyle & Consumer, Scientific Research
Ajah	2010	Montreal, Quebec	Data/Technology
Alberta Data Partnerships Ltd.	1996	Edmonton, Alberta	Data/Technology, Geospatial/Mapping
Alert Labs	2015	Kitchener, Ontario	Data/Technology, Housing/Real Estate
Alongside	2013	Moncton, New Brunswick	Data/Technology, Other
Ambience Data	2014	Mississauga, Ontario	Data/Technology, Environment & Weather
Anagraph	2015	Montréal, Quebec	Data/Technology, Geospatial/Mapping
Analytix Insight	2010	Toronto, Ontario	For-profit
Andrs	2017	Waterloo, Ontario	Geospatial/Mapping, Lifestyle & Consumer
Applyboard	2012	Waterloo, Ontario	Education
AreaVibes	2009	Toronto, Ontario	For-profit
ASL 19	2012	Toronto, Ontario	Developer group
Athena Software	2001	Waterloo, Ontario	Healthcare, Other
Bitbakery	2013	Kitchener, Ontario	Data/Technology
Blackarcs	2015	Moncton, New Brunswick	Education, Geospatial/Mapping

Company	Founded Year	Location	Category
B-Line Analytics	2016	Halifax, Nova Scotia	Transportation
Brisk Synergies Tech Corp	2013	Waterloo, Ontario	Data/Technology, Geospatial/Mapping
Busbud	2011	Montreal, Quebec	For-profit
Canada Businet Co.	2002	Vancouver, British Columbia	Food & Agriculture, Transportation
Canada's Open Data Exchange	2015	Waterloo, Ontario	Nonprofit
Cansel Survey Equipment	1966	Etobicoke, Ontario	Data/Technology
Carey Consultants	2009	Tracadie, New Brunswick	Management Consulting
CARTaGENE	2007	Montreal, Quebec	Nonprofit
Chalklette	2016	Waterloo, Ontario	Data/Technology, Lifestyle & Consumer
Chillwall Inc.	2015	Toronto, Ontario	Data/Technology
Cibc	1961	Waterloo, Ontario	Finance & Investment, Other
City Parking	2016	Montreal, Quebec	Transportation
Clean Cut Energy Corp	2012	Guelph, Ontario	Electricity & Energy Transmission
Colliers International	1898	Vancouver, British Columbia	Housing/Real Estate
Conseil Économique Du Nouveau-Brunswick	1979	Moncton, New Brunswick	Other
Contextere	2015	Ottawa, Ontario	Data/Technology
Cubbyspot	2015	Toronto, Ontario	Education, Other
Curiosity Analysis And Consulting Corp.	2015	Kelowna, British Columbia	Transportation

Company	Founded Year	Location	Category
D2I Corporation	1999	Kitchener, Ontario	Data/Technology, Education
Darkhorse Analytics	2008	Edmonton, Alberta	Data/Technology, Geospatial/Mapping
Data for Good	2015	Toronto, Ontario	Developer group
Data Perceptions	1993	Waterloo, Ontario	Data/Technology
Data Tellit Inc	2016	Waterloo, Ontario	Data/Technology, Lifestyle & Consumer
Dateva	2016	Toronto, Ontario	Healthcare
Deloitte Consulting	1870	Toronto, Ontario	Business & Legal Services
Doctr	2015	Montreal, Ontario	Healthcare
Dynamic Solutions	2015	Victoria, British Columbia	For-profit
Easecommute	2016	Toronto, Ontario	Environment & Weather, Lifestyle & Consumer
Eatsleepride Mobile Inc.	2010	Toronto, Ontario	Data/Technology
Ecopia	2013	Waterloo, Ontario	Data/Technology
Emagin	2016	Kitchener, Ontario	Data/Technology, Governance
Empowered	2015	Kitchener, Ontario	Electricity & Energy Transmission
E-Per-Se International Inc.	1999	Waterloo, Ontario	Business & Legal Services, Data/Technology
Esolutionsgroup	1999	Waterloo, Ontario	Data/Technology
Esri Canada Limited	1984	Toronto, Ontario	Data/Technology, Education
Evidence for Democracy	2015	Ottawa, Ontario	Nonprofit

