Open Data Portals In Northern New England States

by

Bonnie E. Paige

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

Master of Library and Information Studies

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Library and Information Studies)

The University of British Columbia

(Vancouver)

August 2017

© Bonnie E. Paige, 2017
Abstract

As the United States transitions from the Obama administration’s engagement with open government data to the Trump administration’s more closed information strategies, the future support for federal open government data is uncertain. An alternative target for open data initiatives is state-level open government data portals. This study provides preliminary information on state level open data, illustrating challenges faced by small, rural states in supporting an open data portal. The research investigates the current condition of state open data portals: whether their current form and the laws supporting them are sufficient to support their intended use. This study also explores whether the effects of the national political climate can be seen on state portals. This research uses a case study approach, focusing on the northern New England states: Maine, New Hampshire, and Vermont. The case studies use four main methods of investigation: content analysis to determine the goals of the portal, consideration of the policies and context influencing the portal based on the Open Data Policy Framework, inventoring of the data based on the Open Data Barometer, and a review of saved copies of the portals using the Internet Archive.

Based on these methods, we found that these portals fall short of supporting their stated goals. Problems with ambiguous licensing, unclear information organization, unclear project ownership, lack of support for data users, and minimal advertisement of the portal’s existence may have contributed to low citizen engagement with the portals. Portal data is vulnerable as none of the states currently have laws that ensure data will be open and proactively
provided, although Vermont is considering such legislation. National politics may have an influence on state open data, as Maine’s portal ceased updates two days before the federal election.

There is potential for those in the field of library and information science to contribute to state level portals through the provision of support for the knowledge organization and information literacy aspects of the portal that are currently lacking. This study also suggests that evaluative tools more specifically attuned to the state open data context would considerably strengthen the analysis of future research.
Lay Summary

This research focuses on government data made available in open formats on state websites in Maine, New Hampshire and Vermont. It details the data available and the state laws related to open data, and considers whether the portals in their current form are able to support the goals they were created to address. It also looks at the influence of national politics on these portals. The study provides preliminary information on state level open data, illustrating the challenges faced by small, rural states in supporting an open data portal.
Preface

This thesis is the original, unpublished, independent work of the author, Bonnie Paige.
Table of Contents

Abstract ................................................................. ii
Lay Summary ............................................................. iv
Preface ................................................................. v
Table of Contents ......................................................... vi
List of Tables ............................................................ ix
Acknowledgments ......................................................... x

1 Introduction ............................................................. 1
  1.1 Research Questions ................................................. 3

2 Literature Review ....................................................... 5
  2.1 Adoption by Governments: Tensions ............................... 5
  2.2 Adoption by Citizens: Barriers and Opportunities .............. 8
  2.3 Portal Assessment Models and Frameworks ...................... 9
  2.4 Summary .......................................................... 10
  2.5 Expected Outcomes ............................................... 11

3 Research ............................................................ 12
  3.1 Case Selection ................................................... 12
  3.2 Design ........................................................ 13
3.3  Key Documents ........................................... 13
3.4  Documentation ........................................... 15
3.5  Data Collection and Analysis ............................ 15
3.6  Challenges and Limitations .............................. 17

4  Results ..................................................... 20

4.1  Maine ....................................................... 20
  4.1.1  Open Data Policy Framework ....................... 21
  4.1.2  Open Data Barometer ............................... 23
  4.1.3  History ............................................... 25
  4.1.4  Maine Summary ..................................... 26

4.2  New Hampshire ........................................... 26
  4.2.1  Open Data Policy Framework ....................... 27
  4.2.2  Open Data Barometer ............................... 28
  4.2.3  History ............................................... 30
  4.2.4  New Hampshire Summary ............................ 31

4.3  Vermont .................................................... 31
  4.3.1  Open Data Policy Framework ....................... 32
  4.3.2  Open Data Barometer ............................... 33
  4.3.3  History ............................................... 35
  4.3.4  Vermont Summary .................................... 35

4.4  Summary ................................................... 35

5  Discussions ................................................. 37

5.1  Implementation .......................................... 37
  5.1.1  Portal Design and Architecture .................... 38
  5.1.2  Portals in Context .................................. 40
  5.1.3  Data Types ......................................... 43

5.2  Use ................................................................ 46

5.3  Laws and Longevity ....................................... 48

5.4  History and Change ....................................... 50
5.5 Frameworks .................................................. 52

6 Conclusions ..................................................... 54
  6.1 The State of State Open Data .......................... 54
  6.2 Contribution of Research, Strengths and Weaknesses .... 57
  6.3 Application of Findings ................................. 57
  6.4 Future Research ........................................... 58

References ....................................................... 60

Appendix A Dictionary for Open Data Motive Coding .... 65
List of Tables

Table 3.1  Open Data Barometer ........................................ 16
Table 3.2  Open Data Policy Framework .............................. 18

Table 4.1  Examples of datasets within the categories on data.maine.gov .............................. 21
Table 4.2  Dataset frequency for Open Data Barometer categories in data.maine.gov .............................. 24
Table 4.3  Dataset frequency for Open Data Barometer categories in New Hampshire portal .............................. 29
Table 4.4  Dataset frequency for Open Data Barometer categories in data.vermont.gov .............................. 34
Acknowledgments

Thank you to Luanne Freund for your expert guidance, encouragement, and thoughtful revisions.

Thank you to Lisa Nathan for providing the voice of skepticism about open data through your comments.

Thank you to Rick Kopak for helping me realize a thesis was a possible option, and to the SLAIS office staff for fielding my many emails during that process.

Thank you to my parents, Lee and John Paige, for nearly three decades of unwavering support and love.

Thank you to Raj Saha, for all of the logistical, financial, and emotional support that made this thesis possible.

And greatest thanks go to Neil Tenzing Sage, for sitting in the SLAIS lounge while I attended my open government data class before you could even open your eyes.
Chapter 1

Introduction

The global proliferation of information technology has led many governments to look into ways of incorporating technology in their governance. As an outgrowth of the shift to e-government, many governments have chosen to adopt some form of open government, which is defined as “the opening up of government processes, proceedings, documents and data for public scrutiny and involvement” (OECD, 2017). While the term “open government” has existed since the push for Freedom of Information legislation in the 1950s in the United States, it has taken on new importance as an endeavor for government transparency through digital means. The adoption of digital technologies by governments increases opportunities for innovation and efficiency in government (Clarke & Francoli, 2014). Enthusiasm for the idea of open government has spread throughout the international community, with 75 member nations making concrete commitments towards open government through membership in the Open Government Partnership as of 2017 (Action plan, 2017). Open government can take shape through a number of initiatives, including releasing meeting minutes, providing e-services, and creating platforms for deliberation (De Blasio & Selva, 2016). This thesis will focus on open government as achieved through proactive release of open government data.

Open government data is data collected by government organizations in the process of their work that is made publicly available in a way that
facilitates reuse. These datasets can include information on social statistics, transport, meteorology, law, geography, patents and businesses (Ubaldi, 2013). Open data can be used to improve public services, increase oversight of government finances, and inform citizens about the standards expected of government agencies (G. 8 Open Data Charter and Technical Annex - Gov. UK., n.d.). Open government data advocates primarily argue for its social and economic benefits, rather than making the case for openness as a right (Ubaldi, 2013). Access to this data is usually provided through a government open data portal, which allows stakeholders to access the data most relevant to their needs without restricting them to certain views of the data (Alonso et al., 2009).

In the United States, open government data was a major initiative of the Obama administration. President Obama’s first executive action was a memorandum on “Transparency and Open Government” which outlined his vision for a government built on “transparency, public participation, and collaboration” and tasked staffers with developing an action plan to implement these goals (Obama, 2009). This was furthered by an executive order requiring US government data to be open by default (Obama, 2013). In the early 2010s, open government data began to be adopted across many levels of government, from federal agencies to municipalities. These initiatives were supported by administrations across party lines, suggesting that open data could transform into a shared American value.

Trends in the federal government after the 2016 election, however, indicate a shift away from open data and transparency. Obama’s White House portal, now archived at open.obamawhitehouse.archives.gov, provided open data collected at the White House, such as datasets on staff salaries and budget, as well as showcasing President Obama’s commitment to open government by providing information on United States open government initiatives and policies. After the transition to the Trump administration, the open.whitehouse .gov url now redirects to whitehouse.gov/briefing-room/disclosures,
which includes a page for PDF versions of salary data and ethics waivers, as well as a contact form to request financial disclosures. Data on climate change was removed from the Environmental Protection Agency website to align the site with the administration’s priorities and has yet to be replaced (Mooney & Perin, 2017). The Trump administration has given little indication as to what plans, if any, they may have to incorporate open data in their governance. The current Open Government Partnership (OGP) plan commits the United States to actions through 2017, but OGP officials still have not heard whether the project will continue to be supported by the Trump administration (Kroll & Choma, 2017).

Given the uncertainty of open government at the federal level in the U.S., the importance of open data made available by lower levels of government has increased. The G8 Open Data Charter suggests approaching open government data in the widest sense possible, including opening local data by default (G. 8 Open Data Charter and Technical Annex - Gov. UK., n.d.). Open data at lower levels of government may also allow citizens to be the drivers of the project, rather than be merely the recipients of the top-down data release at the national level (Kassen, 2013). While there has been some research into open government data portals for large cities like New York and San Francisco, there has been little investigation of state-level portals. Particularly for smaller and more rural states that are not anchored by a large metropolitan area, we argue that information made available by the state would be the best way to facilitate the goals of transparency and innovation through open government.

1.1 Research Questions

This study will examine three state open data portals to investigate how the push towards open government data initiated by the Obama administration and international organizations such as the Open Government Partnership has been interpreted and implemented at the state level. Research questions
include: Are state-level portals successful in supporting their stated goals? How are they used? Do policies currently enacted at the state level encourage further development of the portals, or is access to currently published data vulnerable? Have there been any changes in state-level portals reflecting the change in federal administration priorities? Case studies will detail the current level of adoption by state governments and determine the degree to which they are following guidelines for implementation as outlined in the Open Data Barometer (Open Data Barometer Global Report 4th ed, 2017). Products or publications created by citizens using the open data will also be included. By looking at the current citizen use of the portals, we can consider whether the current status of the portals is sufficient to encourage reuse. Additionally, policy documents will be discussed. We will investigate whether the current state policy framework would ensure continued access to open data in the long term.
Chapter 2

Literature Review

We will focus on literature discussing the adoption of open government data by governments and by citizens, as well as methods for the assessment of portals and data policies.

