Social Media and GIScience: collection, analysis, and visualization of user-generated spatial data

by

Matthew K. McKitrick

B.A., University of Ottawa, 2019

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Science

in the
Department of Geography
Faculty of Environment

SIMON FRASER UNIVERSITY
Summer 2021

Declaration of Committee

Name: **Matthew K. McKitrick**

Degree: **Master of Science (Geography)**

Title: **Social Media and GIScience: collection, analysis, and visualization of user-generated spatial data**

Committee:
Chair: **Kirsten Zickfeld**
Distinguished Professor, Geography

Nadine Schuurman
Supervisor
Professor, Geography

Valorie Crooks
Committee Member
Professor, Geography

Meghan Winters
Examiner
Associate Professor, Health Sciences

Abstract

Over the last decade, social media platforms have eclipsed the height of popular culture and communication technology, which, in combination with widespread access to GIS-enabled hardware (i.e. mobile phones), has resulted in the continuous creation of massive amounts of user-generated spatial data. This thesis explores how social media data have been utilized in GIS research and provides a commentary on the impacts of this next iteration of technological change with respect to GIScience. First, the roots of GIS technology are traced to set the stage for the examination of social media as a technological catalyst for change in GIScience. Next, a scoping review is conducted to gather and synthesize a summary of methods used to collect, analyze, and visualize this data. Finally, a case study exploring the spatio-temporality of crowdfunding behaviours in Canada during the COVID-19 pandemic is presented to demonstrate the utility of social media data in spatial research.

Keywords: social media; spatial-temporal analysis; location-based social media data; geographic information systems; cartographic visualization; crowdsourced data collection

Acknowledgements

First and foremost I would like to thank my supervisor Dr. Nadine Schuurman, who has been there with me every step of the way. Without her encouragement and wise advice over the course of my degree I would not have been able to complete this research. Her focus on personal well being, and consistent delivery of calm, collected counsel no matter the circumstances are what gave me the courage and stamina to complete this thesis. Thank you, Nadine, for helping me find my voice. I would also like to extend my sincerest gratitude to Dr. Valorie Crooks, not only for her careful reading, thoughtful critique, and examination of this thesis, but also for sharing her insights and intellectual curiosity in health geography and qualitative methods. Finally, I would like to thank Dr. Megan Winters for her playing a role in the examination of this thesis.

This adventure over the past 2 years at Simon Fraser would have never been possible without the unconditional love and support of my parents, David and Janet, who always believed in me and to whom this work is dedicated.

Table of Contents

List of Tables

List of Figures

List of Acronyms

API	Application Program Interface
CERB	Canada Emergency Response Benefit
COVID-19	SARS-CoV-2
GIS	Geographic Information Systems
GPS	Global positioning system
ID	Identifier
KDE	Kernel Density Estimations
LBSMD	Location-based social media data
PPE	Personal Protective Equipment
SARS	Severe-acute respiratory illness
SFU	Simon Fraser University
ST	Spatio-temporal / Space-time

Chapter 1.

Introduction

From the GPS satellite, to the personal computer, to the world wide web, geographic information systems (GIS) have always evolved in tandem with the innovation and proliferation of other technologies that provided the capabilities to create, manipulate, or communicate spatial data (1). While GIS itself underwent many changes and innovations in internally, it was largely these external technological factors that dictated the capabilities of GIS, and more generally, popular access to GIS technology itself (2). First, global positioning system (GPS) satellites allowed for the collection of massive amounts of accurate spatial information, while basic computing provided the tools to perform spatial calculations at a much faster rate than the human brain. The advent of the affordable personal computer meant that GIS were now available to far more users, and subsequently the creation of the internet meant that data and software could be communicated at faster and faster speeds, and that the average internet user could now access basic GIS web mapping applications. Now, over 2 decades since the creation of web 2.0 technology, the next iteration of this coevolution is taking shape with a new technological force at the fore: social media applications.

Just as the development of GIS has been dictated in large part by the technological capabilities of a given era, so too has the development of social media been largely if not solely dependent on these same external technological factors. In this way, when referencing the impacts of social media platforms, the technological foundations upon which social media operate are implicitly included in this impact assessment. Mobile smartphones with GPS and widespread access to mobile and home internet are the main examples, however just about any innovation related to internet communications, computer storage, processing power, camera technology, and cloud computing could be included here, and that's without mentioning breakthroughs in the social media software design itself, nor the mathematical foundations in machine learning and neural networks from which they sprang. The point is not that social media platforms are the solely responsible technological driving force behind external pressures to GIS, but instead that they sit atop a mountain of innovation and therefore can serve as an analytic focal point for how this amalgamation of powerful hardware and software advancements over the last 20 years have, and will continue to shape not only the GIS toolkit, but also GIScience more broadly as a discipline.

This thesis will explore the GIS-social media interface by identifying areas of GIScience research that utilize social media platforms as a data resource, and investigating how social media has incorporated many GIS elements and in doing so has expanded GIS's reach in the popular culture. First, the specific methods used by researchers to collect, spatially analyze, and visualize data from social media are examined and categorized to provide a functional understanding of the broad research trends in this area of GIScience. Next, a case study is provided within which social media data are collected and used as the main data resource to explore the spatio-temporality of the cultural phenomenon of charitable crowdfunding in the context of the COVID-19 pandemic. Together, there chapters will serve as evidence to the claim that social media and GIS have become increasingly interlinked, both at the scientific level, where GIScience has increasingly embraced social media as a spatial data resource, and at the societal level, where social media has introduced elements of GIS to the wider populace. In doing so, this thesis will serve as a synthesis of GIScience research in the domain of the GIS-social media interface.

GIS and Technological Change

In order to understand the relationship between GIS and social media, we must first observe how GIS technology and the discipline of GIScience have been impacted over the course of the last half-century by advancements in the broader technological landscape. The foundational 2015 text by Longley describes the evolution of GIS between 1960 and 2011 in three distinct eras (2). First, the era of innovation in the 60's and early 70's saw the launch of the first GIS, the establishment of the first academic, government, and commercial GIS institutions, and finally the launch of Landsat, the first of many civilian remote-sensing satellites (2). These innovations paved the way for what we now know as modern geographic information systems (GIS), and laid the technical foundations upon which geographical information science (GIScience), the domain of scientific knowledge surrounding GISystems, would be built. The era of innovation provided the template from which modern GIS would evolve, endowing today's GIS practitioners with not only the tremendous power of these innovations, but also the limitations inherent to the technology of the day, none the least of which remains the primacy of 2D map-based analysis and visualization (3).

These early foundations would be expanded upon in the 80's and 90's, mostly as a direct result of institutional interest and demand for the products of GIS and spatial analysis. Dubbed 'the era of commercialization' (2), GIS were now affordable enough for the average business, university, or branch of government to own and operate, largely due to the rapid proliferation and decline in cost of the personal computer. It was at this stage that the term *geographical information science* was first used by Goodchild in 1992 to describe the growing body of scientific literature concerned with geographic concepts from a technical lens, an area distinct from yet completely interlinked with the traditional concept of GI-Systems as tools used by GISystem practitioners (4). The concept of GIScience was previously unexamined due to the fact that GISystems were specifically a collection of tools and technology utilized by GIS practitioners and spatial analysts, and thus were seen and marketed as 'slave-like processors' used to handle the heavy lifting in calculating spatial statistics and visualizing spatial data for use in other domains of geography and science more broadly (1). Goodchild's idea of GIScience came at a time when the average university could now afford to own and operate GIS applications of their own, which was leading to a growing body of scientific inquiry dealing with the results derived from the rapidly widening umbrella of GIS technology that lacked any substantial epistemological underpinnings. This new idea of a scientific discipline specific to the analytic fruits of the GIS environment, therefore, was driven at least in part by the advancements in the broader technology market, and sparked rigorous debate within the geographic and GIS communities that would animate these areas of study for years to come, and ultimately shape the theory of knowledge behind GIS as a whole (5).