Company	Founded Year	Location	Category
Evollia	2012	Quebec City, Quebec	For-profit
Eyedro Green Solutions Inc.	2009	Kitchener, Ontario	Data/Technology
Ffunction Inc.	2008	Montreal, Quebec	Data/Technology
Fishbuoy - An Environx Inc. Company	2015	Waterloo, Ontario	Data/Technology, Environment & Weather
Fleetcarma	2007	Waterloo, Ontario	Data/Technology, Transportation
Focus21 Inc.	2014	Kitchener, Ontario	Data/Technology
Funnelcake	2014	Kitchener, Ontario	Data/Technology, Other
Gary Little, Realtor (Royal LePage Sussex)	2006	Sechelt, British Columbia	Housing/Real Estate
Gatineau Open (Gatineau Ouverte)	2012	Gatineau, Quebec	Developer group
Gdm Pipelines	1996	Calgary, Alberta	Data/Technology, Oil & Gas
Geoanalytic Inc	1996	Calgary, Alberta	Geospatial/Mapping
Geolytica	2015	Ottawa, Ontario	For-profit
Gnowise Inc.	2016	Kitchener, Ontario	Data/Technology, Finance & Investment
Gnowit Inc	2011	Ottawa, Ontario	Data/Technology, Media
Greenbelt Fund	2011	Toronto, Ontario	Food & Agriculture
Greenland Engineering And Technologies Group	1994	Collingwood, Ontario	Data/Technology, Environment & Weather
Greenowl Mobile	2009	Markham, Ontario	Data/Technology
Greybox Solutions Inc.	2012	Montreal, Quebec	Data/Technology, Healthcare

Company	Founded Year	Location	Category
H2O Geomatics	2014	Waterloo, Ontario	Data/Technology, Geospatial/Mapping
HackForge	2015	Windsor, Ontario	Nonprofit
Hatfield Consulting	1974	Vancouver, British Columbia	Environment & Weather, Geospatial/Mapping
Hexoskin	2006	Montréal, Quebec	Data/Technology, Healthcare
Hidden Harvest	2011	Ottawa, Ontario	For-profit
Hijinks	2016	Toronto, Ontario	Lifestyle & Consumer
Home Tribe	2002	Edmonton, Alberta	For-profit
Hometrics Corporation	2014	Toronto, Ontario	Data/Technology, Housing/Real Estate
Hotspot Parking Inc.	2013	Fredericton, New Brunswick	Data/Technology, Geospatial/Mapping
Ibi Group	1974	Toronto, Ontario	Data/Technology, Geospatial/Mapping
Ibm Canada	1917	Markham, Ontario	Data/Technology
Imagine Canada	2015	Toronto, Ontario	Nonprofit
Infomagnetics Technologies (Imt)	1991	Winnipeg, Manitoba	Data/Technology, Healthcare
International Development Research Centre (IDRC)	1970	Ottawa, Ontario	Nonprofit
Intuit Canada	1995	Mississauga, Ontario	Data/Technology
Iotheatre	2006	Montreal, Quebec	Data/Technology, Media
Isabel Avery & Company Inc.	2013	Kitchener, Ontario	Media