2.1 Adoption by Governments: Tensions

Claims have been made that governments benefit from open data, as it allows government processes to function more efficiently and transparently, increasing legitimacy in the eyes of citizens (Ubaldi, 2013). In addition to social benefits like transparency, proponents argue that open data can provide economic benefits to governments by encouraging innovation, as well as operational benefits by optimizing government processes (Janssen, Charalabidis, & Zuiderwijk, 2012). The motivations of governments opening their data, however, are often more complex than simply viewing open data as a public good. The use of open data can also be a strategic move to appear open while pursuing a neoliberal agenda of privatization (Bates, 2014).

Yang, Lo, and Shiang (2015) identified the legal and policy framework as the most important factor inhibiting open data. Difficulty navigating various privacy regulations and ensuring appropriate licensing discourage governments from opening their data. They also found that useful data in non-
machine readable formats, such as articles or pictures, pose a major problem for data publication. While the information may be valuable, it is difficult to include these sources within an open data information environment without significant effort to prepare them into a machine-readable form. Pressure from authority, organizational norms of information sharing and concerns of liability for data misuse also work to discourage the opening of data (Yang & Wu, 2016). In local initiatives, it may be more difficult to identify data that is non-identifying and has copyright that allows for release (Conradie & Choenni, 2014). It may be easier for local projects to begin with data that could be publicly observed, such as the location of public art projects, to work through the basic technical aspects of the project before dealing with the legal issues of releasing more sensitive data.

A study by Yang and Wu (2016) shows that the organizational capability coming from existing staff gives governments confidence to proactively open their data. Particularly, Lee (2013) found that stronger working relationships between IT and program staff improved the ability to collaboratively solve issues, leading to better e-government effectiveness. The proactive opening needs to be facilitated by adequate resources and support in order to be effective (Yang & Wu, 2016). Perceived usefulness motivates governments to open their data, but governments often underestimate data’s usefulness by failing to consider uses such as inter-agency data sharing (Yang & Wu, 2016). Other governments undertaking open data initiatives can provide a positive influence, as well as inspiration as to what datasets could be shared (Yang et al., 2015).

When a government has decided it has the capacity to begin an open data project, the project is shaped the context and attitude in which it is being created. De Blasio and Selva (2016) comparison of European open data policy documents found that the two main models of open data approach the issue from either an innovation- or economic-based framework. The innovation-based policies focus on transparency and public-private collaboration, while
the economic models anchor open data policies within previously existing economic policies. De Blasio and Selva (2016) suggest that more democratic policy approaches, like collaborative projects involving other stakeholders in government processes, are beginning to be implemented in some local governments. This may lead to a new participatory trend in open data policy.

A case study comparison by Dawes, Vidiasova, and Parkhimovich (2016) explores open data as an ecosystem where the government interacts with data producers, innovators and users. In addition to providing high-quality data, governments can provide consultations to steer stakeholders towards real community needs that could be addressed with data. They also found that taking advantage of a broad network of existing working relationships within governments and agencies was a more effective foundation for open government data that relying on select activists to push the effort forward. Smaller governments may encounter additional difficulties due to a lack of resources and few locals with data expertise. One potential strategy is to create regional cooperatives to share costs and expertise with neighboring governments (Lassinantti, Bergvall-Kåreborn, & Ståhlbröst, 2014). These cooperatives currently exist primarily between municipalities, counties and organizations, such as the Western Pennsylvania Regional Data Center, but the same model could be applied to state-level portals.

After the initial difficulties in getting buy-in are overcome and the philosophy is cemented, the project can proceed to actually publishing the government data. According to Janssen and Zuiderwijk (2014), it is more effective to incorporate multiple different business models into an open data project, rather than simply putting raw data into portal. Applications, aggregators, and service platforms can allow the data to be more accessible to users without high levels of comfort with technology. These services can bring government information into users everyday life, such as a trash collection alert or

---

1www.wprdc.org
a restaurant health code violation notification system, or help users monitor
government, as with a legislation tracker. Different business models can also encourage different types of participation and dialogue. Although it may initially be simpler to release information such as financial data that is already in tabular form, it also provides additional value to citizens to publish open data such as meeting transcripts and extract, analyze and visualize the data (Steinbauer, Hiesmair, & Anderst-Kotsis, 2016). This allows citizens to see trends in government deliberation without fully reading each text, which will encourage them to access the primary source data rather than relying solely on infomediaries.

2.2 Adoption by Citizens: Barriers and Opportunities

Once a government has embraced open data, it is often difficult to engage potential users. Zuiderwijk, Janssen, and Dwivedi (2015) carried out a study of open data users based on the Unified Theory of Acceptance and Use of Technology (UTAUT). The team found that a belief that open data would help the user succeed in their job and data use amongst respected peers and colleagues are the most significant predictors of the intent to use open data. Perceived difficulty of using open data and a lack of work that requires the use of open data lessened the intent to use open data. Based on these findings, an open government data portal could increase its use by clearly communicating and advertising the potential value of its datasets, creating a user community, keeping portal navigation clear and consistent, and encouraging schools and businesses to use open data as part of their regular workflow.

Engagement with the public can also be improved by opening up datasets that are most interesting to the target groups of citizens. According to Both (2012), potential users are most interested in data about city planning, administration and the environment. While this might vary by community, identifying these high interest datasets is an important aspect of attracting users to the project. Since many users may not know enough about avail-
able data to make informed requests, civic organizations that work with the public, particularly with underserved populations, could play a role advocating for data they identify as useful to their constituents (Ubaldi, 2013). Balancing the needs of citizens with the goals of the government agencies opening their data may require a participatory design or community informatics approach to creating the portal, to ensure that all stakeholders have a real chance to influence the design process (Maruyama, Douglas, & Robertson, 2013). Although these approaches require a large amount of time from both portal creators and participants, the investment may be worthwhile if it significantly improves the portal’s efficacy and likelihood of being useful to the community.

Data literacy can be a major hurdle for would-be users of an open data portal. One potential way to overcome this is by having an “infomediary” who can process some of the open data into forms that are more understandable to the average citizen and also provide some support for users learning to work with data (Chan, Johnson, & Shookner, 2016).

2.3 Portal Assessment Models and Frameworks

For effective adoption, open data portals need to cater to the multiple forms of democratic activity and different user groups that the portal can support, and help uncover the information need behind the question asked (Ruijer, Grimmelikhuijsen, & Meijer, 2017). A model for understanding the context of these intersecting roles is the Democratic Activity Model of Open Data Use, developed by Ruijer et al. (2017). In this model, the open data portal serves as a tool that connects the separate activity streams of citizens and governments for different types of democratic actions. As described in the model, monitorial democracy largely relies on the proactive release of information by the government, and continued oversight of this information by citizens, often facilitated by the media. Deliberative democracy requires dialog between citizens and governments about policy issues in which citizen
opinion has some possibility of contributing to substantive change. Participatory democracy requires the active contribution of citizens in bettering their communities in collaboration with the government. These different democratic processes require a context-sensitive design for an open data portal.

The evaluation of open data adoption is frequently done using benchmarks. Susha, Zuiderwijk, Janssen, and Gronlund (2015) compared the open data benchmarks proposed by several organizations, and found that benchmarks designed to measure open data in one way, such as implementation, are not necessarily adequate to measure other aspects such as impact. Therefore, a deep understanding of what aspects of an open data project should be measured is key to adequately selecting benchmarks. To measure the implementation of open data, Susha et al. (2015) suggest using the Open Data Barometer developed by the World Wide Web Foundation. For a researcher or portal creator, understanding the types of democratic activity the portal intends to support and the measure by which success is defined is critical for the evaluation of a portal.

2.4 Summary

Governments may adopt open data portals to appear transparent, to encourage innovation or to optimize processes. Concerns about licensing, format and legality, along with a lack of in-house expertise needed for the project, often discourage governments from opening their data. Governments also often underestimate the utility of open data unless they are given examples of what it can achieve. Those governments that do open their data usually take either an economic- or innovation-based approach, but a participatory approach is increasing in popularity. This approach is best achieved by providing support to existing networks rather than relying on activists to drive progress. Providing applications and service platforms in addition to the raw data enables more citizens to participate, and extracting data from text sources such as meeting minutes adds additional value to the portal. Regional
cooperatives can help governments share costs and talent.

Citizens are more likely to use open government data if they believe it will help them with their work, find it easy to use, and know respected peers are using it. Identifying high value datasets to release, either through consultation with citizens and organizations representing them or through a participatory design process, will increase citizen use of the portal. Providing some support for users as they learn to work with data and creating more user-friendly products such as visualizations from the data can help more people get involved.

An effective portal will be designed to support monitoring of, dialog with, and participation in the government, despite the differing data needs of these activities. The method of evaluating an open data portal needs to be appropriate to both the processes it is meant to support and the stage of open data the researcher wishes to measure.

### 2.5 Expected Outcomes

Based on the existing research literature to date, we expect that the current content and organization adopted by the state governments will vary widely, but that all states could benefit from more clearly articulated goals and policies. Furthermore, we expect that reuse of data is low and limited to a small group of users. We expect that the policies governing state open data are minimal, and that further legislation would be necessary to protect the future of state open data. We also expect many of the internal policies on open data publication to not be available online, despite the importance of these documents for transparency (*G. 8 Open Data Charter and Technical Annex - Gov. UK.*, n.d.).
Chapter 3

Research

To consider issues in state open data portals, we have conducted case studies of three sample states—Maine, New Hampshire, and Vermont.

3.1 Case Selection

The lack of a clear partisan stance makes these New England states interesting cases in the context of the change in federal administration and its shift with regard to open data policy. The states are fairly centrist politically. All three states currently have Republican governors, despite having voted for the most recent Democratic presidential candidate. Maine and Vermont are also the only states represented by Independent senators. The history and present condition of these states are similar, making it easier to draw comparisons between them. These states are among the least populous and most rural in the US (Annual estimates of the resident population: April 1, 2010 to July 1, 2016, n.d.). Therefore, they lack the access to human capital that a major city like nearby Boston attracts to an open data project. The cases were also chosen, in part, for their geographic proximity to the author.

Naturally, the very qualities that singled these states out as “neutral” also set them apart from many states that have urban cores and different political makeups. It is also worth noticing that these states are considerably less
diverse than the United States as a whole, with a white population of over 90% vs. 77% in the US generally (Annual estimates of the resident population: April 1, 2010 to July 1, 2016, n.d.). With these cautions in mind, the selected cases can be illustrative of some of the issues in creating and sustaining a state open data portal.