The final era noted by Longley covers the first decade of the 21st century, and was dubbed *The Era of Openness and Pervasive Use*, largely due to the innovations surrounding the internet and mass global digital communications (2). While the era of commercialization saw the migration of GIS technology from being prohibitively expensive to something affordable at the level of the institution, the era of openness saw GIS launch from exclusivity in institutions and organizations to widespread public access to GIS via the internet. Web mapping websites like www.mapquest.com and 3d earth visualizers like Google Earth put the most fundamental tools of GIS visualization into the hands of nearly everyone, while open source GIS applications and cloud-based computing were beginning to take off, equating to nothing less than a quantum leap in the speed and volume at which spatial data could be created, processed, and shared by an also exponentially increasing GIS userbase. The 2015 text ends with its characterization of GIS eras here in the early years of the 2010's, thus leaving the decade largely open to interpretation as to how GIS has been shaped by external technological forces. Now, in 2021, with the decade behind us, we can begin to perform this analysis.

The Rise of Social Media

In addition to the proliferation of GIS, the 2000's and early 2010's also saw the creation of the significant majority of the large social media platforms that we know today. Of the top five social media platforms by users, all of them were created in 2010 or earlier (Facebook: 2004, YouTube: 2005, Whatsapp: 2009, Instagram: 2010, WeChat: 2010), and of the top 15, only three were created later than 2010 (Snapchat: 2011, Telegram: 2013, Tiktok: 2016) (6). The most popular, Facebook, has over 2.7 billion active users worldwide, and also owns two of the other top five major platforms, Instagram and Whatsapp with 1.2, and 2 billion users respectively (6). As of today in 2021, Facebook boasts a market capitalization of nearly $900 billion, a 10-fold increase since its initial stock market listing in 2012; Alphabet, Google's (and therefore Youtube's) parent company has a market cap. of $1.57 trillion. All this is not to point out a potential monopoly in social media communications, but instead to demonstrate both the exponential rise in popularity of digital social media, and the fact that nearly all of the big names in social media that sprang up around the same time (and often in similar locales) have remained the dominant platforms to this day, and show no signs of relinquishing their hold on the market.

This pivotal time period also saw the launch of the first mobile smartphones, at least in the form we recognize them today, most notably with the IPhone in 2007. The touchscreen, app-based model of the IPhone would go on to become the dominant smartphone design, with most smartphones sold today following this pattern. This design would also serve as a perfect platform for the use of social media applications, which as discussed previously were also being established at this time. It quickly became apparent that users preferred the mobile app versions of social media, as opposed to the original desktop-based website equivalents, as the share of social media visits via mobile devices reached 83% in 2019 (7). It was this turn to mobile technology that would finally introduce in earnest the of 'spatial angle' to social media, as it now allowed anyone with a smartphone (which as of 2019 represented 81% of US adults (8)) to digitally engage with their spatial whereabouts in real-time through a wide variety of smartphone applications, none the least of which remain social media applications.

Once location-based services were incorporated into social media applications, they often became an integral part of the user experience. Photos could now be tagged to a specific location on the earth, and the users personal location could be shared via a status update, check-in, or even by direct message. Given the massive scale and reach of social media and the smartphone, the GIS integration into the most basic aspects of these systems means that social media are likely one of the most common and profound ways most people will interact with GIS as a technology. As social media grew increasingly location-based, the convergence of GIS and social media became apparent, sparking interest and inquiry into the manifestations of this union and the challenges for GIScience as a discipline (9). Thus, given social media's significance as a conduit by which billions will utilize elements of GIS, it becomes important to consider the history of social media and how it may impact the evolution of GIS moving forward. Perhaps akin to the 2d-based reality endowed to the entire body of GIS by the technical limitations at the time of its founding (5), so too might the legacy of our current modes of social media, and therefore of GIS, be dictated by the technological limitations, both those known and those unknown, of today.

The GIS-Social Media Interface

The interface between GIS and social media is the point at which the two systems converge, and can be understood by looking at how each system has influence the other. First, the degree to which social media applications have utilized spatial services, i.e. the ways in which GIS has influenced social media, remains heterogenous across platforms, and varies widely depending on the nature of the platform. YouTube, for example, as a site is mostly aspatial except perhaps when videos are recommended as a consequence of one's location. A platform like Twitter, on the other hand, gives users the option to include a geotag containing one's location when tweeting, meaning that for some users the Twitter-verse is entirely spatially enabled with respect to the locations they are in when posting tweets. This periodic posting of one's location is the main mechanism by which GIS has been incorporated into social media, and, simultaneously, the main reason why social media platforms have recently become relevant as a data resource in GIScience analysis. While the latter will be the main focus of the body of this thesis, the former, i.e. the injection of GIS technology into social media and mobile communications more broadly, remains an important idea is reminiscent of past eras where a shift in the capabilities of technology has lead to a substantial demographic transition in terms of aggregate GIS usage across the world.

The integration of GIS into social media is extensive, however the client-side user experience of GIS is only one aspect of its relationship to GIScience. Much more applicable to applied GIS research and spatial analysis is the product of this location-based user experience, namely the location-based data generated by every user interaction within a given social media platform. This 'location-based social media data' (LBSMD) has been the subject of much GIS analysis in recent years, and has served as a significant data resource for a large body of research looking at everything from detecting earthquakes and the flu to mapping floods and sport-related crime events (10–13). The appeal of using social media data to discover and measure spatial phenomena lies in the fact that, first of all, social media data tends to have spatial qualities, second, that those spatial qualities usually also accompany some other content (text, image, video, etc.) created by a human, and third, because it can be accessed at in real-time at incredible volume and in many cases for free. As will be discussed in chapter 2, Twitter data was a favourite amongst researchers because it fulfills all these characteristics (a wide variety of human generated spatial data at low cost and high volumes and velocities).

Thesis Outline

This thesis is composed of four chapters. The first chapter provided a background on GIS, its relationship to technology, and the growing significance of social media both as a GIS tool and data source. The second chapter is a scoping review of the methods used to collect, spatially analyze, and visualize location-based social media data, which provides both a summary of the utilizations of LBSMD in the geographic literature and an operationalized understanding of the methods required to transform this data into GIS research outputs. The third chapter is a case study that demonstrates the applications of LBSMD in spatial research by gathering, analyzing, and visualizing data on COVID-19 related crowdfunding campaigns from the website gofundme.com. The fourth and final chapter summarizes the conclusions from chapters 2 and 3 and reflects on how this relates to themes discussed in the introduction. The thesis concludes with final remarks and insights on areas of potential future research.

Chapter 2.

Collecting, Analyzing, and Visualizing Location-Based Social Media Data: *review of methods in GIS-social media analysis*[1]

Abstract

With billions of active users, social media platforms now generate spatial data on a massive scale, which presents researchers with opportunities to use this data in new and innovative ways. There is an emerging body of literature that reports on the use of location-based social media data used to spatially analyze a range of phenomena. Due to the low cost, high availability, and substantial range of content and geographic coverage, location-based social media data have the potential to play a major role in GIS research. To this end, our scoping reviews and charts methods used by GIS researchers to collect, spatially analyze, and cartographically visualize location-based social media data.

Introduction

Social media platforms have risen to become some of the most popular mediums of communication in recent years, attracting billions of users worldwide. As a result, these platforms now produce a massive amount of user-generated content that can be downloaded by third parties. As some of this data is spatially referenced, this development has posed a promising opportunity for GIS researchers to gain insight into the spatial patterns and relationships that manifest online in social networks. There is a growing body of research that utilizes location-based social media data (LBSMD) to investigate these patterns and relationships, however the specific methods used to achieve meaningful analysis of LBSMD remain scattered and difficult to find. Of particular interest to GIS researchers are methods of collecting, spatially analyzing, and visualizing data gathered from social media platforms.

Thus, the objective of this review is to synthesize a summary of methods used by GIS researchers to exploit LBSMD as a data resource. To achieve this synthesis, a systematic

[1] Authors: Matthew K. McKitrick, Nadine Schuurman, and Valorie A. Crooks.

search strategy was employed to gather relevant GIS literature within which LBSMD was a significant research component. The articles were then reviewed to identify key trends and answer research questions surrounding the methods used to collect, analyze, and visualize LBSMD. To conclude, key themes within the literature are discussed and recommendations for future research are given.