Company	Founded Year	Location	Category
Keboola	2010	Vancouver, British Columbia	Data/Technology, Management Consulting
Kleurvision Inc.	2004	Port Perry, Ontario	Business & Legal Services, Data/Technology
Knomos	2014	Vancouver, British Columbia	Business & Legal Services, Data/Technology
Knowledgehook	2014	Waterloo, Ontario	Education
Kubra	2004	Toronto, Ontario	For-profit
Lake Winnipeg Basin Information Network (LWBIN)	2009	Winnipeg, Manitoba	Nonprofit
Latitude Geographics	1999	Victoria, British Columbia	Geospatial/Mapping
Les Solutions Cyclélabs Inc.	2014	Montréal, Quebec	Data/Technology, Geospatial/Mapping
Local Logic	2015	Montreal, Quebec	Data/Technology, Housing/Real Estate
Manifold Data Mining Inc.	2001	Toronto, Ontario	Data/Technology
Maplesoft	1988	Waterloo, Ontario	Education, Scientific Research
Mapyourproperty	2014	Toronto, ON, Ontario	Business & Legal Services, Data/Technology
MapYourProperty	2014	Toronto, Ontario	For-profit
MaRS Discovery District	2000	Toronto, Ontario	Nonprofit
Media Edge Communications	1985	Toronto, Ontario	Media
Mediaboro Kiwi Social Wifi	2016	Barrie, Ontario	Media, Management Consulting
Mediadoc Inc.	1997	Waterloo, Ontario	Data/Technology

Company	Founded Year	Location	Category
Metasoft Systems	1988	Vancouver, British Columbia	For-profit
Midnight Illusions Ltd.	2000	Guelph, Ontario	Data/Technology, Education
Milieu Technologies Ltd	2015	Ottawa, Ontario	Data/Technology, Geospatial/Mapping
Miovision	2005	Kitchener, Ontario	Data/Technology, Transportation
Mnubo	2012	Montreal, Quebec	Data/Technology, Other
Montreal Ouvert	2011	Montreal, Quebec	Developer group
Mycampsite-review.Com	2007	Burlington, Ontario	Other
Neatware Company	1997	Toronto, Ontario	Data/Technology, Media
Neil Squire Society	1984	Burnaby, British Columbia	Nonprofit
NestReady	2012	Montreal, Quebec	For-profit
Niagara Connects	2014	St. Catharines, Ontario	Other
Niew Labs	2015	Toronto, Ontario	Developer group
Norima Consulting Inc.	2006	Winnipeg, Manitoba	Business & Legal Services
Omx	2012	Toronto, Ontario	Data/Technology, Governance
Open Data Institute Toronto	2015	Toronto, Ontario	Nonprofit
Open Data Ottawa	2015	Ottawa, Ontario	Developer group
Open Edmonton	2015	Edmonton, Alberta	Developer group
Open Quebec (Quebec Ouverte)	2015	Quebec City, Quebec	Developer group
Openconcept Consulting Inc.	1999	Ottawa, Ontario	Data/Technology

Company	Founded Year	Location	Category
OpenDataSK	2011	Regina, Saskatchewan	Developer group
OpenGovGear	2014	Victoria, British Columbia	For-profit
Openlabyrinth Development Consortium	2007	Calgary, Alberta	Data/Technology, Education
OpenNorth	2011	Montreal, Quebec	Nonprofit
Opensports Inc.	2014	Toronto, Ontario	Other
Parallel 42	2013	Windsor, Ontario	For-profit
Parlay Ideas	2016	Toronto, Ontario	Education
Paybysky Inc	2013	London, Ontario	Transportation
Peekapak Inc.	2013	Toronto, Ontario	Education
Pelmorex Media Inc.	1989	Oakville, Ontario	Environment & Weather
Philanthropic Foundations of Canada (PFC)	1999	Montreal, Quebec	Nonprofit
Piinpoint	2013	Waterloo, Ontario	Geospatial/Mapping, Housing/Real Estate
Pitstop	2013	Kitchener, Ontario	Data/Technology
PlaceSpeak	2011	Vancouver, British Columbia	For-profit
Plasticity Labs	2012	Waterloo, Ontario	Data/Technology
Playticipate	2016	Kitchener, Ontario	Lifestyle & Consumer
Plot And Scatter	2015	Vancouver, British Columbia	Data/Technology
Profound Ideation Inc.	2012	Markham, Ontario	Business & Legal Services, Data/Technology

