Experimenting on Mapping the Digital Divide in Canada: The Data Desert and The Unreported Inequality

Zein Hajj-Ali*
Leslie Zijun Rao*
zein.hajjali@carleton.ca
leslierao@cmail.carleton.ca
Carleton University
Ottawa, ON, Canada

CCS CONCEPTS

• **Social and professional topics** → **Broadband access**; **Geographic characteristics**.

KEYWORDS

Digital Divide, Broadband, Internet Access, Online Learning

ACM Reference Format:
Zein Hajj-Ali and Leslie Zijun Rao. 2021. Experimenting on Mapping the Digital Divide in Canada: The Data Desert and The Unreported Inequality. In *Proceedings of ACM Conference (Conference'17).* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

The digital divide is a multidimensional concept that embeds complicated social dynamics. It often refers to the Internet "haves" and the "have-nots," which demonstrates the unequal access to knowledge and opportunities through disproportionate access to information and communication technologies. Various factors can contribute to the digital discrepancies, such as lacking broadband infrastructures to cover the region, the shortage of good quality network services, or lacking proper digital devices to access Internet services.

The Government of Canada has devised a series of programs to reduce the digital divide, including expanding internet infrastructures and increasing the affordability of internet services. However, the problem still remains unsolved in various areas. Inaccessibility issues are frequently being reported by Canadians across the country, while the situation is often unmonitored, as the digital divide and many of its related factors are not officially recorded. The disparity between rural and urban areas is the most commonly recognized trait of Canada's digital divide. However, the further details beyond this dichotomous scope are not often discussed in the public, because the detailed information and data are generally absent from the public sphere.

Our project is an attempt to map and analyze the digital divide in Canada by using the available data in the public, and to examine the issues that prevent us (as well as the public in general) from obtaining an accurate assessment of digital divide. Correspondingly, this project is conducted based on two central research questions:

- How to map digital divide in Canada and what information about the digital divide can be derived from the analysis?
- What are the barriers that prevent us from obtaining an accurate assessment of digital divide in Canada?

The following report will be arranged into five sections: the first section is a literature review of the multidimensional concept of the digital divide; the second is the methodology we adopted to assess the digital divide in Canada; the third is the illustration of our attempts to map the digital divide, as well as the data results derived from the data analysis; the fourth is the observation and analysis of the mapping experiment, in which we reflect on the process, further discuss why the data analysis did not render sufficient new knowledge about digital divide, and identify the barriers that prevented us from gaining a more accurate assessment of the digital

*Both authors contributed equally to this research.

Conference'17, July 2017, Washington, DC, USA

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

divide; lastly, we will conclude with a series of recommendations for the improvement of the public knowledge about the digital divide in Canada.

Due to the limited semester time frame, this project could potentially miss dataset resources and data processing tactics that might be more optimal for gaining better data results from the data mapping analysis experiment. With the scope that we had conducted the mapping experiment, our report attempts to reveal the lack of easy access to usable and reliable datasets related to the digital divide in Canada, which is a major barrier that fundamentally prevented us—as the public—from having more detailed understanding of the digital divide.

2 THE DIGITAL DIVIDE AS A MULTIDIMENSIONAL CONCEPT

The digital divide is a multidimensional concept with complex social implications. This term is raised since the 1990s with the rapid expansion of the networked computer technology, described as "discrepancies between social groups in access to, use of, and empowerment by networked computers and other digital tools [10]." Such digital discrepancies are caused by various factors and are addressed differently by a series of alternative definitions.

The alternative definitions can be roughly categorized into three types of access [4]. First, the physical access divide, refers to whether or not one has the necessary hard and software to access the Internet, often correlated with the demographic factors such as race, education and income. Here, both the availability and affordability of digital resources can lead to the physical divide. Second, the skill divide evokes the inquiry beyond the provision of physical access, shedding light on technical literacy (how to use computers) and the information literacy (the ability to navigate, evaluate and use information)—necessary educational resources and technological awareness must be facilitated to close the digital divide. Third, the "actual" usage divide focuses on whether ones can appropriate technologies for their own purposes. Quan-Haase [10] further divided the usage access: the economic opportunity divide (if ones are advantageous to economic opportunities by having the access), the democratic divide (if ones can use the Internet for political engagement) and the social media divide (if ones are capable of identity management and creating social and cultural capital) [5]. These sub-divides specifically address the different outcomes that could derive from the Internet use.