LBSMD as a Spatial Data Resource

Technological innovation has always played a pivotal role in the development of GIS technologies and GIScience (1). In recent years, social media has proven to be one such technological force and has created a series of challenges and opportunities for GIS researchers (2). These opportunities have been driven in large part by widespread access to GPS-enabled mobile devices, cellular internet connections, and high performance computing power, which have facilitated the development of a variety of client-side location-based service technologies (3). These technologies allow users the opportunity to post their precise locations online without prior training, thus making it possible to digitally crowdsource spatial data from a large sample of users. This phenomenon was dubbed 'Citizens as Sensors' by Goodchild, as users were now acting as voluntary sensors in the field, each gathering small amounts of spatially referenced information that can later be aggregated and analyzed to reveal meaningful spatial patterns and relationships (4). As social media platforms became more and more popular as a result of access to these technologies, they also became some of the largest producers of this crowdsourced, human-generated spatial data.

The proportion of spatially referenced data produced by social media platforms remains relatively small (5), however, given the significant number of social media users, spatial data can still be accessed in volumes large enough to support spatial analysis (6). Typically, an LBSM-dataset presents as an aggregation of many individual digital records generated by users during their interactions with the platform. Each record contains content information pertaining to the activity or post of the user, the user's unique identifier (ID), as well as a geotag and timestamp denoting when and where the user performed the interaction. When gathered in large samples, analysis of these simple records can reveal spatio-temporal trends relevant to specific phenomena because the data can be preferentially collected by searching for key terms or specific times and location. Therefore, when available, LBSMD can serve as a versatile spatial data resource given its high volumes, large variety, and low cost, in contrast to spatial data from traditional sources, which usually have significant costs associated with access and production.

Location-based social media data is available at high volumes due to the sheer quantitative magnitude of social media users, high velocities due to its digital nature, and in extreme variety due to the wealth of topics discussed online, but understandably lack veracity, the fourth factor of big data. While the crowdsourced nature of LBSMD make it extremely plentiful and highly variable, it also creates questions surrounding the accuracy or viability of any data collected in this way, especially when assumptions are made and conclusions are drawn from crowdsourced big data in general (7). This is especially true with data derived from social media platforms, and researchers must consider the pitfalls and methodological hazards that can arise when performing secondary analysis on big data (8). The considerations surrounding user over-representation, exclusion of non-user individuals, reliance on secondary data analysis, and spatial accuracy are critically important and should always be considered when utilizing LBSMD, however these are not insurmountable obstacles. Indeed, LBSMD will never be a replacement for traditional data in most cases, however, when used correctly it can be a viable spatial data resource with widespread geographic coverage available at low cost.

Having established the potential utility of LBSMD for geographic research, this review will now explore the methods researchers have used to exploit LBSMD as a spatial data source. Following the general architecture for archiving and exploring LBSMD outlined in 2014 by Huang & Xu, where data moves from the platform, to the repository, to the processor and then to the end client (9), this review will focus on the three areas of data manipulation that are required for this workflow. The first of these areas, data collection, covers the methods required to gather LBSMD from a platform and save it locally in a repository. Next, the methods spatial analysis used by researchers to identify relevant spatio-temporal patterns and trends in the LBSMD will be summarized. Finally, transformation and visualization techniques will be investigated to produce a synthesis of methods used to communicate the findings of the preceding spatial analysis. These three sections will be informed by a thorough review of research articles gathered and selected through a systematic search process which will identify articles that utilize LBSMD as a data resource. Thus, the following review presents the scope of methods used to collect, analyze, and visualize LBSMD for GIS research.

Existing Literature Reviews

Prior to the main search for literature, a non-systematic search was performed to identify any similar reviews concerned with the utilization of social media data for GIS research. Although several reviews containing relevant information were gathered, few of them were directly related to GIScience and none were specifically concerned with the scope of methods discussed in this review. Related to geospatial analysis was Steiger et al. (2015), which performed a systematic review on spatiotemporal analyses of data specific to Twitter and explored which academic disciplines were focused on researching Twitter as well as applications and methods of analysis for Twitter data (3). Stock (2018) provides another systematic review concerned with methods of location mining from social media, and explored overall trends in which social media platforms data had been extracted from, what methods were used to extract location, and which research questions had been addressed by geographic data extracted from social media (10). While both of these reviews present detailed information on methods of manipulating LBSMD, the former focuses specifically on Twitter and the former specifically on the mining of location information, rather than LBSMD in general. Neither review encompasses a full GIS workflow, from collection through analysis and visualization.

A number of other reviews covered the utilization of social media data in researching niche topic areas. For example, Weigmann et al. (2020) provides a systematic review of the opportunities and risks of mining disaster data from social media (11), while Mirzaalin & Halpenny (2019) presents a systematic review of the role of social media analytics in hospitality and tourism (12). Further, Kamalich et al. (2020) shows a systematic review of social media utilization in emergency response to natural disasters (13) , while Wilkins et al. (2021) presents a systematic review of the uses of social media in informing park use management (14). While these reviews specifically deal with articles that utilize LBSMD, their scope is much more narrow and not done from a GIScience perspective. Although not a review, Stefanidis et al. (2013) provides an early outlook of the capabilities in harvesting and harnessing geospatial information from social media (15). This foundational article provides an overview of many of the same techniques to be discussed in this paper. As such, this review will consider the main themes present in this article and will serve as an update to this early work.

Methods

This review follows a scoping review protocol outlined by Arksey & O'Malley (2005) (15) and advanced by Pham et al. (2014) (16), which seeks to 'map' relevant literature in a field of interest. Drawing from the original Arksey & O'Malley framework, this review will seek to summarise and disseminate research findings contained within the body of literature, thereby providing a summative synthesis of evidence for practitioners and consumers who might otherwise lack the time or resources to perform a similar exercise (15). To this end, we pose the following research question: what methods have been used for the purposes of collecting, analyzing, and visualizing location-based social media data? This section outlines the procedure used to gather and analyze literature relevant to answering this question, and utilizes a systematic database literature search alongside collaborative literature review to generate a summative synthesis of findings. While this procedure is reflective of the scoping review protocols previously discussed, it deviates from the prescribed protocol because the included literature was reviewed only by a single reviewer.

Inclusion / Exclusion Criteria

This review included peer-reviewed articles that utilized location-based social media data as a resource for an analysis of a particular phenomenon. A broad definition of social media was mobilized to include any online platform where users can post content and directly interact with one another. Data was considered location-based when it contained locational metadata, i.e. latitude/longitude coordinate pairs, place names, zip codes, or any other spatial identifier for a location on the surface of the earth. Articles were included regardless of LBSMD sample size or phenomenon of interest, so long as they utilized LBSMD as a resource in the investigation or analysis of said phenomenon. Articles that did not perform a specific analysis using LBSMD were excluded. We did not include articles that provided commentary or theoretical input on the use of LBSMD and its general implications, rather, only articles that specifically gathered a sample of LBSMD from a platform and then used that sample to investigate an occurrence in the world were included. Articles that proposed an advanced method specific to a small area of the workflow, or articles concerned only with workflow performance and efficiency were also excluded.

Database Sources and Literature Search

A search strategy was developed and deployed across two online literature databases: Web of Science and GEOBASE. These databases were chosen to provide a broad body of evidence, with a focus on geographic literature. The search query included 3 terms relating to 'methods', 'social media' and 'GIS', each of which were queried with appropriate synonyms (see appendix for full search strategy). The search was designed to preferentially gather methods-focused papers utilizing GIS analysis on social media data. The query was searched on each database in January 2020, with a follow-up search being performed in May 2021 to gather any literature published since the previous search was conducted. No restrictions on language or date of publication were used in either search.

Article review and identification

A single reviewer screened all initially resulting articles by title and abstract for inclusion criteria. When it was unclear whether or not an article fit the inclusion criteria, multiple authors were consulted. Both reviewers then met to discuss and adjudicate any discrepancies between their respective title and abstract analysis. Articles deemed eligible for inclusion were then reviewed in full by a single reviewer, where data on methods of collection, analysis, and visualization were stored in a table. Fig. 1 illustrates the results of the paper collection process.