Company	Founded Year	Location	Category
Proliteracy.Ca	2015	Richmond Hill, Ontario	Data/Technology, Education
Psd	2003	London, Ontario	Data/Technology, Governance
Purple Forge	2008	Ottawa, Ontario	Other
Quandl	2011	Toronto, Ontario	For-profit
Quote Kong Ltd	2016	Elora, Ontario	Data/Technology, Management Consulting
Recollect	2009	Vancouver, British Columbia	For-profit
Recollect Systems Inc	2012	Vancouver, British Columbia	Data/Technology, Education
Reelyactive	2012	Montreal, Quebec	Data/Technology
Refinepro	2014	Toronto, Ontario	Data/Technology, Insurance
Refractions Research	1998	Victoria, British Columbia	For-profit
Rel8ed.To	2016	St. Catharines, Ontario	Data/Technology, Management Consulting
Revery Architecture	1981	Vancouver, British Columbia	For-profit
Selkirk Geospatial Research Centre	2003	Castlegar, British Columbia	Academic institution
Sensor Up	2014	Calgary, Alberta	Data/Technology, Geospatial/Mapping
Sensorsuite Inc.	2012	Toronto, Ontario	Electricity & Energy Transmission, Housing/Real Estate
Service Canada	2006	Edmonton, Alberta	Other
Sesame	2013	Kitchener, Ontario	Data/Technology, Education
Shinydocs Corporation	2013	Kitchener, Ontario	Data/Technology

Company	Founded Year	Location	Category
Silicon Halton	2009	Halton Region, Ontario	Data/Technology, Other
Silota	2014	Vancouver, British Columbia	Data/Technology
Simptek Technologies	2014	Fredericton, New Brunswick	Electricity & Energy Transmission
Sitata	2012	Toronto, Ontario	Data/Technology, Healthcare
Skywatch	2014	Waterloo, Ontario	Data/Technology, Environment & Weather
Spotzi Canada Ltd.	2013	Kitchener, Ontario	Geospatial/Mapping
Stellar Optics Research International Corporation (Soric)	1993	Markham, Ontario	Data/Technology, Scientific Research
Surgically Clean Air	2008	Mississauga, Ontario	Environment & Weather
Swell Advantage	2014	Halifax, Nova Scotia	Data/Technology, Transportation
Systematics	2009	Toronto, Ontario	Data/Technology
Tabnex Inc.	2016	Kitchener, Ontario	Data/Technology
Terrene	2015	Waterloo, Ontario	Data/Technology
The Globe And Mail	1844	Toronto, Ontario	Media
The Triple Effect Inc./Digilearn	2016	Fredericton, New Brunswick	Education, Lifestyle & Consumer
The Unite Project Inc.	2015	Kelowna, British Columbia	Data/Technology, Geospatial/Mapping
The World Council on City Data	2014	Toronto, Ontario	Nonprofit
Thinkdata Works	2014	Toronto, Ontario	Data/Technology
Third Sector Publishing	2001	Orillia/Waterloo, Ontario	Data/Technology, Other

Company	Founded Year	Location	Category
Thomson Reuters	1934	Toronto, Ontario	Business & Legal Services, Data/Technology
Tibet Data	2015	Toronto, Ontario	Developer group
Tides Canada	2015	Vancouver, British Columbia	Nonprofit
Toronto Star	0	Toronto, Ontario	For-profit
Townfolio	2015	Saskatoon, Saskatchewan	Data/Technology
Transit App	2013	Montreal, Quebec	For-profit
Trustpoint Innovation Technologies, Ltd.	2013	Waterloo, Ontario	Data/Technology
Uav Geomatics	2012	Calgary, Alberta	Geospatial/Mapping
Ulula	2013	Toronto, Ontario	Data/Technology
Urban Analytics	2009	Vancouver, British Columbia	Data/Technology, Housing/Real Estate
Value Connect Inc.	2015	Guelph, Ontario	Data/Technology, Finance & Investment
Vidya Consulting Inc.	2015	Waterloo, Ontario	Data/Technology, Education
Virtual Squared	2014	Kitchener, Ontario	Data/Technology, Finance & Investment
Voilà: design d'information révélateur	2013	Montréal, Quebec	Management Consulting
Waterline Resources Inc.	2000	Nanaimo, British Columbia	Environment & Weather, Geospatial/Mapping
Waterlix	2015	Cambridge, Ontario	Data/Technology, Geospatial/Mapping
Waterloop	2015	Waterloo, Ontario	Scientific Research, Transportation