3.2 Design

This study follows a multiple-case study design, as described in (Yin, 1994). In these cases, the unit of study is a state portal. A case study approach was chosen to allow for a deeper consideration of how policy and politics affect open data portals than would have been possible in, for example, a survey of all state portals. The ability to combine more specific data from content analysis of the portals and policies with qualitative observations about these relationships is the case study’s strength. The study was exploratory in nature.

The case studies focused on portals in the Northern New England states of Maine, New Hampshire and Vermont. We detailed the types and forms of data available in each portal and any current use by citizens, as well as discussed the policies governing their release. Information was all gathered through reviewing documents either published or available through the Freedom of Access/Information process. We predicted that these cases would follow the replication logic of a literal replication, which means that many of the same issues are present across the case studies, suggesting underlying similarities (Yin, 1994).

3.3 Key Documents

For each research question restated below, we will rely on several key documents.
1. How has the push towards open government data initiated by the Obama administration and international organizations such as the Open Government Partnership been interpreted and implemented at the state level?

This question is addressed through introductory materials on the open data portal websites. This background information provides evidence of what open government data means to the creators of the portals. State level policies also inform this line of questioning.

2. Are state-level portals successful in supporting their stated goals? How are they used?

This information is available primarily through looking at the portals themselves and considering what they publish, with what frequency, and to what degree it facilitates reuse. Projects using the data, or a lack thereof, are also included in answering this question.

3. Do policies currently enacted at the state level encourage further development of the portals, or is access to currently published data vulnerable?

This question uses information from policy documents, and also includes federal policy documents and newspaper articles to put state policies in context.

4. Have there been any changes in state-level portals reflecting the change in federal administration priorities?

Looking at previous iterations of the portals through the Internet Archive documents any shifts in the portals since the change in federal administration.
3.4 Documentation

General note collection was done through a word processor document file, in which all relevant observations while looking through materials are noted. Two reading lists were kept in a citation manager, one for background literature and one for items actively involved in the research such as policy documents and portal screenshots. A spreadsheet was used for content analysis to track information gathered from the portals and the policies.

3.5 Data Collection and Analysis

Data collection was done through a mixed-methods approach, combining document analysis and content analysis (Krippendorf, 1980). The coding for the content analysis, included as Appendix 1, assumes knowledge of social science research and U.S. politics.

We performed several content analyses as a part of this research. First, themes in the introductory information on state open data portals were categorized to look for patterns in the portal creators’ stated goals. Based on the literature, expected major themes were economic accountability and citizen innovation. Units in this section were syntactical, focusing on words that occur within the dictionary of terms for coding.

Next, the content of the portals themselves was analyzed for information content, format, publication date, and licensing information. We looked through the units of the datasets for patterns in content type and publication schedule and compared these portals to benchmarks available from the Open Data Barometer (Open Data Barometer Global Report 4th ed, 2017). These benchmarks focus on inventorying whether different types of data are open, reusable, and updated regularly, as detailed in Figure 1. We also noted if the data is available from the government but not included in the portal. Coding for all other questions followed a simple strategy of 1-present, 0-absent for unambiguous reporting (Krippendorf, 1980). We used the Open
Table 3.1: Open Data Barometer

<table>
<thead>
<tr>
<th>Does the data exist?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is it available online from government in any form?</td>
</tr>
<tr>
<td>Is the dataset provided in machine-readable and reusable formats?</td>
</tr>
<tr>
<td>Is the machine-readable and reusable data available as a whole?</td>
</tr>
<tr>
<td>Is the dataset available free of charge?</td>
</tr>
<tr>
<td>Is the data openly licensed?</td>
</tr>
<tr>
<td>Is the dataset up to date?</td>
</tr>
<tr>
<td>Is the dataset being kept regularly updated?</td>
</tr>
<tr>
<td>Was it easy to find information about this dataset?</td>
</tr>
<tr>
<td>Are data identifiers provided for key elements in the dataset?</td>
</tr>
</tbody>
</table>

Data Barometer Research Handbook to ensure we evaluated the resources in a way that is consistent with other researchers to increase the ease of comparison between portals. We also, however, adjusted the questions from the Open Data Barometer to better suit the examination of state portals by changing the initial question “Does the data exist?” to “Is the data available on the portal?” While this does decrease the ability to directly compare with other research, it is necessary to keep the scope of the thesis focused on the portals themselves rather than all state government information. To address any misrepresentation coming from this focus on portals, the question “Is the data available online from the government in any form?” was answered with a link to the appropriate data, whether on or off the portal. This strategy also has the benefit of identifying data for possible future inclusion. The use of the Open Data Barometer also helped with closure, as it was easy to find patterns in what types of information are absent from the portals. If any projects made use of open data from the portals, we checked for patterns in the content of datasets that were reused.

The final content analysis for this project was of the policies governing these portals. To guide this investigation, we used the Open Data Policy
Framework developed by Zuiderwijk and Janssen (2014). The framework allows for comparison of the context and content of different policies, as well as their performance indicators. The elements of the framework are listed in Table 3.2.

Document analysis, or close reading of text with a goal of interpreting meaning, was used throughout to support the content analysis, as well as being the main approach for the archived versions of the portal sites (Olsen, 2012). Any differences between previous versions of the portals and their current state were noted and compared.

3.6 Challenges and Limitations

The main challenge we experienced in the process of this research was difficulty communicating with portal staff. We were never able to establish contact with staffers knowledgeable about any of the portals, which significantly limited the scope of information we had hoped to access. Specific policies relating to issues like how data was selected for inclusion and how frequently data was updated were not available online, and without these contacts in government we had to extrapolate these policies based on what was evident in the portal. It may be that these policy documents do not exist, but the impossibility of confirming this was a frustrating result of the lack of communication.

The other major obstacle this project encountered is the changing landscape of open data in the United States. Federal open data is in the process of changing to fit the current administration’s priorities. This may mean that some of the context provided in this thesis becomes obsolete before publication.

The completed research still has limitations. It is difficult to generalize the results of case studies, as they focus on the case’s unique aspects. Although the cases may provide a snapshot of certain portals and their policies, no conclusion can be drawn about state level open data in general from this
Table 3.2: Open Data Policy Framework

<table>
<thead>
<tr>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level of Government/ Mission Type</td>
</tr>
<tr>
<td>Key motivations, policy objectives</td>
</tr>
<tr>
<td>Open data platform launch</td>
</tr>
<tr>
<td>Resource allocation and economic context</td>
</tr>
<tr>
<td>Legislation</td>
</tr>
<tr>
<td>Social and political context, culture in which the opening of data is institutionalized</td>
</tr>
<tr>
<td>Policy strategy and principles for opening data</td>
</tr>
<tr>
<td>Policy measure and instruments</td>
</tr>
<tr>
<td>Processing of data before publication</td>
</tr>
<tr>
<td>Amounts of open data</td>
</tr>
<tr>
<td>Types of open data</td>
</tr>
<tr>
<td>Way of presenting data</td>
</tr>
<tr>
<td>Fee charged for data access</td>
</tr>
<tr>
<td>Target group(s) for the open data</td>
</tr>
<tr>
<td>Technical standards and formats for open data</td>
</tr>
<tr>
<td>Provision of metadata</td>
</tr>
<tr>
<td>Types of data not publicized</td>
</tr>
<tr>
<td>Technical support for the use of publicized data</td>
</tr>
<tr>
<td>Active encouragement of data reuse and promotion of open data</td>
</tr>
<tr>
<td>Data quality</td>
</tr>
<tr>
<td>Data license</td>
</tr>
<tr>
<td>Availability of data without application or registration and without requiring user detail</td>
</tr>
<tr>
<td>Structure of relationship between information suppliers and users</td>
</tr>
<tr>
<td>Usages of publicized data</td>
</tr>
<tr>
<td>Risks of publicizing data (possible negative impacts)</td>
</tr>
<tr>
<td>Benefits of publicizing data (possible positive impacts)</td>
</tr>
</tbody>
</table>
work. The states selected also create limitations, as the information gathered here may not be relevant to open data portals in more diverse and densely populated states.
Chapter 4

Results

In this chapter, the results of the three cases, Maine, New Hampshire, and Vermont, are presented separately. Results are compared and synthesized in the following Discussion chapter.

4.1 Maine

The Maine open data portal, available at data.maine.gov, is a site powered by the commercial subscription-based platform Socrata, which provides a pre-made layout for open data catalogs as well as tools to create charts and maps without additional software.

The site includes 38 datasets and visualizations. The site’s landing page highlights the categories of government, finances, recreation, and statistics, implying that these are the priority datasets of the site creators. Examples of datasets illustrative of the contents of these categories are included in Table 4.1.

The site also has a prominent link to the financial information hosted at opencheckbook.maine.gov. Although Open Checkbook is hosted separately, the importance it is given on data.maine.gov’s landing page suggests that the two sites are intended to work in tandem. On data.maine.gov, the only introductory text is a nine word tagline within the welcome banner- “Access
Table 4.1: Examples of datasets within the categories on data.maine.gov

<table>
<thead>
<tr>
<th>Data Category</th>
<th>Examples of datasets with category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Government</td>
<td>- Veterans’ services offices</td>
</tr>
<tr>
<td></td>
<td>- Median household income data</td>
</tr>
<tr>
<td>Finances</td>
<td>- State expenses 2011</td>
</tr>
<tr>
<td></td>
<td>- State revenue 2004 2009</td>
</tr>
<tr>
<td>Recreation</td>
<td>- Maine lighthouses</td>
</tr>
<tr>
<td></td>
<td>- Wildlife management districts</td>
</tr>
<tr>
<td>Statistics</td>
<td>- Maine population by counties</td>
</tr>
<tr>
<td></td>
<td>- Crime by county</td>
</tr>
</tbody>
</table>

to public information to facilitate transparency and knowledge”. Content analysis of this phrase shows the intended uses of this portal as stated are for transparency and education. For Maine Open Checkbook, we determined that the first two paragraphs of the “About” page are introductory in nature. These paragraphs give equal attention to issues of transparency and fiscal responsibility, most simply expressed in the portal’s own explanation of the idea behind its creation, “Every Maine citizen has a right to know how their hard earned tax dollars are spent”. Given that this portal houses open financial data, it is unsurprising that the introductory content focuses on transparency in government spending. This page’s introductory content also includes a minor theme of participation, as it “encourage(s) Mainers to share their proposals to save money in state government”.