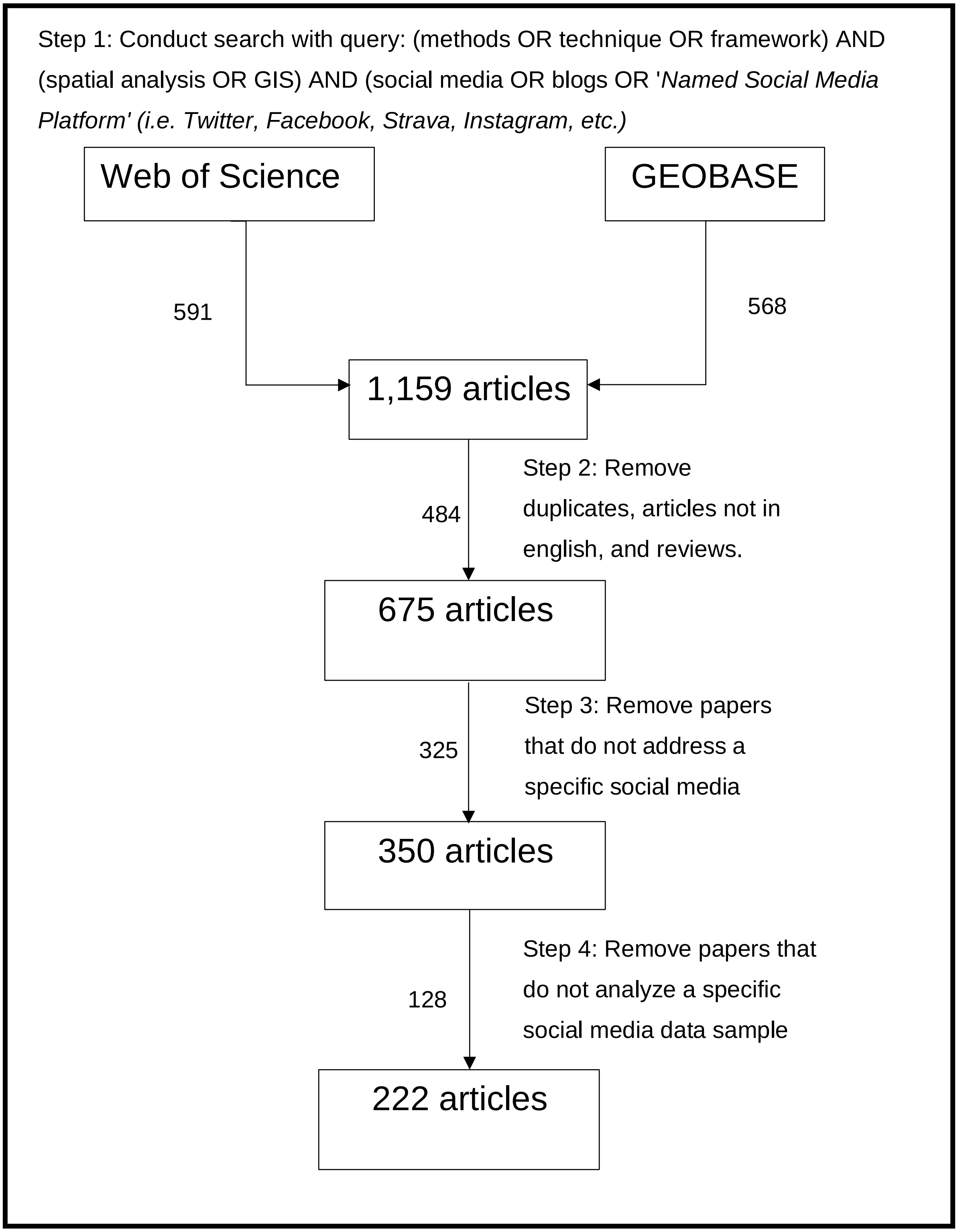

Fig. 1. Summary literature search diagram.
Articles were aggregated across two different searches in each database, and were subsequently filtered by title and abstract for full-text inclusion in the review.

Results

RQ1: What were the methods and variables of LBSMD collection?

Platforms as a spatial data source

Social media platforms used to collect LBSMD fell into three categories: Microblogs (e.g., Twitter, Weibo, Facebook), Photo/video sharing (e.g., Flickr, Instagram, YouTube), and Point-of-Interest (POI) (e.g., Open Street Map, Foursquare, Yelp). Microblogs were by far the most popular source of LBSMD: Twitter 'tweets' were used measure or model everything from natural disasters like earthquakes (17–19) and flooding (20–22) to influenza outbreaks (23–26), and when counted alongside studies that used Weibo, the two microblogs were the main data source for the large majority (>70%) of studies. In contrast, Facebook's inward facing, friend-oriented structure made it useful as a data source in only a small number of cases (27–29); Twitter, meanwhile, allowed anonymous viewing and download of public profile data by default, meaning users had to explicitly make their data private if they wanted it as such (30). Similarly, Flickr accounted for most of the studies using photo/video sharing platforms as a data source and was the second most popular data source overall. Data from collected from Flickr was used to model human activity (31–33) and natural landscapes (34–41). Data from POI platforms was most often used to measure site activity patterns (33,42–44) by comparing the number and frequency of visitation of POIs between areas, and to measure public knowledge and sentiment associated with specific locations (45–47).

Types of collection

There were two ways to access data on social media platforms: by gathering the data from the front-end of a platform, or by using an application-program interface (API). Collecting data from the front-end of a platform (i.e., using the user interface) involves either manually or algorithmically visiting and recording information from the platform the way a normal user would via a web browser or smartphone app. Collecting data algorithmically has been a contentious legal issue in the past (48) and platforms concerned over corporate data propriety have implemented digital security measures that have spurred advancements in proxy-based web crawling strategies (49). In contrast, utilizing an API entails communicating directly with the servers of a particular platform, and in the case of studies in this review, querying and downloading specific subsections of data that the API gathers automatically (9). The former method is rarely used, as it is a labour intensive method requiring either a custom built web scraping tool or the manual collection of data by a researcher. However, the manual method has been applied effectively for niche research objectives such as mapping tornado trajectories using manually classified Facebook posts (27), or by gathering information from otherwise non-indexed content such as YouTube videos (50–52). In contrast, the latter method of utilizing an API when free and available offers significant research incentives in terms of the volume and velocity of data, as well as in ease of data accessibility and acquisition, which are exactly the qualities that make the Twitter and Flickr APIs so popular and reliable when compared to manual or custom-built scraping strategies.