Company	Founded Year	Location	Category
Wave Digital	2013	Toronto, Ontario	Other
Wishpad (Casa-nomic Ventures Inc.)	2015	Toronto, Ontario	Housing/Real Estate, Lifestyle & Consumer
Zensurance	2016	Toronto, Ontario	Insurance
Zora Computing Inc.	2014	Halifax, Nova Scotia	Housing/Real Estate

Appendix G. NVivo-based Coding Process

This appendix provides examples of the coding files and categories of the participating Canadian data intermediaries as generated with the NVivo software used in this thesis, especially in section 3.11. Figure 15 first highlights some of the files in NVivo.

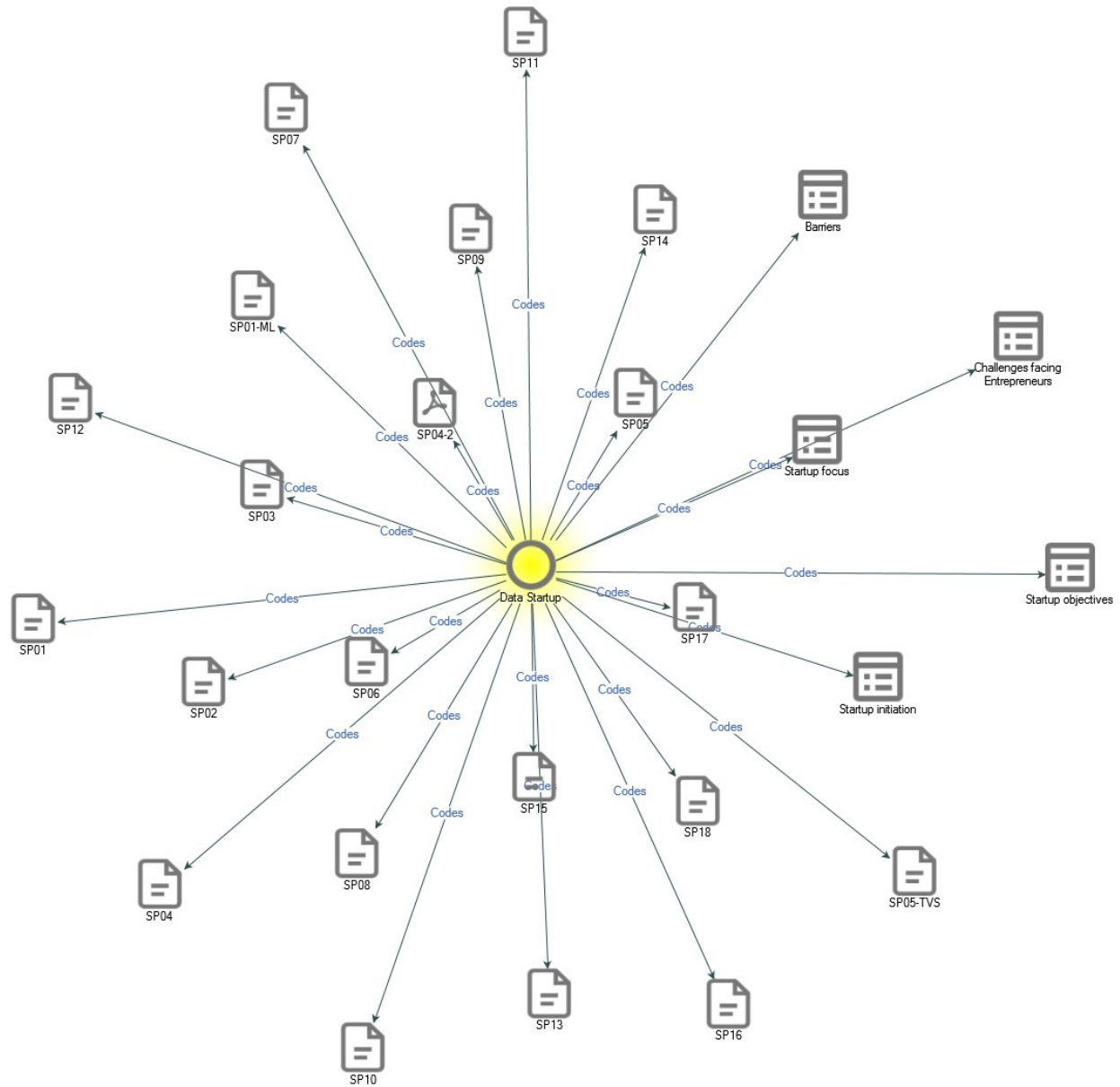


Figure 15: Coding Files (NVivo Screenshot)

Figure 16 shows two NVivo screenshots that display the final selection of nine categories and the number of their corresponding sub-categories (top), with one example of sub-categories (bottom).

The top screenshot displays the 'Codes' list in NVivo. The table below represents the data shown in this list:

Name	Files	References	Created on	Created by	Modified on	Modified by
Data Startup	26	1050	2021-05-19 9:51 AM	SM	2021-05-20 11:46 AM	SM
Government Dimensions	20	178	2021-05-17 3:41 PM	SM	2021-06-07 3:55 PM	SM
Investment, Funds and grants	17	92	2021-05-06 12:47 PM	SM	2021-06-07 3:41 PM	SM
Open Data	22	373	2021-05-04 8:03 AM	SM	2021-06-08 10:44 AM	SM
reflection on business	16	55	2021-05-02 10:08 AM	SM	2021-06-09 1:01 PM	SM
risks and opportunities	6	11	2021-05-03 1:35 PM	SM	2021-05-25 3:43 PM	SM
successful outcomes	7	15	2021-04-28 6:29 PM	SM	2021-05-20 9:50 AM	SM