Norris [6] notes that digital networked technology embeds the potential to broaden the access to information and communications. By fostering the capability to leverage digital technologies, the remote and poor areas can join in the information age economically, culturally and socially. That is, the digital divide is not merely a problem of lacking infrastructures and information, but also a problem of equal right to development in the long term.

The physical access divide is the most basic digital divide, it is the entry-level barrier to pass in order to use the Internet. Without the necessary physical infrastructures and digital tools, including the broadband network, appropriate quality of Internet services and digital devices, people in the given region could suffer from unequal informational access and slow development in the long term. For such social significance, the digital divide mapping experiment will be conducted with the goal to examine the physical access in Canada.

3 METHODOLOGY

To illustrate the process of measuring the physical access divide in Canada, we first introduce our dataset exploration, identify three datasets that are related to the examination of the physical access divide: 1.) National Broadband Data, 2.) Pseudo-Household Demographic Distribution, and 3.) the 2016 Census of Canada. All three datasets are official datasets made publicly available by the Government of Canada through the Open Data program. We then detail the procedures and techniques used to pre-process and analyse the data. This includes merging the relevant sections of the datasets, cleaning up the records, and performing some calculations to measure the level of correlation between the demographic data found in the Census dataset and the level of internet access from the National Broadband Data.

4 THE MAPPING EXPERIMENT

4.1 Dataset Description

To identify datasets that can be applied to analyze the physical access divide in Canada, we explored a series of reports and data that are made available by various government agencies, Internet service providers (ISPs) and non-profit organizations. Due to the time constraint of this project, we might miss some other useful and higher-quality datasets in the public. We decided to chose the following three datasets as the raw material applied to the data analysis of the physical access divide in Canada by considering the fit of the research topic and the overall quality.

4.1.1 National Broadband Data. The National Broadband Data [12], or NBD, is the main dataset chosen to apply to the next step data analysis of the physical access divide in Canada. The National Broadband Data was produced by Innovation, Science and Economic Development Canada (ISED) and released on the Government of Canada open data portal. The dataset records the status of broadband coverage across the country in accord to 6 levels of speed: None, <5/1, 5/1,

10/2, 25/5, 50/10 (download/upload) in Mbps—speed level data are provided to both wired and wireless broadband.

The dataset records data by splitting Canada's geography into hexagonal areas, each with an area size of around 25 square km. The broadband speeds that are recorded in the dataset for each hexagonal area are defined by the speeds that are available for more than 75% of the dwellings in that area. This way of counting overlooks the outliers where some households have speeds significantly above or below the average available speeds within the area. The dataset also includes the available broadband speeds along 250m segments of road, but we did not adopt this part of data, because it repeats the data that are already available in the hexagonal representation in a manner that is less useful to the goals of the project. The segment of the NBD that represents Ontario includes over 1.3 million records. The dataset uses an ID called the "PHH_ID", which is a unique identifier for each of the hexagons that correspond to a pseudo-household representative point. Figure 1 shows an example of how the hexagons are represented on a geographical map.