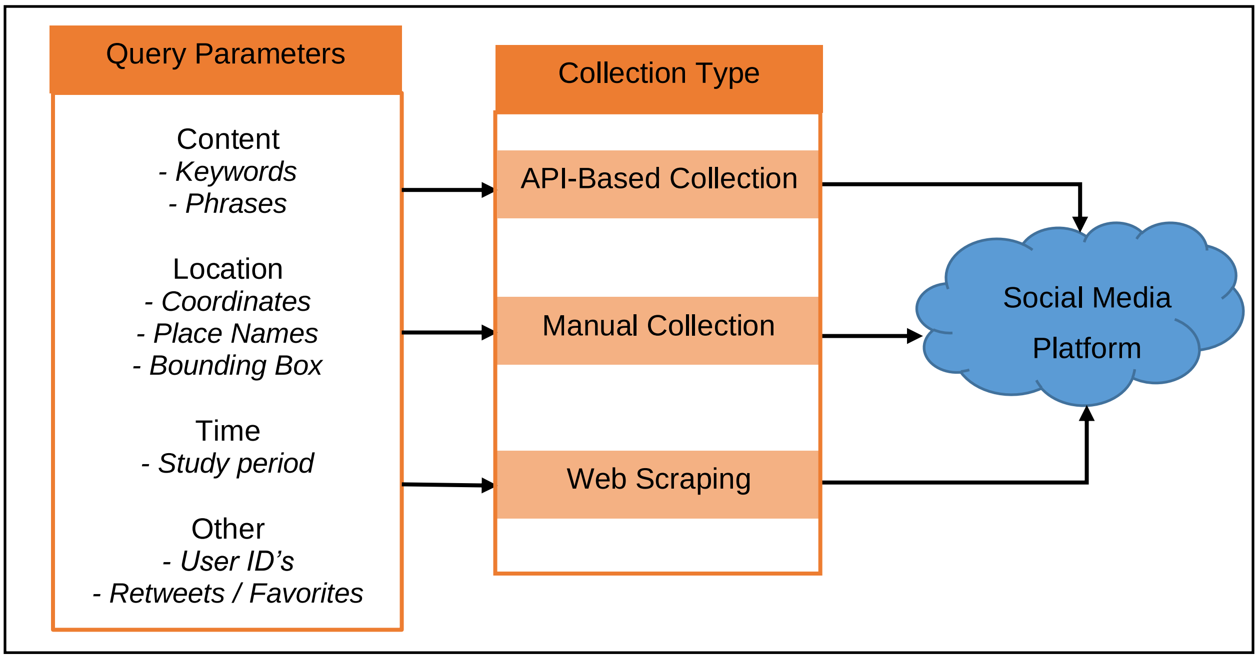

Fig. 2. Collection query parameters and the three types of data collection.
After various parameters are selected, they can be mobilized through either of the various collection types to access the desired data from the social media platform.

Data queries

Data can be retrieved from APIs using parameterized queries designed to gather desirable posts based on location (area research of interest), time (period of interest), content (the actual text or image content of the post), and a variety of other attributes associated with the post, such as user ID (17). All studies that gathered data from an API utilized some form of query parameter to filter their data. In many cases, researchers preferred to collect large swathes of data by using broad parameters, restricting collection only to time frame within a country or continent and relying on subsequent filtering to narrow down the corpus after it has been retrieved (53,54). Alternatively, queries can be quite specific across one or more parameters, resulting in a more focused dataset that can be used without significant post-collection filtering (17,55). In any case, the specific query parameters utilized to gather social media data are a critical component to understanding the implications behind research reliant on this source of data as even small changes can significantly change the content of the sample, and therefore the results of any subsequent analysis.

Query parameters

Querying location-based data always requires a location parameter, which often places significant limitations on the amount of data available for download. Not only does a location filter exclude posts from outside the prescribed area, but it also eliminates the possibility of downloading the significant majority of total posts due to privacy settings. For example, the location of tweet creation is not published alongside a tweet unless the user has explicitly enabled the geotag feature, and as a result estimates put the proportion of geotagged tweets between 0.85% and 2% of total tweets (5,56). However, while a geotag (which includes the precise latitude/longitude location from which a tweet was posted) is the most precise and therefore preferred geolocational identifier, location can also be queried via locations mentioned in the body of a tweet, or the location associated with profile of the user that posted the tweet (56), both of which are a less reliable, yet more readily available measure of location. Geolocational identifiers can be queried by supplying a specific location (place name or coordinate pair) parameter in addition to a search distance radius, or alternatively, four coordinate pairs denoting a bounding box within which all posts are collected.

Content is most often filtered using specific keyword parameters, which when used alongside locational and temporal parameters can greatly narrow down a sample of tweets to those pertaining to a common subject area in a common time and place. Studies utilizing the content of twitter feeds to model influenza trends combined the keyword parameter 'influenza' (including synonyms such as flu, etc.) with a temporal parameter limiting the search to one full flu year, and a locational parameter limiting results to those within the target country or to those with a geotag (23–26). In the case of Twitter, hashtags can also be used to effectively gather content related to a specific topic as hashtags signify a commonly accepted context or topic present in the tweet. For example, using #earthquake combined with a short time period parameter following a major earthquake yielded a sample of tweets whose spatiotemporal distribution (i.e. when and where certain the tweets were created) mirrored that of the earthquake epicenter (17).

Rate limits

Location-based social media data can be gathered either via a stream or downloaded from an archive. Studies utilizing Flickr often used the archive approach, gathering all posts within a selected area (usually a bounding box) and timeframe (35,38). In contrast, studies using Twitter data often utilized the Twitter streaming API, which allows direct download of tweets as they are being created, usually within a specific area of interest. The stream provides either a general or filtered sample of all tweets in real-time, and has been used by researchers to gather large tweet datasets over several weeks or months that can then be filtered after collection (22,53,57). The advantage of the Twitter streaming API as opposed to an archive download is that Twitter imposes a rate limit of 1% on all API requests for standard users, meaning that no more than 1% of tweets can be returned without service interruption, or paid premium API access privileges (53). The 1% rate limit, combined with the low overall prevalence of geotagged tweets means that collecting a broad sample over time with few criteria other than a locational filter results in a more robust dataset than can be gathered with a single API archive request.

RQ2: What were the methods used to analyze LBSMD?

Categories of analysis

Methods of data analysis were reflective of the attributes of the collected data; the different types of analysis are identified here accordingly. For the purposes of LBSMD, a dataset can be examined by the time and location at which each post was created (spatio-temporal (ST) metadata), the text or image content of each post, and/or the user identifier of the post's creator. These three attributes and the methods of analysis that stem from them are non-mutually exclusive, and therefore any combination of them can and were utilized alongside one another. For example, the seminal work first exploring analysis of geo-Twitter data to detect events in real-time utilized both the spatial and semantic content of tweets (58). For the purposes of this section, methods of analysis will be examined via the lens of each LBSMD attribute, with the understanding that the observed studies commonly utilized analytic methods from each category in unison.

Spatio-temporal analysis

The ST attributes of LBSMD are used to assign a geo-temporal dimension to the semantic content in the dataset so that it can be analyzed for patterns and compared to other relevant spatial features. Thus, analysis of this LBSMD component is central to the observed studies as it grounds the online, user-generated information in real space-time so that it can be meaningfully compared to other real-world data. In many cases, ST analysis of LBSMD was successfully used to detect the time and place of specific events by observing the locations of anomalous quantities of related posts (58–64). In others, LBSMD-derived spatial indices of specific phenomena were produced and statistically compared with ground-truth data to mathematically assess possible ST correlations between the LBSMD-derived variables and confirmatory evidence (28,35,37,39,42,44,55,65–75). In any case, the ST metadata within LBSM-datasets enables ST analysis of social media content by giving the user-generated, semantic information points of reference in real space-time. These reference points can then be analyzed to produce measures of human mobility and activity (76–90), emergency events (63,91–93), and even prevalence of infectious illness (23–26,64,94–101).

An illustrative example of comparative analysis is demonstrated in studies seeking to utilize LBSMD to detect or model specific events or areas, such as traffic accidents or natural disaster footprints. The basic premise of comparatively analyzing LBSMD is that if a given subset of tweets, for example, can be analyzed and manipulated to produce spatiotemporal patterns that match or reflect the patterns of verifiable ground-truth data on the same topic, social media data can be relied upon to act as a sensor for similar events in the future (17,58,102). In 2013, Crooks et al. adapted the framework established by Sakaki et al. (2010) and measured the spatiotemporal creation patterns of tweets containing '#Earthquake', comparing them to the time and epicenter location of the 2011 Mineral, VA earthquake (17,58). By plotting the data spatially and temporally, researchers are able to visually demonstrate correlations between LBSMD-derived measures of an event and actual confirmatory indicators of that same event, thus establishing the veracity of LBSMD as an alternate data resource (22,103,104).

A variant of visual comparative analysis is found in papers that use regression techniques. Regression, both geographically and temporally weighted, was commonly applied to compare LBSMD to other datasets, including other LBSM-datasets. This technique was used to create numeric and sometimes mappable indices of correlation between two variables to quantitatively examine the relationship between the spatio-temporal dimensions of two related datasets. For example, in 2018 Wu et al. utilized both check-in data and POI data to map how land use type, time, and location affected the check-in derived measure of urban vibrancy, employing a regression equation to mathematically measure the heterogeneity of vibrance against these variables (44). Similar approaches were taken by a variety of studies utilizing regression to measure the correlation between their collected subset of LBSMD and a validation dataset of choice to produce indices of spatio-temporal correlation, heterogeneity, or perceived value (28,35–37,42,66,67,69,105).