technology	11	22	2021-05-03 10:34 AM	SM	2021-06-02 5:37 PM	SM
United States	4	12	2021-05-04 5:19 PM	SM	2021-05-19 11:06 PM	SM

The bottom screenshot shows a detailed view of the 'business objectives' category. The table below represents the data shown in this view:

Name	Files	References	Created on	Created by	Modified on	Modified by
business objectives	19	69	2021-05-16 11:41 AM	SM	2021-06-08 9:55 AM	SM
focus of the company	12	29	2021-04-30 4:13 PM	SM	2021-06-03 9:45 AM	SM
Business Opportunities	8	8	2021-04-30 11:39 AM	SM	2021-06-02 3:54 PM	SM
Company Background	23	62	2021-04-29 12:48 PM	SM	2021-06-03 9:10 AM	SM
COVID-19 impacted us.	7	13	2021-05-02 12:38 PM	SM	2021-05-25 4:04 PM	SM
criteria of Success	13	43	2021-04-30 4:04 PM	SM	2021-05-26 12:32 PM	SM
teamwork	9	16	2021-05-07 9:00 AM	SM	2021-05-19 10:17 PM	SM
Data Intermediaries	5	7	2021-04-27 12:48 PM	SM	2021-05-20 11:37 AM	SM
Human Resources	16	29	2021-04-29 1:43 PM	SM	2021-05-27 10:19 AM	SM
Innovation	1	5	2021-05-04 8:56 AM	SM	2021-05-19 10:15 PM	SM
Learning	3	5	2021-05-04 5:15 PM	SM	2021-05-20 11:02 AM	SM
motivation	20	30	2021-05-05 8:59 AM	SM	2021-08-09 9:57 AM	SM
not for profits	7	11	2021-05-07 8:03 AM	SM	2021-06-02 5:37 PM	SM
Organizational Development	25	399	2021-05-04 12:32 PM	SM	2021-05-25 3:40 PM	SM

Figure 16: Final Code Categories (Top) and Sub-categories of One Higher-level Category (Bottom)

Figure 17 shows two NVivo screenshots that display the sub-categories of two high-level categories among the nine final ones.