4.1.1 Open Data Policy Framework

In applying the criteria outlined in Zuiderwijk and Janssen (2014) Open Data Policy Framework, it becomes clear that while the portal is relatively user friendly, it lacks the user support, legal framework, and depth of data to be
fully utilized. The Maine portal was created in 2014. The legal context in which the portal exists consists of national laws combined with the state’s Freedom of Access laws. It is clear on the Open Checkbook page that the state’s Republican governor at the time of creation (and also at the time of this study), Paul LePage, was in support of open financial data as a part of transparency initiatives. While the Open Checkbook is funded and run by the Office of the State Controller, the data.maine.gov portal seems to be generally under the auspices of the state government office InforME (Information Resources Maine), which builds state websites and applications, but whether they provide the funding and oversight for the portal or merely the technical know-how is not stated. There is no information provided about the strategy used in opening data, or how data is processed before publication.

The design of the portal does seem to take user needs into account to some degree in its presentation. The Socrata platform provides tools for information visualization within the portal, making it easier for users to understand the data provided with minimal background or additional software. There are tools for filtering, which for example can display only the deer killed after 2011 within the larger hunting dataset provided on the portal. There are also tools for creating calendars, maps, graphs and charts. These tools are quite user-friendly, as options are presented as check boxes, and data that does not make sense on a given axis is greyed out from selection. The data formats vary, including CSV, RSS, KML, Shapefile, and many other available formats. Basic metadata such as creator, last update and source is usually included with the datasets. Based on the information about sources provided, the risks of publishing data on the portal should be minimal, as the data on the portal is generally gathered from previous publications by government entities.

While no information about technical support is included on the portal, there is a general InforMe help desk that might be able to assist with the por-
tal (InforMe 2014) Attempts to reach the help desk through email, however, were not successful. There is no record of any promotion of reuse of data on the Maine portal online. Usage statistics from the portal indicate that three community users have used the data to create filtered views, maps, or charts with Socrata. No applications or outside uses of the data have been mentioned on the portal. Because it is possible for portal creators to see how data is being used within the portal, it creates a limited feedback loop in which creators could provide more instances of popular datasets.

### 4.1.2 Open Data Barometer

Applying the Open Data Barometer’s evaluation criteria, which focuses on the types of data provided as well as their licensing and currency, to the dataset records on the data.maine.gov catalog, reveals that four types of data are available within the portal. Map data, census data, government spending, and crime statistics are all represented within the portal. The most frequent dataset type, as shown in Table 4.2, is government spending data.

All of these datasets are provided in machine-readable, reusable, downloadable data formats at no charge. All of the datasets are licensed as being in the public domain, but this information was somewhat difficult to uncover. No details are given as to whether this licensing allows for modification of the data or requires attribution. The datasets do all include basic metadata and some include a link to the original data source. Some useful metadata, like tags and categories, are provided for some datasets but not consistently. Datasets do not seem to be up to date, besides map data which is unlikely to need major changes with time. Census information came from the most recent full census, 2010, but does not include any population statistical information gathered since then, such as the American Community Survey (American Community Survey, n.d.). All other datasets used data collected no later than 2013. The datasets frequently show the same created and last
**Table 4.2:** Dataset frequency for Open Data Barometer categories in data.maine.gov

<table>
<thead>
<tr>
<th>Data Category</th>
<th># of Datasets in Portal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Map</td>
<td>1</td>
</tr>
<tr>
<td>Census</td>
<td>2</td>
</tr>
<tr>
<td>Government spending</td>
<td>3</td>
</tr>
<tr>
<td>Crime Statistics</td>
<td>1</td>
</tr>
</tbody>
</table>

updated dates, meaning that they have not been updated since being added to the portal. Additionally, many of these datasets share a created or last updated date around April 16, 2015, which shows that there was an initial push to populate the site around this time. The most recent additions of new datasets was in February 2016, over a year ago at the time of research. No updates were made from November 2016 to the date of this research (July 2017).

The portal also includes a variety of datasets that are not represented by the Open Data Barometer’s metrics. Information about hunting and lighthouses falls outside of the government focus of the Open Data Barometer, but is included in the portal as a part of the recreation focus. The locations of Veteran’s Service offices add human services information to the portal, while datasets about railroads and natural gas pipelines add infrastructure data. It is also important to note what information inventoried by the Open Data Barometer was not present in the portal. Some of these data types, like education performance data, are available in open data formats on their department’s website[^1]. Others, like budget data, are only available through non-open formats, like PDFs or a website search engine. Still others, like health performance data, do not have any data easily available online for public review. Even among those types of data present in the portal, there was often a more complete data repository elsewhere online. For example,

while the portal had some basic map data, the Office of GIS website has much more detailed datasets available for download\(^2\).

### 4.1.3 History

Data from the Internet Archive provides an overview of the development of the portal over time. Examining the 19 saved copies of `data.maine.gov` in the Internet Archive’s Wayback Machine\(^3\), it seems that the Maine portal has not undergone major changes since its initial creation. Between the initial upload of 20 datasets in 2015 and the next archived version of the portal in 2016, 20 datasets were added. Some datasets, such as a heat map of moose-involved crashes, were taken down. Some minor changes in design also appear to have been made by Socrata for the data catalog presentation. Besides these adjustments to the datasets and layout, the page has remained unchanged, including the items highlighted on the front page and major design elements such as the header photo.

In addition to looking at the historical version of `data.maine.gov`, we also found reference to a previous open data portal, Maine DataShare\(^4\). This portal has not been updated since 2010, and most links redirect to `data.maine.gov`, suggesting that `data.maine.gov` was intended as replacement for DataShare. However, the site is still in existence, and its contents include a variety of datasets that are not replicated on `data.maine.gov`, such as the locations of boat launches and data on families unable to afford housing. Whether this data was intentionally left off of the new site or just never was migrated is unclear. Given the broken image links and seven years without updating, it seems likely that this site has been abandoned, but no information is included encouraging visitors to check `data.maine.gov` or explaining that it is no longer being maintained.

\(^2\)http://www.maine.gov/megis/catalog/
\(^3\)web.archive.org
\(^4\)http://www.maine.gov/cgi-bin/data/index.pl
4.1.4 Maine Summary

While Maine’s portal shows some intention of user-friendly design, the portal is missing some elements necessary to support its goals of transparency and education. While some data of special interest to Maine residents and visitors is included, a great deal of important datasets are not. Technical support and education is lacking, as are clear details of licensing and legal framework encouraging open data. Reuse of portal data has been limited. Maine previously had a different open data portal, which was abandoned without fully migrating the data. The current portal was not updated for over 8 months.

4.2 New Hampshire

New Hampshire’s open data has not been worked into a formal portal like Maine and Vermont’s. Instead, a directory provides links to 117 open datasets made available by 17 different departments\(^5\).

The introductory information available on this site is in a PDF on the main open source page: “DoIT Open Data Standards Policy”. The second section of the policy, titled “Open Standards/Open Data Formats”, is the most applicable to open data. The content analysis of this section shows a focus on fiscal responsibility with an additional theme of transparency. According to this document, the driving motivation for New Hampshire’s adoption of open data is to use a “cost-effective, build once, use many times’ approach” to data, as well as to “facilitate the interoperability of State systems” and “support transparency in government” (Open source software and data formats policy, 2012).

\(^5\)https://www.nh.gov/doit/open-source/data-sets.htm
4.2.1 Open Data Policy Framework

Analyzing the New Hampshire portal through the Open Data Policy Framework (Zuiderwijk & Janssen, 2014) shows that although New Hampshire has a strong policy commitment to open data formats, this does not necessarily translate into a strong open data portal because of the choice not to employ user-centered design. No date is available for the platform launch, as New Hampshire’s data is not on a purpose-built portal, but the website copyright date is 2015. The state legal context for the open data is robust, as New Hampshire has a law, RSA 21-R:10-13, which requires that all New Hampshire data comply to open standards “unless specific project requirements, or excessive cost, preclude use of an open data format” (RSA 21-R:10-13, 2012). This law also has a requirement that all existing data not available in open format be reviewed every four years to assess the feasibility of converting it to an open format. The bill was sponsored by former State Representative Seth Cohn, a member of the libertarian Free State movement (Cohn, 2012). This law adds an additional layer to national laws and state Right-to-Know laws.

Since the data is hosted on the sites of individual agencies, and oversight of the data collection and publication comes from these agencies, they essentially provide the funding for the project as well, with the Department of Information Technology providing the minimal funding to host the single page listing the data. Because state data is required to be in open formats, it is already in an appropriate format to be linked to on an open data directory without additional preparation. Acceptable types of open data, such as CSV, SHP, and XML, are specifically defined by the IT department (Open source resources and open data formats, 2015).

The method of presenting the data does not take a user-centered approach, but rather reflects a more traditional government-centric approach. Data is presented by the government department that hosts it, rather than by topic or any other category that is more user needs-based. The list format
aids users in identifying and accessing available datasets, but does not add value for users with respect to using the data, as a platform with capabilities for visualization would. The directory itself also does not provide any metadata aside from the title. The metadata standards adhered to by the contributing departments vary, so no expectation of metadata can be assumed in this directory. This same problem occurs with the data license: the lack of any overarching licensing information on the directory makes it unclear what is available for reuse and what might have more restrictive licensing conditions, such as prohibiting commercial reuse. Even when a user follows the link to a department’s landing page of datasets available for download, licensing information is not clearly stated and the information provided rarely goes beyond the title of the dataset. The model for the directory is decentralized, with individual departments deciding what information to provide and how. For users, this is a low-engagement, informational model without the opportunity to give feedback, share, or be aware of previous uses of the data. There is no known support for working with the data, and the only contact information provided on the directory is the email of the site’s webmaster. Individual department sites also show no sign of providing support for data usage or contact information for staff who can address data-related questions. There is no apparent promotion of the open data on this site or its active reuse. There is no evidence that this data has been used, nor are there any applications or products built with it that can be discovered.