Human activity and mobility were also common subjects of ST analysis using LBSMD, as space-time variability in user posts can be strongly indicative of real-world patterns in human behaviour (77). Some research analyzed mobility through general ST variations in online posting activity, utilizing the quantitative shifts in the general number of posts being made throughout the day as inferential data suggesting equivalent shifts in the number of people in those areas (44). Other researchers utilized the user identifiers associated with posts to derive aggregate ST trajectories, which show the typical mobility patterns of specific populations through an area (77,80,106,107). The former method was typically used to measure ST fluctuations in municipal population densities or to generate occupancy curves for establishments (78,79,108,109), while the latter method of analyzing trajectories was used to model aggregate movements of specific user populations of interest, such as those potentially infected with influenza (23,25,26,64,94,96). Activity analysis was performed to detect ST variability in hot-spots of specific user actions, and was frequently used to show ST patterns in tourist activity behaviours (31,33,110–115).

The reviewed ST analyses were facilitated by a variety of computational techniques used to detect spatial clusters, measure ST correlation between datasets, and analyze patterns in social networks. Spatial clustering algorithms such as k-means and DBSCAN were used to identify variations in ST points derived from LBSMD to develop spatial trajectories and detect anomalous clusters of posting activity (72,74,77,86,110,116–118), which was particularly essential for mobility analyses. Kernel density estimation (KDE) was also utilized in detecting ST hotspot activity, as well as for transforming discrete ST points derived from LBSMD into continuous spatial surfaces (31,42,44,109,111,119–121). Geo-temporally weighted regression was the preferred method for measuring the ST relationships between comparable datasets and assessing spatial correlations (28,35–37,42,66,67,69,105). Finally, social network analysis was used to examine the ST variability of interactions between users online and develop models of social connection across digital space (55,122,123).

Content analysis

Where comparative spatio-temporal analysis addresses the patterns of post creation with respect to other spatial features, content analysis utilizes the actual text, image, or keyword identifiers associated with a post to enhance or derived further information from the dataset . As discussed, these types of analysis are non-mutually exclusive, meaning it was not uncommon for researchers to utilize both spatiotemporal and content analyses within the same project. In most cases, content analysis was used alongside spatiotemporal techniques to strengthen or expand upon the inferences made based upon the information contained in each post. Content analysis differs from the other analytic categories in that the methods utilized are concerned only with the text or image content of the post, and are thus distinct from methods used to analyze spatio-temporality of posts or users.

Beyond simple keyword matching used at the collection stage, studies utilizing LBSMD have applied more advanced content analysis techniques, such as using text-mining and text analysis to improve the reliability of the dataset with respect to relevant context. For example, after collecting flu-related tweets using an API-based keyword search for 'Influenza' or 'flu', Allen et al. (2016) applied machine learning methods to further analyze the text content of tweets in order to remove those tweets that, although containing the word 'flu', do not indicate that the user is actually sick with the flu (ex: *I'm getting the flu shot today* (invalid) vs. *I gotta get over this flu!* (Valid)) (23). In this way, content analysis via filtering produces a more robust dataset, removing false positives gathered during the collection process and enabling greater confidence in the conclusions drawn from subsequent spatiotemporal analysis by ensuring only relevant posts are included. The basis for this type of content filtering is found in machine learning techniques, which enables algorithms to 'learn' how to distinguish between cases based on their content by training them on example datasets (26,101,124).

Machine learning approaches can also be used to detect the intended underlying sentiment in a text corpus through training on similar datasets with human-verified answer keys. Sentiment analysis, where posts are assigned a 'sentiment' based on the interpreted meaning of their contents, was commonly used in studies that sought to model human thoughts and emotions present in social media datasets (53,71,125–129). The sentiment of a post refers to the meaning or feeling that the user was trying to convey with their words, which can be determined using either manual (54,127) or the aforementioned algorithmic, machine learning-based approaches (130). For example, Seltzer et al. (2017) classified Zika virus-related Instagram post content by hand to assign a sentiment of either humour, fear, positive, and negative to investigate characteristics of public discourse on the topic (127), while in 2016 Jiang et al. analyzed Weibo posts to gauge citizen's feelings about a large public infrastructure project by using a machine learning-based approach (125). Earlier research analysed sentiment through manual classification of posts, exemplified in 2015 with Yang & Mu's identification of depressed Twitter users through researcher-classified depressive tweets (54).

Finally, post content can be analyzed to derive a representation of the main topics or meanings presented in the text, or in some cases images or videos. In the case of text data, 'topic modelling' is a machine-learning, more specifically a natural language processing-based analysis technique that generates a statistical model of the abstract topics present in a corpus of text (53,131–136). This type of model, based in Latent Dirichlet Allocation, assigns a number of related words to each generated topic area, categorizing words based on their interchangeability observed in the body of text so as to reduce a large body of text to a researcher-assigned number of topics (53,137). Alternatively, when the desired information resides in images rather than text, machine learning-based image recognition techniques can be applied to identify posts based on specific objects present in the picture. In 2019, Di Minin et al. trained a deep learning algorithm with manually verified data to produce a model capable of identifying posts on Instagram that contained images of illegal wildlife trophies (ivory, etc.) (138). Videos from social media have also been analyzed to derive summative information, similar to text and video, however most of this work has been done manually by researchers viewing and recording notes for videos, such as in 2018 where Basch et. al used YouTube as a source for information on a specific illness (50).

RQ3: What were the methods used to visualize LBSMD and communicate findings?

Data collection and spatial analysis have ostensibly more impact on the overall output of the research, however the resulting visualizations have the most significant impact on reader communication, and therefore dictate the uptake of a given idea presented by a study. After the data has been analyzed, it needs to be transformed and visualized into a format that clearly demonstrates the utility of the LBSM-dataset in emulating a chosen phenomenon. Where data analysis consists of methods used to glean information and meaning from the data, the methods of data transformation and visualization are the vehicle that allows the information to be communicated and understood in the context of the research objective. Studies in the review frequently utilized cartographic maps, graphs, and tables to communicate research findings visually. The visual outputs contain the main thrust of the results and research findings in general, as they are the direct product of the collected LBSMD upon which the study is based. As such, these methods and the outputs they produce are of significant importance when considering the utilization of LBSMD, as they are the main mechanisms by which research findings can be communicated and evaluated in print. The following will describe the prominent themes in data transformation and visualization found in the included studies; instead of presenting an exhaustive list of every methodological variant found in each research paper, select examples will be used to illustrate what were found to be the main ways of achieving LBSMD-derived visual output.

Basic visual outputs

A combination of tabular, graphic, and cartographic methods were used in almost all cases to communicate key evidence and information derived from collected LBSMD. Similar to techniques of analysis, categories of transformations and visualizations were non-mutually exclusive, with studies employing multiple methods to communicate findings. Tables were commonly used to display and relate spatio-temporal attributes with other variables such as correlation coefficients (23) and other summative data (70,73–75,75,90,123). Time-series were often shown graphically in addition to tabularly: in the most clear-cut cases, studies showed the temporal incidence of two variables (usually one derived from the social media dataset, another from an alternative confirmatory source) to demonstrate how closely the social media data emulated the patterns observed in the closest authoritative account of the actual phenomenon (17,24,139). For example, in 2017 Wang et al. utilized messages collected from the platform Weibo to infer information about air quality, and then plotted the inferred values alongside weather station data in-series to show how closely the inferences based on LBSMD matched to the reality of air quality on the ground (139). Although graphs and tables were utilized heavily by studies in the review and served as a main type of information visualization, their discussion will end here given their simplicity and widespread usage across all fields of quantitative study.