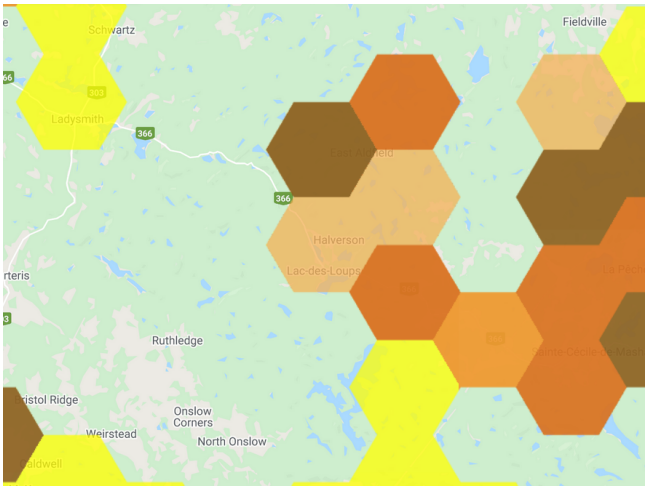

Figure 1: Example of Hexagonal Area Representation in National Broadband Data

4.1.2 Pseudo-Household Demographic Distribution. The Pseudo-Household Demographic Distribution [13] dataset, or PHH, is produced by the ISED, published on the Open Canada website. It will be used as a supplementary dataset to the NBD. This dataset will allow us to associate the hexagonal area IDs of the NBD to geographical locations. Each record in the PHH data includes the same hexagon identifier found in the NBD along with their coordinates and the ID of the corresponding 2016 Census subdivisions. The inclusion of the Census subdivision ID makes possible the retrieval of more detailed demographic data, as the Census data is available freely on the Statistics Canada website.

The PHH also includes the population of each recorded hexagonal area. This is helpful to proportion some of the Census data since the geographic areas of the datasets do not line up perfectly. Since the geospatial data is partitioned in the same way as the NBD, the segment of the PHH that represents Ontario is over 1.3 million records, corresponding to the NBD records. The dataset includes a number of columns, one of which is the "DBUID" which is the dissemination block ID, a 10 digit number that is used to identify the smallest scope of Census data records. The "PHH_ID" of each record is a unique ID that corresponds to a pseudo-household representative point and can be used to match to the "PHH_ID"s of the NBD. The dataset also includes latitude and longitude, which can potentially provide us more precise location information of the physical access divide.

4.1.3 2016 Census of Canada. The 2016 Census of Canada [1] released by Statistics Canada as partitioned datasets, and it was used as the second main dataset. To limit the scope of the data analysis, the digital divide mapping will be primarily focused on Ontario as an illustrative example. The finer-grained demographic data of Ontario are available in the 2016 Census, including information on population, age, number of dwellings, income, language, and more. The data chosen for this project covered all of Ontario. The raw dataset included over 46 million records for Ontario alone, and we used around 27.3 million of them. The dataset uses a few different columns to represent all the variables.

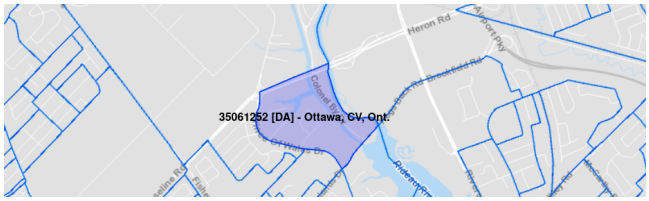

Figure 2: Example of Dissemination Area Representation in 2016 Census of Canada

There are a number of different ID formats included with each record, but the only one of note is the "GEO_CODE". The "GEO_CODE" could be presented in many different formats, but the ID that points to the finest grain information available publicly is the dissemination area ID. This is a 8 digit number that points to areas of a "relatively stable geographic unit [2]," made up of one or multiple dissemination blocks, totaling a population of somewhere between 400-700. An example of the dissemination area represented geographically can be seen in Figure 2. The dataset also has a column named "DIM: Profile of dissemination areas" which is used as the variable name identifier. Rather than creating hundreds of columns to represent each characteristic that is recorded

Characteristic	Total	Male	Female
	Counts (unless ot		
Income of individuals in 2015			
Total - Income statistics in 2015 for the population aged 15 years and over in private households - 100% data [21]	720	380	340
Number of total income recipients aged 15 years and over in private households - 100% data	690	355	335
Median total income in 2015 among recipients ($)	33,856	42,752	24,896
Number of after-tax income recipients aged 15 years and over in private households - 100% data	685	355	330

Figure 3: Census data as it is found on the Statistics Canada website

GEO_CODE (POR)	DIM: Profile of Dissemination Areas (2247)	Dim: Sex (3): Member ID: [1]: Total - Sex	Dim: Sex (3): Member ID: [2]: Male	Dim: Sex (3): Member ID: [3]: Female
35190002	Employment income (%)	70.0	71.8	67.8
35180802	Athabaskan languages, n.i.e.	0	0	0
35480074	Marathi	0	0	0
35200332	Bachelor's degree	115	40	70
35211401	Philippines	60	35	20
35130072	English and non-official language	0	0	0
35600210	Total - After-tax income groups in 2015 for th...	x	x	x
35390832	65 to 69 years	20	10	10
35211720	Only regular maintenance or minor repairs needed	185	...	...
35202133	Dakota	0	0	0

Figure 4: Example of the format of the publicly available Census data from a CSV file

for each record, a column was made to contain the name of the variable being looked at. Three other columns were made to hold the values of that variable for each record. Figure 3 shows a representation of how the data appears on the Census website, while Figure 4 shows the format of a segment of the downloaded data.