4.2.2 Open Data Barometer

Since no data is actually hosted by New Hampshire’s open data directory, for the purposes of this thesis any dataset the page links to counts as “in the portal”. Inventorizing the datasets using the criteria of 15 dataset types from the Open Data Barometer shows that the datasets cover five different topics. Links are provided for datasets on census data, budget, government spending, education performance and environmental statistics. As shown in
Table 4.3: Dataset frequency for Open Data Barometer categories in New Hampshire portal

<table>
<thead>
<tr>
<th>Data Category</th>
<th># of Datasets in Portal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Census</td>
<td>25</td>
</tr>
<tr>
<td>Budget</td>
<td>1</td>
</tr>
<tr>
<td>Government spending</td>
<td>3</td>
</tr>
<tr>
<td>Education Performance</td>
<td>9</td>
</tr>
<tr>
<td>Environmental Statistics</td>
<td>14</td>
</tr>
</tbody>
</table>

Table 4.3, census data is the most frequently occurring type of dataset, with a significant number of environmental statistics and education performance datasets as well.

All of these datasets are provided in machine-readable, reusable, downloadable data formats at no charge. No clear licensing information is included, neither on the portal itself nor on the pages hosting each specific dataset. All datasets appear to be up to date and regularly updated. Since the datasets are maintained by the individual departments, this updating schedule would depend on the information practices of the departments rather than on the creators of the open data site. The directory does not include explanatory information about the datasets.

Additional datasets that do not fall under the Open Data Barometer categories highlighted on New Hampshire’s open data page could support both governmental and personal business. Bank charters, employment data and projections, and inspections data could all be useful to the business community. Polling places and traffic information are useful to citizens. Interlibrary loan statistics and plant species round out the data provided. Data not linked to on the page but easily available from the government includes GIS data, election results, and government contracts. Additional datasets are available through a search feature, such as environmental data. Some data, while nominally available to access, is extremely difficult to get, such as deeds.
information that requires detailed troubleshooting tips and additional software downloads to be accessed. Other datasets, like the company register, require a log-in for access. No New Hampshire data sources could be found for international trade statistics, public transit, or crime.

4.2.3 History

The Internet Archive has limited information on the history of the New Hampshire open data directory. A check of the single saved copy of the site on the Internet Archive shows that the New Hampshire site has largely remained the same over time. The one archived version of the site is from September 2015. The only difference between this saved page and the current version, as of July 2017, is that one department and dataset—the Insurance Department and New Hampshire Population Under 65 By Zip Code—has been removed from the current page. Aside from this, the page remains unchanged in content or layout, although Internet Archive saved copies of the data the directory links to on the agency pages suggest that the data itself has sometimes been updated in that time.

While the open data directory on the Department of Information Technology website is the most current and complete source for New Hampshire open data, it worth noting that the list of state open data portals compiled by the federal government at www.data.gov/open-gov does not link to this site. Instead, it links to nhopengovt.org, a now-defunct open data portal for government financial information run by a think tank. The last archived version of the page that was actually related to open data was in April 2016. The site, run by the Josiah Bartlett Center for Public Policy, housed a database of 4.5 million state financial transactions (NHOpenGovt lets you Google your government, n.d.). Although the database is no longer available, the datasets seemed to be grouped into state spending, payroll, and pension data. There is no indication of whether this data was moved to a new location online or has been lost. The nhopengovt.org web address,
from October 2016 to present, now sells motor fleet insurance.

4.2.4 New Hampshire Summary

New Hampshire has chosen to host a directory of open data on state agency websites, rather than a full open data portal, in support of its goals of fiscal responsibility and transparency. Due to a law that requires all government data to be made available in open formats, there is a great deal of open data available from New Hampshire. The decentralized model means that things like metadata and licensing are inconsistent. The directory format is not user-friendly, as it does not categorize data in a way that facilitates discovery or offer any support for working with the data.

4.3 Vermont

The Vermont open data portal, available at data.vermont.gov, includes 160 datasets and visualizations hosted by Socrata.

The site’s landing page highlights select datasets, as well as providing a list of 12 categories of data. The portal itself does not include any explanatory text, but a page on the Department of Information and Innovation site does provide this background information (Open Data Portal, 2017). The content analysis of the “What is this Service?” and “What is Included?” sections required contingency analysis, as many words included in the coding dictionary were included in negative phrases (eg. no cost). With the contingency analysis controlling for such statements, the portal appears to focus on participation and education. The initial description of the purpose is to “publish and store publicly accessible data,” which implies providing access as an end rather than a means. However, further discussion promotes using visualizations to “share the data in a meaningful way and in context,” encouraging users to actively create new information products and use the tools to communicate information.
4.3.1 Open Data Policy Framework

Using the Zuiderwijk and Janssen (2014) Open Data Policy Framework, an analysis of Vermont’s portal shows that although the portal does utilise user-friendly design, the policy surrounding open data is still developing and the support system for users is unclear. Vermont’s portal was created in 2014. Although there was an attempt to introduce legislation to make Vermont an “open data democracy state”, this legislation has not been adopted (Open Government Data Legislation, 2017). This means that currently, the portal is supported just by national laws and the state Public Records law. In addition to this legal context, there was a public interest around open data, with a summit that occurred in 2013 (Vermont Open Data Summit, 2017). Vermont’s major city, Burlington, also has its own open data portal which further demonstrates interest in open data within the state. Vermont’s portal is run by the state’s Department of Information and Innovation, which states that the portal has no ongoing maintenance costs (Open Data Portal, 2017). The portal follows an opt-in strategy for government departments, which can choose to include any non-private information in the portal. Those uploading the data are responsible for ensuring sensitive data is not shared. Since the data is opt-in from government agencies, there should be very little risk to publishing data on the portal.

The portal is somewhat user needs driven, with Socrata’s built-in visualization tools to encourage reuse and frequent use of tagging and categorization to aid discovery. The portal also makes full use of its landing page, highlighting several datasets that might be of interest to users. All data provided within the portal is in open formats, but the portal also links to data in closed formats such as PDFs. Basic metadata is included with all datasets, although not all include links to the original data source. Training and support are offered by the Department of Information and Innovation, but it is

\[\text{data.burlingtonvt.gov}\]
unclear if this is available for all users or only government employees wishing to add datasets to the portal. The portal forms a limited feedback loop, as portal creators can see what datasets are popular and reused, and community users can contribute through creating visualizations. The licensing for the datasets varies, with most sets using either the Open Database License or Public Domain, but many datasets provide no licensing information at all. There are no clear instances of the portal being promoted or the reuse of the open data being encouraged. Usage statistics show that the portal to date has 14 community users who have contributed 18 filtered views, maps and charts. No applications or other projects created using this data have been recorded.

4.3.2 Open Data Barometer

The application of Open Data Barometer criteria to the datasets included in the data.vermont.gov catalog shows the portal includes seven types of data. These datasets include map data, census data, budget, health sector performance, crime, environmental statistics, and public contracts. As shown in Table 4.4, census data is the most common dataset type on the Vermont portal, followed by crime statistics.

All of these datasets were provided in machine-readable, reusable, downloadable data formats at no charge. Data identifiers are present for all datasets. Licensing information varied by dataset. Some, such as health performance data, had an unspecified license. Those datasets that did have a license listed used Open Data Commons Open Database License. This license allows the user to share, create with, and adapt the data, as long as the data is attributed, also provided under the same license, and provided in a open format (ODC Open Database License (ODbL) summary., n.d.). Most datasets appear to be up to date, but some datasets do not show signs of being regularly updated. Census data includes up to 2014, and has not been updated since its original addition to the portal in 2015. Health data
Table 4.4: Dataset frequency for Open Data Barometer categories in data.vermont.gov

<table>
<thead>
<tr>
<th>Data Category</th>
<th># of Datasets in Portal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Map</td>
<td>1</td>
</tr>
<tr>
<td>Census</td>
<td>9</td>
</tr>
<tr>
<td>Budget</td>
<td>1</td>
</tr>
<tr>
<td>Health Sector Perf.</td>
<td>1</td>
</tr>
<tr>
<td>Crime Statistics</td>
<td>6</td>
</tr>
<tr>
<td>Environmental Statistics</td>
<td>3</td>
</tr>
<tr>
<td>Public Contracts</td>
<td>3</td>
</tr>
</tbody>
</table>

also has not been updated since being added to the portal. Although environmental data has been updated after its addition to the portal, it has not been updated for 10 months. Information about the datasets is usually easy to find, with links provided, but for some data, like census and budget, the information is lacking. New datasets have been added and updates have been made during the time of writing, suggesting that the portal is still in active use and development.

Additional datasets made available by the Vermont portal include fuel prices, library locations, trade licenses, traffic and accident reports, public records requests, infrastructure, childcare providers, baby names, SNAP retailers, energy use, hospital spending, public works requests, house price index, police incidents, and hazardous waste sites. Some of these data topics are recent additions to the portal, leading to the possibility that additional data topics could be expected in the future. The portal did not include several types of data that were available from the government but not in open data form, such as election data, which is available through a search function on a government website, and transit data, which is available in a PDF. Trade data did not seem to be available from the Vermont government. Notably absent from both the portal and the Internet at large is data on government
spending.

4.3.3 History

Examining the 13 iterations of data.vermont.gov on the Internet Archive, there have been some changes in this portal. The portal is still being updated with new datasets. The total number of datasets and visualizations increased from 116 in August 2016 to 143 in November 2016 to 160 in July 2017. Frequently changing data such as traffic stops are kept updated to the most recent year. One major subtraction from archived versions of the site is the removal of a link to the Governor’s Dashboard, a page providing progress reports on various issues from the office of the governor. The link was removed from data.vermont.gov some time between August and November 2016. That site itself was taken down some time between December 2016 and April 2017 after being unavailable in August 2016 due to a redesign.

4.3.4 Vermont Summary

Vermont’s portal is intended to support participation and education. The portal includes tools to support visualizations, and has fairly complete tagging and categorization of datasets. Vermont introduced legislation in 2015 for proactive release of open data, but it has not been adopted as of yet. New and updated datasets continue to be added in the portal.