Point mapping

Ultimately, cartographic representation (mapping) was the most significant vehicle used to communicate spatial findings due to its visual nature and high information density. Figure 4 contains an example of each type of cartographic representation used in publications to communicate data derived from social media. In some cases, the spatial incidence of the collected data alone was sufficient to demonstrate the target phenomenon without any cartographic techniques beyond plotting the location of post creation with respect to a specific time and place. This results in simple point-mapping, as shown in Figure 3a, where each point represents the location of an individual tweet. In 2013, Crooks et al. were able to show how the spatiotemporal creation patterns of tweets related to an earthquake mirrored the incidence of the earthquake itself by plotting the locations of the tweets at various time intervals shortly after the event (17). This produced cartographic maps clearly displaying the progression of the earthquake's impact as it rippled outwards from the epicentre using nothing but the time and location of collected tweets superimposed over the continental US, and then additionally plotted against alternative crowdsourced means of earthquake impact measurement to clearly demonstrate the correlation between the datasets (17). Aside from this specific example, points were often used to portray the locations of the LBSM-dataset, situating the sample inside the study area (5,35,55,72,74,140).

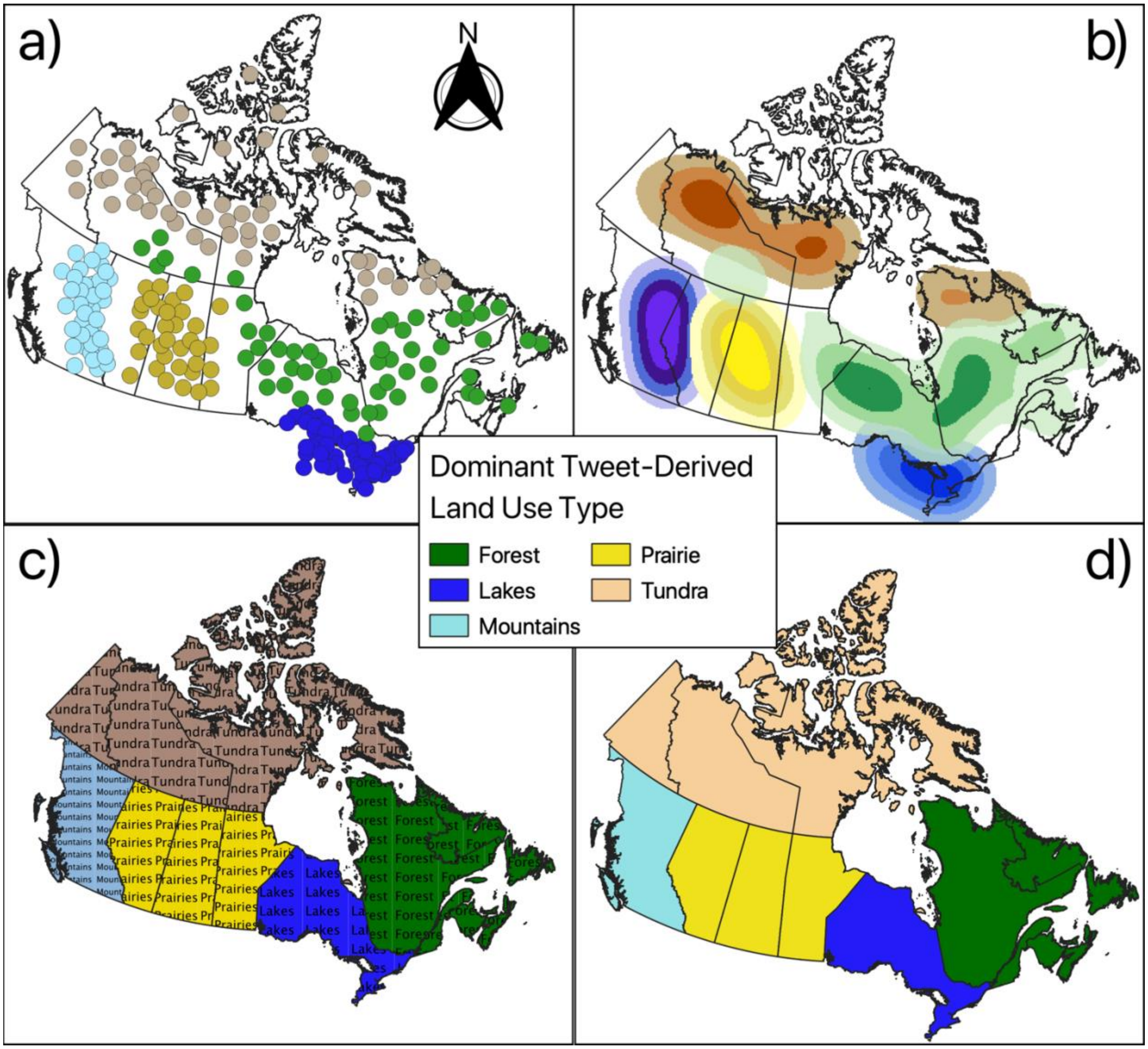

Fig. 3. Examples of transforming pseudo Twitter data into a variety of visualizations.

a) Shows an example of basic locational plotting, where each point represents the location and classified land use type contained in a single tweet. b) Shows these same points spatially interpolated into a raster surface representing location and content. c) Shows a qualitative visualization where the content is reflected via text, and the location via Provincial/Territorial divisions. d) Shows a thematic choropleth map using the Provinces and Territories of Canada as zonal units, where color reflects the dominant tweet-derived measure of land use type in each zone.

Raster surfaces

Point mapping is an efficient way to visualize data because most LBSMD is point-based, as is most spatial data more generally, however, there are many cases where this technique does not provide sufficient spatial information due to the nature of discrete point-based locations. To overcome the limitations of point-based data, researchers used methods of spatial interpolation, which allows for the prediction of cell values based on a limited number of sample data points (141). There are a wide variety of spatial interpolation methods, each varying based on different mathematical approaches or assumptions about the data, and as such the specific methods of spatial interpolation utilized by studies in the review were largely heterogenous. Mitas & Mitasova (1999) outlines many of these variants and can serve as a technical manual for understanding and applying the specifics of spatial interpolation techniques utilized by studies included here. The main output of spatial interpolation is a raster surface displaying the coverage area and point density derived from the initial point sample (69,70,72,74). Shown in Figure 3b, an interpolation method is used on the points shown in a) to create a continuous surface whose area and color intensity are estimated based on the locations and proximity of each point relative to one another. Raster surfaces were commonly generated using kernel density estimations, and were used to visualize social media-user density with respect to public park use (41,69), spheres of influence in adjacent cities (107), perceived value of POIs (42), and a variety of other different patterns of human activity and behaviour (131,142–144).

Thematic maps

Thematic maps, as shown in Fig.3 d), use established boundaries or zones to eliminate the need for interpolating point area and density by classifying each zone according to the number of observations (points) within each zonal polygon. In the case of Figure 4d, the Provinces and Territories of Canada are used as these zonal polygons, each being categorically color coded according to the most numerous type of tweet contained within. This strategy is commonly used when predefined boundaries, such as those at the level of municipal wards or state counties, are available and serve as meaningful delineations within which the data can be shown (35,69,71,73,74). In 2015, Yang & Mu utilized this mapping strategy to investigate the incidence of depression amongst Twitter users in the New York metropolitan area at the county level by color-coding counties according to the number of Twitter users showing signs of depression within each zone (54). Similarly, Ghosh & Guha (2013) classified obesity-related tweets at the county level to show the relative rates of discussion on the topic occurring across the U.S (131). Tu et al. (2020) used area-based choropleth maps extensively in their exploration of LBSMD-derived measures of urban vibrancy in Shenzhen, China, using traffic analysis zones as the areal boundaries for the analysis (73).

Qualitative content diagrams

The types of visualization discussed thus far are effective at displaying spatio-temporal, or otherwise quantitative data, however many of the methods of analysis discussed above (and therefore many papers in the review) produced qualitative analytic outputs that don't map onto quantitative schemes of transformation. Therefore, many studies in the review utilized cartographic methods that produced figures depicting the content or topics contained within a corpus of LBSMD in a spatial manner. To achieve this, studies used LDA-based content analysis to generate topics and topic probability which were then visualized in 'code clouds' or 'content clouds' which display each topic visually, often showing higher probability or more numerous topics with larger text (53,120,145–148). Then, to add a spatial dimension to the code cloud topic models, the content of the clouds would be inputted into a GIS environment where the topic content can be geo-locationally visualized according to the locations aggregated within each prominent topic (53,146,147). A simple example of this can be seen in Figure 3c, where the dominant land-use types are visualized graphically on the map itself at the level of the provinces, rather than being symbolically described using a color-coded legend. This is obviously an oversimplified example; however, the main idea is that the topics or words themselves are being represented visually on the map as opposed to being shown with symbols depicting the quantitative classification derived from the such text.