4.2 Data Processing

When working with the National Broadband Data, we noticed a number of redundant columns and variables. The dataset listed the availability of each level of internet speed for both wired and wireless broadband technologies as a separate column that takes a boolean value. As we were investigating the speed level of internet access, the only values we needed were the maximum available speeds and the technology used to achieve those speeds. In this way, the NBD was transformed to keep the information relevant to our purposes while reducing dimensionality and increasing readability. The "PHH_ID" column is kept to allow for merging with the Pseudo-Household Demographic Distribution data.

The transformed NBD was then merged with the PHH data to add the dissemination Block ID (DBUID) column as well as the latitude, longitude, and portion of the population. As a result, the merged data was cleaned and any records that are not in Ontario were removed. The "PHH_ID" was then dropped. Since the dissemination blocks are the smallest areas that are used in the Census, and the dissemination Areas are composed of up to 999 adjacent dissemination Blocks, the dissemination Block ID is equal to the dissemination Area ID but with three additional Block identification digits added on. Since the smallest unit of area found in the Census data is the dissemination Area, the last three digits were dropped from the merged data's "DBUID." The inclusion of the latitude and longitude records allowed us to display the maximum available speeds on a map of Ontario as seen in Figure 5.

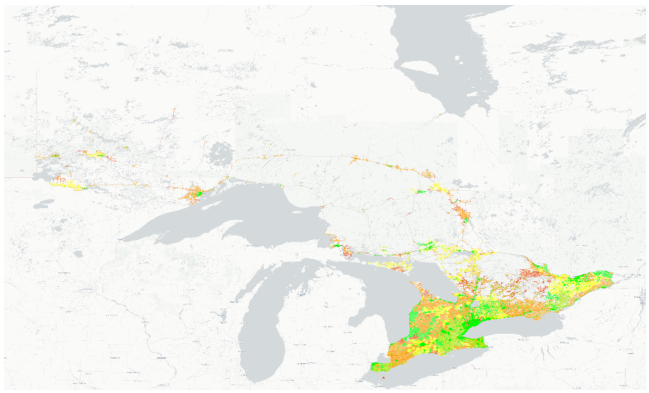

Figure 5: Visualization of the maximum available internet speeds across Ontario

The Census data, consisting of over 46 million records for Ontario, was around 4.37GB and therefore needed to be reduced for ease of use. Many columns were dropped, including: "CENSUS_YEAR," "GEO_LEVEL," "GEO_NAME,""GNR," "GNR_LF," and "ALT_GEO_CODE." According to Statistics Canada, the "GEO_CODE" for areas in Ontario all begin with the digits "35." Since we are looking at the most detailed data of the Census, we want only the "GEO_CODE" beginning with "35" and 8 digits long to filter out everything except the dissemination area codes. The Census data was then partitioned into smaller two-column data frames to make merging with the transformed NBD less resource intensive. This partitioning was done by searching the Census data "DIM: Profile of dissemination areas" column for the desired "characteristic" (as per the left side of Figure 3) and saving the "GEO_CODE" and the "Dim: Sex (3): Member ID: [1]: Total - Sex" values in two different columns. For example, if we were looking to create the dataframe containing the average age of the population at each "GEO_CODE," we would search the "DIM: Profile of dissemination areas" for "Average

age of population," and store the resulting "GEO_CODE" and "Dim: Sex (3): Member ID: [1]: Total - Sex" values of each found record in the "DBUID" and "Average age of population" columns respectively. This is due to the way the Census data is set up by Statistics Canada.