4.4 Summary

This chapter presented the results of three case studies of the provision of state-level open data. These case studies included three analytical approaches: considering the policy and design aspects of the portals using the Zuiderwijk and Janssen (2014) Open Data Policy Framework, categorizing the data presented following the criteria of the Open Data Barometer, and ex-
ploring the history of changes made to the portal using the Internet Archive. These cases demonstrate a range of responses to the trend of open data and open government in the United States. Maine’s case shows initial investment in open data, followed by stalled growth and engagement, possibly resulting from a lack of clear oversight of the portal. New Hampshire’s case shows a different approach, in which providing open format data is a legally-mandated priority, but moving beyond supplying data into encouraging citizen engagement with the data does not seem to be a goal of the directory. Analysis of the Vermont case shows the most consistent interest in engaging citizens in open data, with updates to the portal continuing throughout the period of study, but the portal still has issues related to oversight and consistency.
Chapter 5

Discussions

This chapter compares the results of the three cases and relates the insights gained to the initial research questions and existing body of open data literature. The initial section will discuss the current implementation of the portals and whether this supports the portal’s stated goals as determined by the content analysis. Next will be a discussion of how the portals are used by citizens. Following this, we will consider the laws governing the open data portals and whether these are sufficient to guarantee the longevity of the portal. We will then discuss the changes to the portal throughout time and whether these show the influence of political changes on the federal level. Finally, we will evaluate the frameworks and criteria used in this research and consider their applicability to the state open data portal context.

5.1 Implementation

How has the push towards open government data initiated by the Obama administration and international organizations such as the Open Government Partnership been interpreted and implemented at the state level? Are state-level portals successful in supporting their stated goals?

All of the northern New England open data portals were created within roughly the same time frame, 2014 or 2015. This timing is likely linked to
Obama’s 2013 order making federal data open by default (Obama, 2013). These portals were created at a time of momentum for the open data movement. Although the states were developing these portals within the same context, their implementation of open data ideas varies.

5.1.1 Portal Design and Architecture

The portals represent three different models for presenting open data information. Vermont’s portal is most in line with the strategy taken at the federal level, with all data presented through a single portal with a consistent user interface. Maine’s portal is similar, but the prominence given to the separately-hosted Open Checkbook site leads to what is essentially a dual portal. The financial data on Open Checkbook is an important supplement to data.maine.gov, but it is a separate site with a completely different user interface. The reasons that Maine chose to follow this dual structure are unclear, but given that, according to the Internet Archive, Open Checkbook was created a year before data.maine.gov, we would speculate that the development process was going on concurrently in two separate government offices, and that merging the products was deemed too challenging. New Hampshire’s list of datasets is the simplest way to present open data information. All of the data is centrally accessible, and the portal creators do not need to perform ongoing work to maintain the datasets. This model, however, does not provide the same consistency of formatting and navigation as having all data uploaded in the same portal. It also lacks the additional features that a platform provides. New Hampshire’s minimal approach signals that this project is largely to fulfill the letter of the law, and the directory shows the least evidence of interest in centralization or collaboration across departments out of the projects studied. Vermont’s seeming enthusiasm for open data translates well into a successful portal. Maine is between the two extremes, with their choice to use Socrata signaling their interest in user participation in working with open data but their organizational choices making
it difficult for users to succeed.

Of the portal design and architecture choices, Vermont’s seems the most effective at supporting user needs, with a design that supports monitorial democracy through frequently updated datasets and a minor degree of participatory democracy by enabling the reuse of data (Ruijer et al., 2017). The Socrata platform used by the Maine and Vermont open data portals does provide opportunities for users to filter data and create visualizations within the portal, but does not provide additional ways of using data for those less comfortable with technology (Janssen & Zuiderwijk, 2014). The method of organizing information Maine employs does not cater to users without a pre-existing knowledge of the extent of the data. The application of categories and tags is inconsistent, which, combined with the choice to have some information on a separate portal, decreases the likelihood that a user will find datasets of interest. New Hampshire’s page does not have the added value of visualizations, and its design does not seem to reflect any user orientation. Information is presented by department, rather than by a method of organization that has more meaning to the user. There are no keywords, subject categories, or search bars included in this portal, which essentially means that in order to find a dataset, a user would need to already know what department would have collected it or read through the entire page. Vermont includes tags and categories for nearly every dataset, making it simple for users to find data through several different pathways. The information architecture within these portals can make information easier to find, as in Vermont’s case, or add additional confusion. Providing services based on categories useful to the user, such as life events, improves the effectiveness of an e-government service, but requires a reframing of government as one collaborative entity rather than siloed departments (Vintar, Kunstelj, & Leben, 2002). The services examined in this study do not seem to have embraced this change of vision. This is clearest in the case of New Hampshire’s department-centered directory but even in the case of Vermont, categoriza-
tion provides some pathway for users but does not make explicitly clear how the data could be applicable to users.

The front pages of the portals also show varying strategies towards open data. Although New Hampshire does not have such a landing page before its data catalog, Vermont and Maine both have datasets highlighted visually at their main URL before presenting the full data catalog. Maine focuses on content areas (Government, Finance, Recreation and Statistics) and content types (Datasets, Charts, Maps and Open Checkbook). Vermont instead features individual datasets, with a menu of categories to the right. One major advantage of Vermont’s approach is that as additional categories of data are added to the portal, they can be added to the category menu without requiring a redesign of the homepage. The static nature of the selected datasets on the homepage, however, is a missed opportunity, as only those datasets have ever been able to benefit from the serendipitous discovery that results from being featured on a landing page. While the content areas of Maine’s landing page are useful, the content types do not seem natural categories for a user and raise questions as to why the portal creators would want to highlight the types of information products rather than the information. This might be due to a disconnect between the agencies producing the data and the employees carrying out the project.

5.1.2 Portals in Context

The government entities responsible for the portal are often not clearly identified. As might be expected, the state IT offices play an important role in the portals. The New Hampshire Department of Information Technology hosts the open datasets page, although since they do not host or oversee any of the individual datasets, this was primarily a one-time investment of effort to compile the links. The Vermont Department of Information and Innovation oversees data.vermont.gov, with other governments departments being able to request accounts to add datasets. While the accompanying
documentation for the portal stresses that there are no ongoing costs associated with the portal, we believe they mean to an agency choosing to include their data, and that the Department of Information and Innovation is paying for the platform and any administrative staff time. Maine’s portals have both the clearest and the most obscure ownership. Open Checkbook states clearly that it is run and funded by the Office of the State Controller’s transparency program and provides contact information for the portal staff. In contrast, data.maine.gov provides no information about who is responsible for the portal. The contact email is for the creative team at InforME, the office in charge of Maine’s websites and e-government services, but emails to this account were never returned. While it makes sense that the portal services were all created by IT teams, keeping the portals as an IT project may lessen the portal’s exposure for both citizens and government workers, as IT is not the most visible component of government. A partnership with the IT department and other government programs may be needed, as Lee (2013) found that e-government services are most successful when these interdepartmental working relationships are strong. A lack of clear ownership and departmental focus contributes to many of the issues on the portal such as unclear procedures.

The processes the departments use for selecting and preparing the data for inclusion are not clearly stated anywhere. The closest thing to such a statement is Vermont’s reiteration that any data added by an outside department needs to be checked thoroughly for any identifying information. New Hampshire is not required to do a great deal of preparation at the stage of inclusion, as all of their state data is in open formats by default. Data on all of the portals seems to largely have been uploaded based on convenience for the portal creator, but more thought may have gone into the selection than is initially apparent. None of the portals seem to meet Kassen (2013) goal of having the release of data driven by user input rather than a top-down method. Making the criteria and process for adding open data on
these portals public would improve transparency and the effectiveness of the portals.

None of the portals publish any sort of updating schedule, so the updating practices must be gathered from what has occurred on the portals to date. New Hampshire is the most hands-off with its update strategy, with all of the responsibility for keeping data up to date borne by the outside departments. Maine does not seem to have a plan for frequent updating. Many of the datasets have never been updated since they were added to the portal, and eight months have passed since the last update to any dataset. Given the length of time since the last update and the minimal development since the portal’s initial creation period, it seems likely that the project is no longer being pursued. Vermont has the most consistent updating of the portals, but it is difficult to determine whether this follows a set schedule, or is up to the discretion of the dataset owners or portal managers. Making the updating schedule clearer would help users ensure they are using the most up to date data.

Based on the content analysis, transparency seems to be the most common stated aim of these portals, followed by education, fiscal responsibility, and participation. The Maine portal’s aim to support “transparency and knowledge” is not clearly evident in the portal, as there is little support to aid a novice user in building knowledge and much of the information that would be necessary to transparency is missing. New Hampshire also falls short on its goal of transparency, although it is more successful in fulfilling its aims of being cost-effective and interoperable. While Vermont does make some visible attempt to “share the data in a meaningful way,” this could be further developed by providing more background information, as well as better tools to support understanding. Those portals that include education as a part of their mission do have more opportunities to do visualizations within their portals, perhaps supporting these educational goals. Participation in the portal seems to be supported only in a narrow interpretation of creating
these visualizations, rather than a more robust form in which dialog would be encouraged within the portal. As the portals seem to be struggling to uphold their basic transparency goals, a collaborative approach to the portal, as described in De Blasio and Selva (2016), seems like an unlikely possibility for these portals.

5.1.3 Data Types

The types of data represented within the portal do show some pattern. All of the portals provide access to census data as their first- or second-most frequent category of dataset. This data was probably added to facilitate the comparison of other datasets to population statistics. Maine and Vermont both have some map data represented, although both states have dedicated GIS portals as well. This information is also likely to have been uploaded primarily in support of other datasets, enabling users to overlay data onto state maps. Maine and Vermont both also provide access to crime data. This could partially be because these states have the two lowest crime rates in the nation, leading this to be a dataset they would want to promote (Violent crime rate by state, 2017). Maine and New Hampshire both publish data on government spending on their portals, not including all of the additional financial information provided on Open Checkbook. New Hampshire’s introductory information, as well as the information on Maine’s Open Checkbook, argues for open data on economic grounds, which may explain the emphasis on these datasets. New Hampshire and Vermont both provide a detailed budget and environmental statistics. Additionally, there were some categories of data that only one of the selected portals published: New Hampshire’s education performance data and Vermont’s public contracts and health sector performance.

Several categories of data did were not provided by any portal. Some of these types of data, like international trade data and public transit tables, aren’t particularly relevant to small, rural states. Other types of data, such
as legislation, might not have been included due to the technical challenges that would occur when presenting all-text data in these portals. Although the confines of the portal could support a dataset with the numbers and titles of bills and links to the full text, the text analysis described by Steinbauer et al. (2016) would likely be beyond the scope of possibility for these portals. Land ownership data and company registers may have been kept off the portal because of privacy concerns, but as the information is available through government-provided searches elsewhere, it may just have not been viewed as a priority for upload. Election results, which are available for all states, and already available in open formats for download for New Hampshire and Maine, seem to have also just not made it onto the portal at this point despite the ease of adding these datasets.