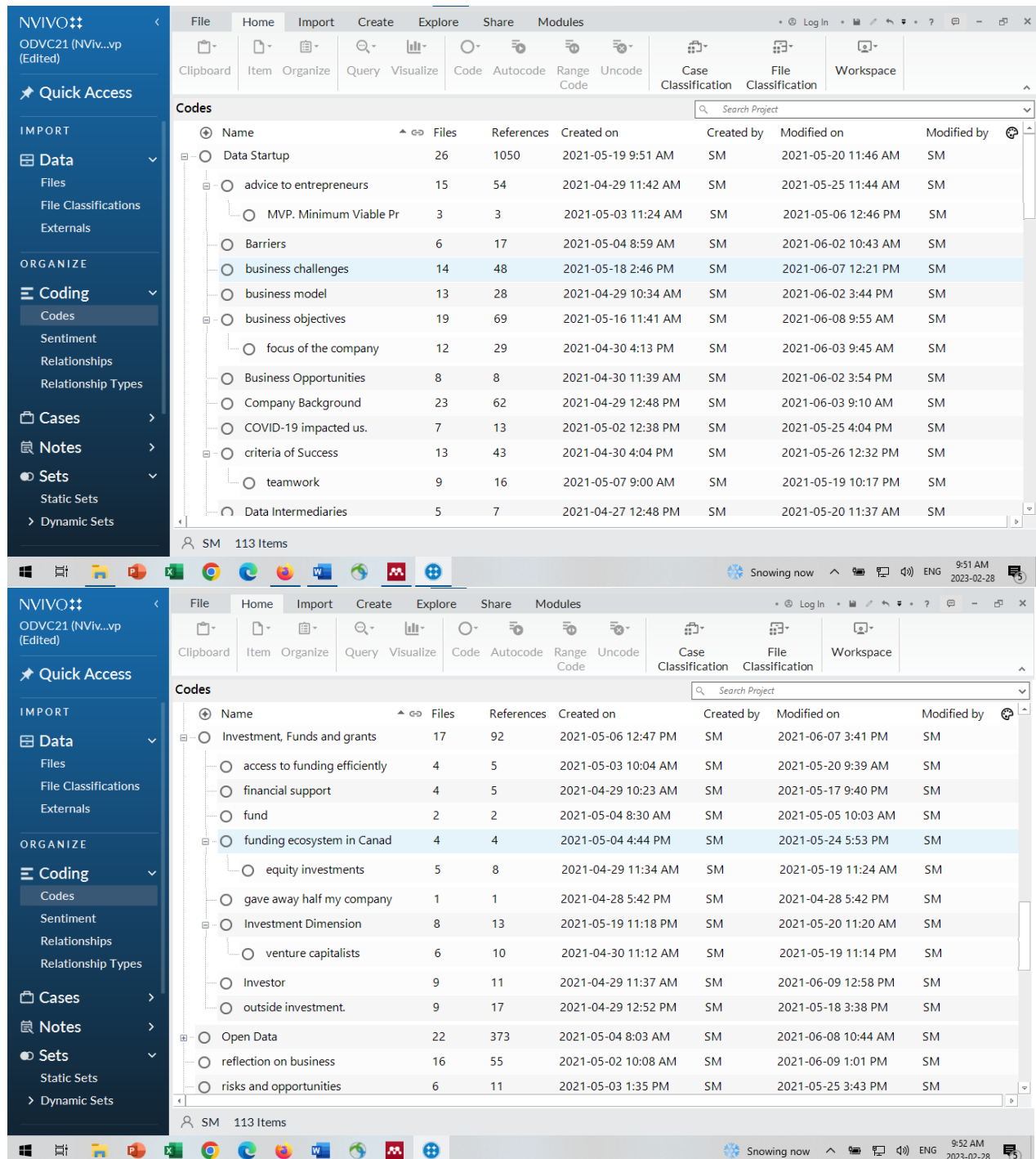


Figure 17: Example Sub-categories for Data Startup (Top) and for Investment, Funds and Grants (Bottom)