This process was done on the Census data to result in 11 dataframes that This process was done on the Census data to result in 11 dataframes that represent a wide range of demographic data: population in 2016, total private dwellings, total private dwellings occupied by usual residents, average age, median total income in 2015 among income recipients aged 15 years and over in private households, percentage of income that comes from government cash benefits, median total income of households in 2015, number of Canadian citizens, number of non-Canadian citizens, number of people that identify as aboriginal, total visible minority population. Each of the dataframes were then cleaned to remove any non-numerical or erroneous values. The dataframes were then merged together with the previously merged and the transformed NBD and PHH data. For the columns that deal with population data, such as "number of Canadian citizens" and "number of people that identify as aboriginal", their values were adjusted to match the proportional population of each record. This is needed due to the fact that the hexagonal areas of the NBD do not line up with the dissemination areas of the Census data. Therefore, a wider area may be covered by one or the other and the population data from the Census dataset needs to be adjusted. The resulting dataset included every record from the NBD, but added the 11 columns taken from the Census data.

Pearson, Kendall, and Spearman correlation tests were performed on the resulting dataset between the demographic data and the speed availability. Averages of the demographic data were also calculated and mapped along the different speed availabilities. Simple regression techniques were tried to find additional trends in the data.

4.3 Results

When looking at the results of the correlation tests, we can see that they do not show a particularly strong linear or monotonic relationship between the Census demographic variables and the level of available internet access. Figure 6 below shows a slight upward trend between the level of internet access and the average household income. Figure 7 shows a slight downward trend between the average age of the population and the level of internet access. The results of the regressions can be seen plotted in Figures 8 & 9, where the speed levels (0, 1, 2, 3, 4, 5) correspond to the available speeds (None, <5/1, 5/1, 10/2, 25/5, 50/10) download/upload in Mbps.

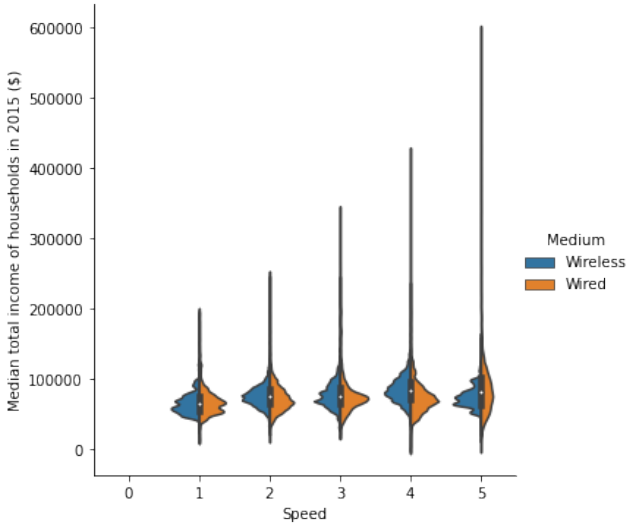

Figure 6: Violin plot of the level of internet speed available and the median household income

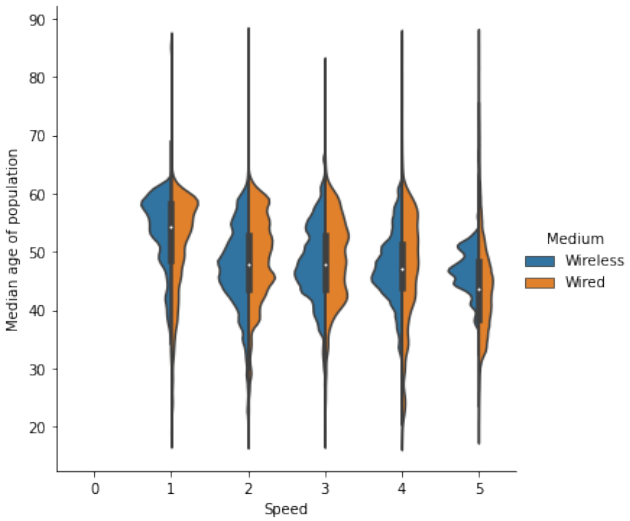

Figure 7: Violin plot of the level of internet speed available and the median age of the population

5 OBSERVATIONS AND ANALYSIS

5.1 Data Results: The Lack of Significant Findings

As the data mapping experiment demonstrates, the findings we can derive from processing the NBD, the PHH, and the Census demographic data are very limited. In the data result section, the variables that reveal correlational relationship with physical access divide include income level, aboriginal