All of the portals also contained datasets that the Open Data Barometer had not included in its evaluative criteria. Maine included datasets about hunting and wildlife management that fit its reputation as a hub for outdoor sporting, as well as information about veteran’s services to cater to the high percentage of Maine residents who are veterans (Chokshi, 2014). New Hampshire’s additional datasets focus more on its economic interests, presenting bank charters, employment projections, and inspections, but also include other datasets that could prove useful to citizens such as traffic data and the location of polling places, plus information for cultural enrichment, such as interlibrary loan statistics and data on native plants. As Vermont’s portal has the most datasets of the states studied, it is unsurprising that it also has the most varied additional datasets, of which we will only highlight a few. A major focus of this portal that could be replicated on other portals is licensing information for tradespersons and child care providers. These datasets are frequently viewed and frequently updated, showing that they are useful to the community. Other popular datasets show the locations of community resources such as libraries and retailers who accept Supplemental Nutrition Assistance Program (SNAP) payments. A dataset detailing pub-
lic records requests could be useful for transparency, but does not seem to being updated with regularity. Some datasets seem as if they were added to enable the portal creators to troubleshoot early on, but were never deleted, such as the Social Security name data. None of the datasets provided by the states that were not included in the Open Data Barometer were repeated across portals, which implies that these datasets are reflective of the states’ unique characters. Overall, the data included seemed to follow the pattern identified by Yang and Wu (2016) in which datasets are added based on their perceived usefulness, but portal owners fail to recognize a great deal of potentially useful data.

The datasets varied with regard to providing metadata and appropriately linking to source materials. On the Maine and Vermont portals, the Socrata interface automatically generates some metadata, such as the date the portal was last updated, but metadata such as tags and keywords are often not provided by the dataset uploader. The dataset owners on these portals also often neglect to link to the original source of the data, making it difficult to verify information. New Hampshire does not provide any metadata within the portal beyond title. The individual datasets, however, sometimes include useful metadata depending on the department. The lack of metadata for many datasets may be because the information is still hosted by the agency providing that information- they may see less need to formally cite their own agency as the source. In all of the portals, inconsistent metadata can be confusing to users and make it unnecessarily difficult to find required information. The lack of consistent metadata is unsurprising given the investment of time and money that would be required for high quality documentation and indexing. Without this investment, however, it is unlikely that the portal will be useable for the average citizen. There is potential for state libraries to contribute their expertise to open data projects by taking over the indexing and metadata provision for the datasets.

The current datasets in their existing form are inadequate on all the por-
tals to support the innovation and optimized government processes that are promoted by open data advocates as reason for governments to create a portal (Janssen et al., 2012). While the portals may express intentions to create transparency, in practice they do not provide the necessary support to make this a reality.

5.2 Use

*How are they used?*

It is difficult to tell to what degree the data from the portals is used. The New Hampshire portal does not provide any usage information, so there is no way to determine how frequently it is accessed. For the Maine and Vermont portals, dataset views and downloads can give some window into how much the information is being accessed. Vermont’s most viewed item, a filtered view of fuel prices, was accessed 10,918 times and downloaded 82 times. Maine’s most viewed item, a map of Maine lighthouses, was accessed 7,555 times and downloaded 59 times. This means that, at least at the passive level of looking at data, Vermont’s portal is getting more use. Given the topics of the most commonly accessed datasets, this may be due to Vermont providing information that is more useful in daily life. Vermont’s portal also has more community-created products in the portal. Vermont has 14 contributing users who created 18 filtered views, charts, or maps. Maine only has three community users who have created six information products. There is no evidence of an interactive user community on either portal, as the platform does not have features that support interaction like message boards, or even an ability to comment on datasets. Adding in some collaborative features might increase engagement.

Additional factors inhibit the portal’s potential for reuse. It is difficult for users to determine what kind of technical support is provided for the portals. It is possible that state information technology help desks, like
InforME and the Department of Information and Innovation, would provide some technical support to users, but it is unclear whether they are only available to help government officials add datasets or if their services extend to public use. Even if public support is a part of their mandate, in practice they are unresponsive to inquiries. The Socrata platform does provide some tutorials to support users in the basics of working with data, but the link to these tutorials is not featured prominently on either the Maine or Vermont portals, and is instead hidden in the filters tab of the dataset viewer. New Hampshire does not seem to provide any support for their data, and the only contact is the webmaster of the site. Without a clear pathway for a portal user seeking support, only the most technologically confident are likely to be able to contribute to the portal. A potential solution to this issue is partnering with state libraries to provide this support and training, as well as create information products understandable to those with lower data literacy (Chan et al., 2016). Librarians with a background in data stewardship and information literacy instruction could assist users of the portal without the need for extensive retraining of existing portal technical staff.

A lack of consistent licensing information is also problematic for data reuse on all portals. According to Ubaldi (2013), licenses allowing all reuse, including commercial reuse, are essential to realizing the benefits of open government. Ambiguous licensing will lessen participation in the portals. New Hampshire’s portal does not provide licensing information on the page, and the licensing information is inconsistently included with the datasets themselves. The Socrata portal has a space for licensing information, but both Maine and Vermont leave the license unspecified for many of the datasets. For those datasets that do specify license, many in both the Maine and Vermont portals are licensed Public Domain, with some additional datasets for Vermont using the Open Database license. Without this information clearly stated, users do not know the terms to which they need to adhere when reusing data.
A common problem to all three portals is a lack of advertising and communication about the portals. It is difficult to serendipitously discover the portal through looking at other government websites, so most traffic would have to come through users specifically searching out open data. Common strategies for promoting open data, such as hackathons or app creation contests, do not seem to have been tried in relation to these portals. It also doesn’t seem like there has been any push to draw attention to the portals through social media. As a result, the user interaction with the portals has been minimal, with no recorded instances of the data being reused to make applications or projects. If the portal creators undertook efforts to promote the portal, and then publicised the resulting applications or projects that used the data, it might bring more users into the portal and create a more vibrant culture of use. Since perceived usefulness for work is the main motivator for open data users, portal owners should focus on making the portal’s data a part of users daily workflow, starting by finding ways to use it for internal operations in government (Zuiderwijk et al., 2015).

5.3 Laws and Longevity

*Do policies currently enacted at the state level encourage further development of the portals, or is access to currently published data vulnerable?*

Given Yang et al. (2015) finding that the largest factor inhibiting open data release is difficulty navigating laws, it is unsurprising that the legal foundations of these less-developed portals do not provide clear legal guidance for an open data project. Of the three states, New Hampshire is the only state with laws specifically relating to open data. The effectiveness of RSA 21-R:10-13 2012 is evident since all datasets included on the portal are in open formats. Even beyond the portal, government provided data upholds the law and provides an open format option, even if they also provide information in a formatted document such as a PDF. New Hampshire, however, does
not have any laws that push for data to be published online in the first place. While New Hampshire’s Right to Know laws do guarantee access to most governmental records and electronic communications, there is no requirement of proactive release (*Right to Know Law*, 2017). This means that while government information may be open, it also may not be available unless specifically requested.

Although Vermont does not currently have any legislation to support open data or proactive data release beyond the state’s Public Records law, “An Act relating to creating an Office of Public Policy,” which requires data be made available for analysis by the public by 2025, is still at committee (*An act relating to the creating an Office of Public Policy*, H. 435, 2015). This bill requires that the Office of Public policy it creates will “establish the State of Vermont as an open data democracy state whereby the Office shall redact any personal identifying information from data the State collects and shall organize and share the State’s data with the public online in a manner that enables the public to perform analyses with that data”. If this bill were to be passed into law, Vermont would have an exceptionally strong legal basis for an open data program.

Maine does not seem to have initiated any dialog around open data laws thus far. In fact, in 2013 three bills proposed in the state legislature sought to exempt additional record types from basic Freedom of Access requests (Mistler, 2013). The minimal legal framework for open data may be a contributing factor to Maine’s portal including the fewest datasets out of the northern New England portals. Existing legislation in New Hampshire and proposed legislation in Vermont could serve as an example of how open data can be promoted through law, if Maine were to ever decide this issue is a state priority.

Based on the current laws, open data is vulnerable to a change in political opinion in all states studied. Although New Hampshire requires data in open formats, it does not require that data be published or displayed in a
common directory. Maine and Vermont currently offer no reason to believe that open data portals will outlast their current champions. The portals of all three states could easily disappear, as has happened with many predecessors.

5.4 History and Change

Have there been any changes in state-level portals reflecting the change in federal administration priorities?

Each of the portals has followed a different trajectory of change through time. New Hampshire has left its portal virtually unchanged since its creation, only removing one dataset. Maine had an initial flurry of activity in 2015, followed by a slow trickle of additions to the portal. Vermont has steadily added new material to its portal, keeping the content of the portal in a consistent state of expansion. Some of these differences may be due to the culture of support around the portals at the time of their creation. Vermont has the strongest open data culture, as the only state studied to host a summit on open data and create an open data portal on a municipal level in addition to the state portal. Relying on a small group of activists rather than getting full buy-in from the existing social networks of government agencies, however, is an unstable foundation for an open data project (Dawes et al., 2016). There is not enough sustained participation from enough people to indicate that this isn’t the case in Vermont. Maine’s governor supported open financial data as an extension of his fiscal conservatism, but there has been little indication of interest in open data by government employees or citizens otherwise. New Hampshire does not have a history of an open data culture beyond its open data law. The law’s libertarian sponsor suggests that perhaps this portal is an extension of New Hampshire’s libertarian political leanings, with the government providing the minimum support possible and expecting market forces to do the rest. These cultural factors are important in understanding the motivations behind the portals, given the criticisms
of Bates (2014) that suggest many governments participate in open data projects in order to pursue a neoliberal agenda of privatization and appear transparent to gain political favor without commitment to full transparency. It seems likely that Maine and New Hampshire fall somewhere within this pattern, rather than pursuing an open data agenda for idealistic reasons.

History also illuminates the fates of the portals’ failed open data predecessors. Maine’s DataShare seems to have been discontinued as it has been left without update for seven years and many links on the page redirect to data.maine.gov. However, not all of the data provided on this portal was copied over to the new portal, and there is no redirect for the home page or information on DataShare that points users towards the updated information. The previous New Hampshire site, nhopengovt.org, is now an insurance sales site. The data that was once available through this portal does not seem to have been preserved anywhere online, so New Hampshire citizens have lost easy access to the financial transparency data this portal had provided. Other sites, most notably data.gov, still link to the nhopengovt.org address, highlighting the problem of having an open data site run by an organization outside the government that may allow the URL to be purchased by other organizations for unrelated purposes. Both portals show the need for a plan to retire an open data site, rather than allowing it to go fallow.

We do not see clear evidence of the influence of the federal government’s new attitude towards open data affecting the state level portals. There are, however, two instances of change where the timing coincides with the change in administration. The removal of Vermont’s Governor’s dashboard site occurred in the same time period as the 2016 election. Rather than reflecting the federal change, we think this removal is more likely to do with the priorities of Vermont’s new governor elected at that time. In Maine, the last update of data.maine.gov was November 6, two days before the presidential election. While this is in no way a conclusive signal that the state is moving away from open data because of the federal administration, the timing is co-
incidental enough to merit further consideration, especially since there was no co-occurring change in state government.

5.5 Frameworks

While the Open Data Policy Framework and Open Data Barometer are the most appropriate tools we were able to find currently in the literature for exploring and comparing these portals, neither tool is ideally suited to the examination of state-level open data. This is consistent with the cautions about applying evaluation tools outside of their context expressed by Susha et al. (2015). While the Open Data Policy Framework created by Zuiderwijk and Janssen (2014) raises many important questions about the open data portal, its criteria assume a more developed portal, and therefore many of its questions are unanswerable for the portals studied. This tool also seems as if it would be more useful for a self-study or study in which the researchers had the ability to personally interview portal staff, as many questions require information that is unpublished, or information that requires exhaustive searching, such as inventorying any usage of the data, where it is difficult to determine whether all relevant information has been found. A tool more specifically suited to the aims of this study would consider details of the published policies governing the portal, such as a specific checklist of policy elements and incorporate more factors from the political situation surrounding the portal, such as political movements active in the state.

The Open Data Barometer was too wide-reaching for this study, leading us to modify it by narrowing the focus to only information provided on the portal. Even with this modification, this is an imperfect measure for state open data. Of the 15 categories of data the portal inventories, 6 categories were not present within any of the portals. While some of these categories could be useful additions to a state portal, like election results, several are inapplicable in this context, such as international trade data. This does not necessarily mean, however, that these metrics would need to be edited out in
order to be useful for state-level research, as international trade data might be useful and available in a large state like California. This highlights a major difficulty of research on a state level—what is appropriate for evaluating the portal of a large, resource rich state is unlikely to translate to small rural states such as the ones in this study.
Chapter 6

Conclusions

6.1 The State of State Open Data

In the northern New England states, the future of the state-level data portals is uncertain. Two to three years after the creation of the portals, Vermont continues to update and improve its portal, Maine seems to have stopped its updates, and New Hampshire has made almost no updates to its original directory. Vermont’s model, a single portal with software that supports users’ creation of visualizations, seems to be more effective for users than Maine’s dual portals or New Hampshire’s list site.

The most common goal for the portals is transparency, but exactly what transparency looks like is unclear. Is simply providing open data enough? Or would true transparency require advertising the available information and making sure users have the technical knowledge to make use of the data? Defining what transparency means in this context and what transparency is supposed to accomplish would help focus any future portal development on where it could have the largest impact. The same principle applies to the additional portal goals of education, fiscal responsibility, and participation—without clearly defining what these goals will achieve, it is difficult to determine what a successful portal will look like.

Several common issues occur across portals that limit the ability of the
general public to use the data. One issue is the incomplete and inconsistent licensing information. Without making the terms for reuse extremely clear, few will understand whether or not they have a legal right to create new products with the data. The ambiguity of portal ownership is also a concern. Although all of the portals have some relationship with their state’s information technology department, finding a contact person who is able to answer questions about the portal is extremely challenging. It seems as if the portals are not considered a critical project in any of the states and therefore don’t truly belong to any specialized working unit. Related to this difficulty contacting any portal leadership is a lack of technical support. Portals do not provide any assistance to users having difficulty navigating the portal. The training materials the Maine and Vermont portals provide are minimal and hidden, decreasing the likelihood that they will be used.

Because of these issues, none of the portals show evidence of leading to noteworthy reuse of the data. While Vermont’s portal leads in use of the data, this is achieved with only a small number of visualizations and no larger projects. None of the states have significantly invested in advertising and promoting the portals. The datasets provided can also affect the level of reuse. Most information on the portal seems to have been selected based on the convenience of the portal managers, leading to some datasets with less value being included while datasets that would be more technologically challenging to include but have a large potential for reuse are excluded from the portal. Some datasets that are only on one portal, such as licenses for tradespeople and childcares on the Vermont portal, show a high level of community interest based on number of views. This could be an area of expansion for other portals. Datasets that uniquely represent the state’s character, such as Maine’s lighthouse map, are also popular. The lack of citizen awareness of the portal and the creators’ prioritization of “easy” datasets combined with the previously identified lack of clarity in and support for the portal lead to the low levels of reuse.
While all states have some sort of Freedom of Information law to ensure that citizens have access to government information, this does not guarantee proactive moves such as uploading government data to a portal. New Hampshire is the only state that has specific open data legislation, which requires all data produced by the government to exist in an open data format. This has the advantage of ensuring that data is ready for publication and reuse without preparation by the data portal, but it does not mean that a given dataset will be available in general. Proposed legislation in Vermont would require the state government to remove identifying information from data and then publish it in a format conducive to analysis by the public. A legal framework like this could help open data be adopted as a part of the normal workflow of government, rather than being viewed as an extra that relies on the interest of government workers to continue.

The effects of the political changes at the federal level were not clearly evident. The change on the portals that coincided most clearly with the change in administration is in Maine. The last dataset added in Maine was added two days before the general election, and no updates or uploads have occurred since that point. Whether this is a direct result of the culture created by the federal government’s decreased support of open data, or merely a coincidence in timing is unclear, but it does signal that the initial push for open data in Maine has not led to a lasting inclusion of open data as a part of government operations. The influence of state politics on portals, however, was noticeable, whether this came in the form of Vermont’s removal of a transparency site with the change of governor or New Hampshire’s libertarian approach to the portals. The influence of political factors on open data portals suggests that in order to be viable long-term, portals would need to be woven into the state legal framework and become integrated with the daily operations of government.
6.2 Contribution of Research, Strengths and Weaknesses

The case studies provide a snapshot of some of the difficulties and successes that smaller states experience in creating and maintaining an open data portal. Since state portals have not been the object of as much research attention as national or municipal portals, identifying and articulating some of the challenges they face is an important first step to research.

One strength of this research is the similarity between states and its translation into a similarity of issues in the portals. The lack of engagement with the portals is unsurprising given the rural nature of the states. The northern New England states would need to devote serious resources to training and support in order to get those citizens who may not rely on information technology for their everyday work to be confident in their use of the portal.

A weakness of this study was the lack of input from government officials. We had hoped to be able to access internal documents related to the portals, in order to get clearer ideas of how they approached updating, anonymizing, and other issues, but the only department supervising a portal we were able to get in contact with, Maine’s, seemed to know almost nothing about the portal. Portal owners for New Hampshire and Vermont simply never replied, so whether they have these documents or not is unknown. In any case, the difficulty of getting in contact with the portals for research questions highlights their general lack of communication, which could prove problematic for users.

6.3 Application of Findings

This research could be useful to those in government tasked with creating or developing open data portals. Comparing the three portals highlights each portal’s strengths and weaknesses, a practice that could be extended by a portal developer to other state portals to gain additional ideas for improvement. The primary recommendation for data portals coming out of this
research is to extend transparency to the portal’s own practices by making information such as the criteria for inclusion, updating schedule, licensing information and contact information clearly available. The types of data included in the portals could also inspire portal developers to add new ideas. Open data practitioners should also note the importance of promotion, and include this in their plan from the outset. The findings of this research can also provide background to those seeking to develop open data laws in state legislature. Librarians can apply this research to a consideration of how they could contribute their knowledge organization and information literacy instruction skills to improve state open data.

6.4 Future Research

Because the research presented is preliminary, many areas for future research can be explored. One major potential avenue is looking at the effect of national political shifts throughout a larger sample of portals. Tracking changes in many state portals over time and correlating them with differing political climates could provide interesting insights into the effect of national government policies on state level portals. If more portals show Maine’s pattern of ceasing updates in the time period of the 2016 election, this could provide evidence that there is a political factor in this change. Related to the consideration of the national political context, the effect of the politics within a state could be an interesting area of research. This study included open data portals with conservative, liberal and libertarian influences, and exploring how these different guiding ideologies affect portals of a larger sample of states could clarify the relationship between politics and product.

Another area worth future investigation is the correlation between successful municipal open data portals and successful state portals. In this study, Vermont had the most developed portal, and was also the only state to have a municipal-level open data site in any of its cities. If these conditions occur together frequently, it could also reveal what underlying factors may lead to
the success of both levels of portal, whether this is an active citizen interest, champions in government, or another driving force.

Further research could also consider how the regional cooperatives such as those municipal cooperatives studied by Lassinantti et al. (2014) could be translated into state level portals. Although it might be difficult for different states with different legal frameworks for open data to share one portal, there could be areas in which they could collaborate. For example, the New England portals all share a lack of support for data users. A possible solution could be staff tasked with assisting users for multiple portals, so the cost would be shared. This cooperative model might also require more standardization of design between portals. Research building on municipal portal research and examining the feasibility at the state level could improve service delivery.

Beyond this, more exploration of the general usefulness of state data portals is necessary. The portals investigated in this study were not heavily used, and did not lead to the development of applications or projects. Whether this is an artefact of these being smaller, less developed portals or whether this is common to many state portals is an important question for open data scholarship. If state data portals are generally ineffective, a new approach to sub-federal open data in areas without large municipalities would need to be considered.

To support the investigation of the usefulness of state data portals, it may also be necessary to develop evaluation tools specifically attuned to the state level. Some of the dataset types in the Open Data Barometer were clearly inapplicable to the state context, suggesting that a method developed for evaluating state level data might provide more useful insights. If this method of evaluation is going to investigate usefulness, it may need to move beyond tallying of datasets to include measures of use and engagement.
References


Appendix A

Dictionary for Open Data Motive Coding

All words should also be counted if occurring in a different form eg. transparency or transparent

- **Economic**: Financial, savings, dollars, money, spending
- **Innovation**: Innovate, rethink, streamline,
- **Transparency**: Accountability, right to know, disclosure
- **Participation**: Discussion, interactive, share, reuse, remix, create
- **Education**: Knowledge, student, educator, class, training