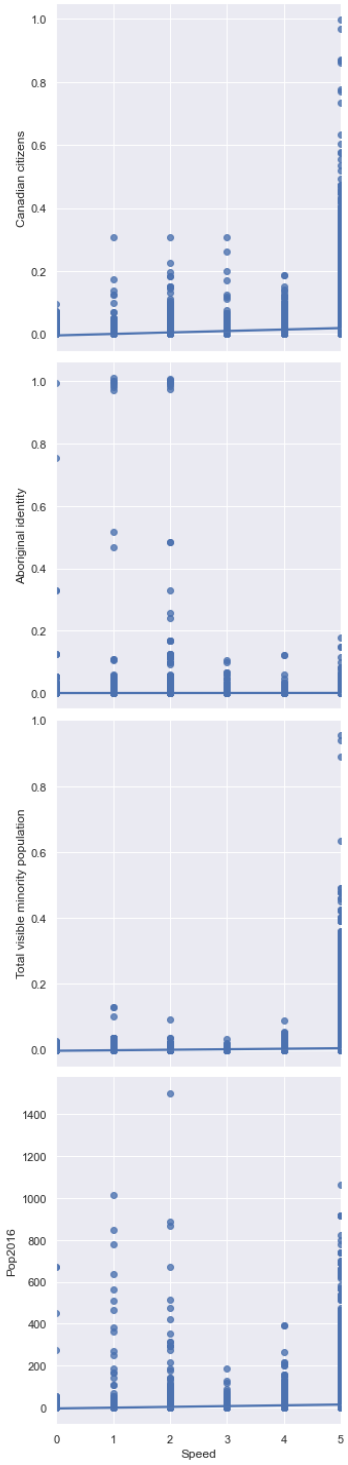

Figure 8: Plots of the regression lines on the data for some of the demographic variables

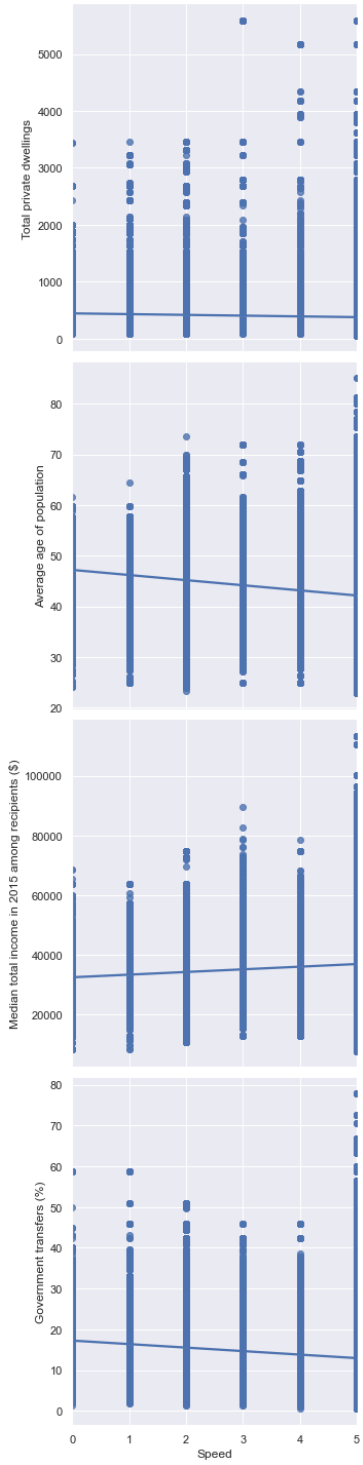

Figure 9: Plots of the regression lines on the data for some of the demographic variables

identity and age. These demographic variables have been recognized as the factors that characterize the general trends of the digital divide [11][8][4]: income level tends to affect the

households' affordability of high-quality Internet services and digital devices; aboriginal communities tend to locate in remote and rural areas, therefore the ISPs (Internet service providers) tend to be disincentivized to invest on broadband infrastructures; individuals' age tends to impact the individuals' technology awareness and the digital literacy, leading to the discrepancies of Internet access. Though this knowledge is not wrong, it is not new either, as it has already been well-circulated in the Canadian public for long. However, even if these demographic factors are commonly recognized as correlating to the digital divide, the data analysis based on these two open datasets still failed to indicate strong correlation.

The CRTC is the main government agency in researching and reporting the progress of the digital divide in Canada. In its annual Communications Monitoring Report, the NBD, the Census demography and the PHH are the significant materials applied for analysis [11]. As our data mapping experiment barely delivered meaningful findings, we consider that the quality and usability of these open government datasets related to the digital divide tends to be questionable.

5.2 The Lack of Usable Datasets

The overall absence of usable datasets that speaks to the digital divide in Canada is a crucial reason for the lack of significant findings derived from the data mapping experiment. In Canada, the digital divide data primarily come from two sources [8]: the government and the Internet service providers (ISPs). The CRTC is the main official institution who provides the most comprehensive data to the public. Its yearly published Communication Monitoring Report is a main official source that focuses on discussing the digital divide. As principal actors to build connectivity and retail Internet services, the ISPs undoubtably hold the most fine-grained information about broadband employment, Internet service subscriptions and the speed (quality) of service. However, the ISPs" data are deemed as commercially sensitive so are not opened to the public, while complaints about the government"s open datasets are far from uncommon [8][10]. That is, in Canada, the publicly available data about the digital divide are not only suffering from low usability, but are also limited in the amount.

Take the CRTC as an example. In the 2020 annual report, the CRTC released 31 open datasets that address the topic of the digital divide. These datasets involve wide range of issues areas, and can be roughly categorized into three types:

- Internet infrastructure coverage ("mobile coverage" "LTE coverage" "broadband coverage")
- Service provision ("subscribers," "service coverage," "services coverage with different levels of speed")
- Community"s Internet access ("official language minority communities (OLMC)," "urban," "rural," "First Nation reserves")

Despite the richness of the issue areas covered, the actual data points contained in the datasets are limited, as they are only available in highly aggregated form: the majority of datasets are presented by percentage of household or population, with the finest geographic unit available is by province or territory; also, the digital divide is only measured in accord to four categories of communities.

With the data that only tell by province/territory, the digital disparity information we can derive from data analysis would at best reflect the already well-known urban and rural inequality, while the regional differences within a province/territory would not be able to be known by using the CRTC open data. That is also the reason we chose the NBD and the Census data for the data mapping experiment, as they are the very few open data with geographic units smaller than province/territory.

Besides, the inability to inform detailed regional differences also reflects on the lack of specificity in addressing community differences. The four community categories (urban, rural, First Nation, OLMC) are still mostly aligned to the urban/rural and rich/poor disparity. For the communities and populations who do not fit these general trends, for example, a high-income community located in a rural area, or a neighborhood in a big city while still constantly having bad network connection, their problems are invisible and so cannot be solved. The open government data principles suggest that open data should be made available in complete information and with the finest possible granularity [9]. The currently open data about the digital divide in Canada are too aggregated to be useful for the public research.

5.3 Mismatched Geographical Areas Between Data Sources

As mentioned in previous sections, the geographical areas used to partition the datasets differ from one dataset to the other. The NBD and PHH datasets both use a hexagonal area covering around 25 squared Km, while the Census data uses what is called the "dissemination area," which is an unspecified "relatively stable geographic unit" that covers a population of around 400 to 700. This was surprising at first, since all the data was sourced from obtained Government of Canada sources. This mismatch in the geographical units can be detrimental to the results of the project. Since the geographic units across datasets could not be perfectly matched, a loss of information in the data analysis became inevitable.

6 CONCLUSION

The concept of physical access divide sheds lights on the availability of the most basic and fundamental infrastructures, which not only enable individuals' use of the Internet, but also make possible for more equal distribution of benefits in the information age. In the ongoing COVID-19 pandemic, many occasions and services such as school, offices and groceries are transferred online, the access to reliable and affordable Internet services has increasingly become a necessity to everyday life. At this critical time, the impact of the digital divide is particularly highlighted. The ability to know the situation is fundamentally important to generate solutions to provide quality lives and wellbeing across various communities. Based on the data mapping experiment conducted in this project, we propose three recommendations to address the digital divide issues in Canada.

First, to provide wider scope and better quality of open data about the digital divide issues, and the data must be more disaggregated based on finer-grained geographic units in order to inform regional differences. The availability of finer-grained data for sources like the CRTC datasets and the open microdata file of the Canadian Internet Use Survey [3] would allow researchers to analyze the digital divide at a deeper level. According to the Government of Canada [7], open data should be able to support innovations and social research, promote public interests and increase government accountability. This project encountered the difficulties of finding appropriate datasets and extracting meaningful findings, showing that the insufficient data provided by the government impose a critical barrier that prevents the public from gaining better understanding—not only the conditions of the digital divide, but also the government's progress in solving the issues. Mitigating this data barrier is important to further address inequality by joining the public forces for community-based solutions and increasing the government's accountability in tackling the issue.

Second, to improve the usability of census data, which is the key source of information that a government can provide for researching demographic characteristics of the digital divide. There is a definite need for readjustment of the formatting of the census data. The use of one column to contain the key of the "characteristic" and another to contain the value can cause issues when two sub-characteristics have the same name/key. A way to fix this situation can be assigning a unique key or ID to each characteristic and sub-characteristic in the dataset. This would allow for easy access to specific characteristics in the datasets, and remove the duplicate naming of variables.

Third, to collect more comprehensive data by leveraging the census infrastructure to include questions about the digital divide in the census survey. A forward-thinking move on the part of Statistics Canada would be to add questions on the future census questionnaires to assess the level of internet access and access to digital devices. This would reduce the need for external and supplementary data sources when studying the digital divide.

REFERENCES

[1] Statistics Canada. 2016. Census Profile, 2016 Census. https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/details/download-telecharger/comp/page_dl-tc.cfm?Lang=E. Accessed: 2021-1-28.

[2] Statistics Canada. 2016. Dictionary, Census of Population, 2016; Dissemination area (DA). https://www12.statcan.gc.ca/census-recensement/2016/ref/dict/geo021-eng.cfm. Accessed: 2021-4-18.

[3] Statistics Canada. 2020. Canadian Internet Use Survey - Public Use Microdata File. https://www150.statcan.gc.ca/n1/en/catalogue/56M0003X. https://doi.org/10.25318/56m0003x-eng Accessed: 2021-4-18.

[4] Jan A. G. M. Dijk. 2017. Digital Divide: Impact of Access. *The International Encyclopedia of Media Effects* (2017), 1–11. https://doi.org/10.1002/9781118783764.wbieme0043

[5] Nicole B. Ellison, Charles Steinfield, and Cliff Lampe. 2007. The benefits of facebook "friends:" Social capital and college students' use of online social network sites. *Journal of Computer-Mediated Communication* 12, 4 (2007), 1143–1168. https://doi.org/10.1111/j.1083-6101.2007.00367.x

[6] Pippa Norris. 2001. *Digital Divide: Civic Engagement, Information Poverty, and the Internet Worldwide.* Cambridge University Press.

[7] Government of Canada. 2020. Open Data 101. https://open.canada.ca/en/open-data-principles. Accessed: 2021-04-18.

[8] Federation of Canadian Municipalities. 2014. Broadband Access in Rural Canada: The role of connectivity in building vibrant communities. https://fcm.ca/sites/default/files/documents/resources/report/broadband-access-rural-canada.pdf. Accessed: 2021-04-18.

[9] OpenGovermentData. 2007. The 8 Principles of Open Government Data. https://opengovdata.org. Accessed: 2021-04-18.

[10] Anabel Quan-Haase. 2013. *Technology and society : social networks, power, and inequality.* Oxford University Press, Don Mills, Ontario.

[11] Canadian Radio-television and Telecommunications Commission. 2020. Communications Monitoring Report. https://crtc.gc.ca/eng/publications/reports/policyMonitoring/2020/cmr4.htm#a2. Accessed: 2021-04-18.

[12] Daniel Winters and Yannick Lacharité. 2021. National Broadband Data. https://open.canada.ca/data/en/dataset/00a331db-121b-445d-b119-35dbbe3eedd9. Accessed: 2021-1-28.

[13] Daniel Winters and Patrick Pilon. 2020. Pseudo-Household Demographic Distribution. https://open.canada.ca/data/en/dataset/b3a1d603-19ca-466c-ae95-b5185e56addf. Accessed: 2021-1-28.