At the most basic level, Jung (2015) generated such topic models from geotagged tweets, and then showed the area of the city from within which the tweets were located (and therefore the location from which the topics were generated) using a bounding box super-imposed over a cluster of tweet locations (146). Martin & Schuurman (2017) took this approach one step further by generating topic models from geo-tweets that were grouped using pre-determined vector boundaries, and then spatially displaying the topics generated from each group within these pre-determined boundaries, resulting in a cartographic representation of Twitter topics at the neighbourhood level (53). Dunkel (2015) used a different approach to visualize the relative abundance of specific tags on Flickr by mapping the text of each tag at the location at which it was tagged, and increasing the size of the text based on the frequency of that term in that location (147).

Transmission diagrams

When human mobility and matters of transmission were studied (be it transmission of pathogens or ideas), the spatio-temporality of point data pertaining to specific users was transformed to create polyline trajectories between points that intertwined to create networks of transmission (fig. 4). These network diagrams display the space-time trajectories generated by investigating post patterns of specific users, inferring user movements when the same user creates posts at differing times and locations. A prime example of this is shown in an influenza surveillance study variant that aimed to model transmission of the flu by generating amalgamations of infected user trajectories to show how transmission of the virus can be modelled using LBSMD (25). Similar techniques were used to visualize the spread of news across the globe from the time and place of a specific event, such as the 2015 Paris attacks (55). In 2016, Williamson & Ruming transformed the results of social network analysis in the context of Twitter data to produce network model visualizations that displayed spatial relationships between users as a function of their tweeting and retweeting behaviours, with each user acting as a node connected by lines, and highly prominent users being displayed using labels and larger sized nodes (149). In the same vein, Wang et al. (2018) used similar types of transformation depicting lines connecting various nodes to visualize aggregate social connections, deriving visualizations of inter-urban connection pathways found in online social interactions in China from Weibo data (150).

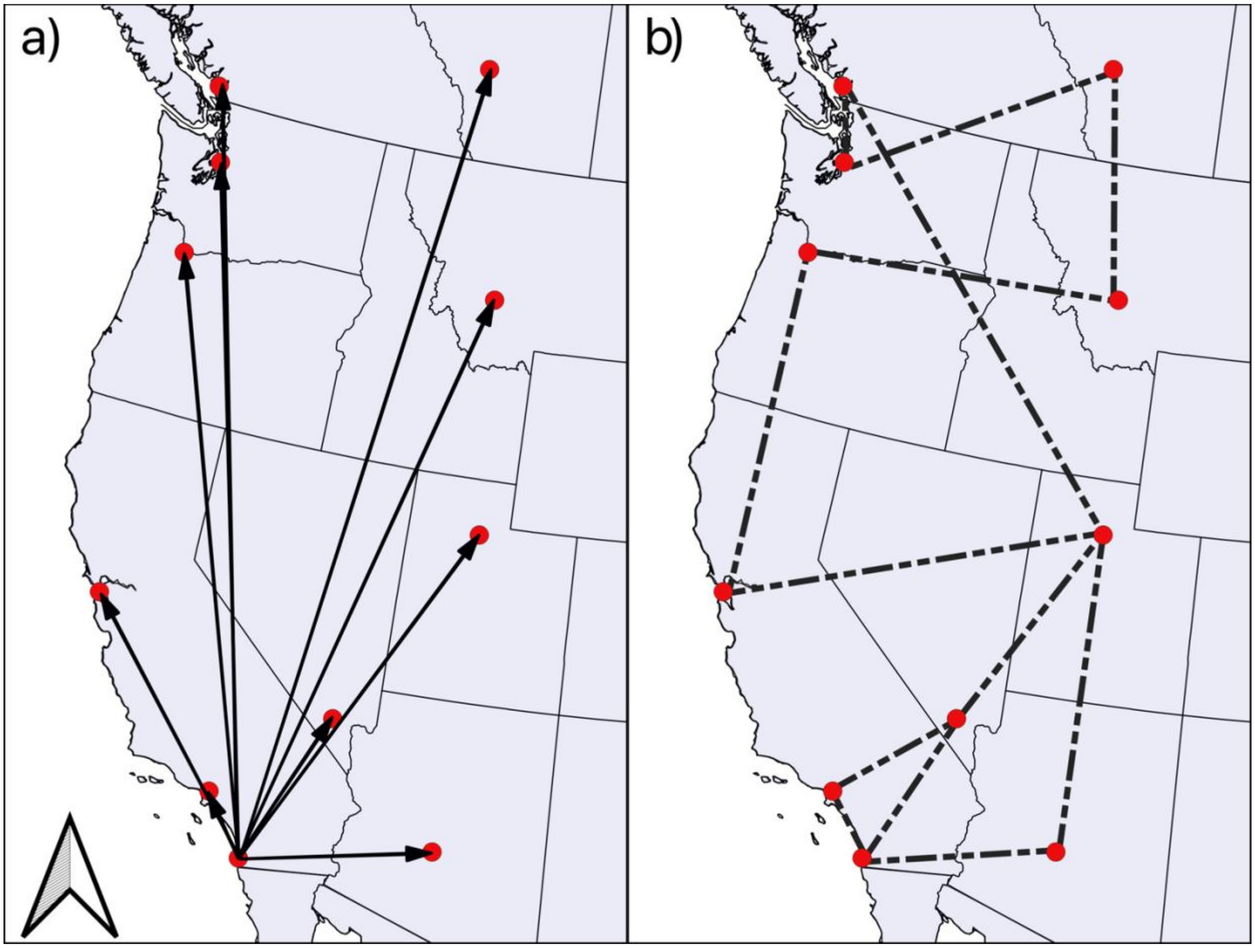

Fig. 4. Two types of LBSMD-generated transmission models.
a) Shows transmission of information from an epicentre to surrounding nodes, in this case depicting the unidirectional broadcast of a tweet from Southern California to other cities in which retweets of that same tweet were located. b) Shows the flow of information (or users) between nodes, in this case displaying the dominant bi-directional paths of transmission between Western Cities.

Web mapping

The final type of data visualization and transformation is defined not necessarily by the content of the figures, but rather the format. Where the majority of studies used static images of maps to display their findings, some researchers operationalized their workflows into online platforms which allows for real-time reader access to interactive cartographic displays of the social media data in question. Researchers include links to webpages that contain the cloud-based workflows, which are able to run autonomously and update in real-time by continuously sampling the data source (i.e. the Twitter streaming API) for posts matching the query parameters. Web maps were commonly utilized in research contexts where real-time data access was necessary or paramount to the research objective, such as in 2017 when Huang et al. developed an online graphical user interface (GUI) allowing real-time tracking and analysis of disaster events derived from Wikipedia entries and social media feeds (151). In 2014, Padmanabhan et al. presented *FluMapper*, which operationalized the trajectory mapping associated with tweet-derived flu transmission into a cyberGIS application allowing for reader interactivity and map dynamism with respect to different locations and time periods (allowing users to pan the map and use the slider tool to investigate the trends in a particular time and place) (25). Several other studies in the review included web mapping as a visual research output for a wide variety of topics, from a platform predicting earthquakes by analyzing and mapping the incidence of anomalous animal behaviours in real-time (19), to web maps designed to improve resource dispatch during a disaster by analyzing Twitter data (152), and more (153–155).

Discussion

Data collection protocols showed a strong preference for utilization of data from Twitter, which was found to be the platform most widely utilized by researchers (5,6,8,17,21,23,26,54,58–60,71,100,126,128,129,156). This is likely due to Twitter's free API, which allows for seamless access to data with minimal investment, and to the massive volume of geolocational tweets generated every day. This observation was commensurate to that of Stock 2018, where Twitter and Flickr were found to be the data sources for about three-quarters of the studies found by the systematic review protocol (10). Steiger et al.'s 2015 review also clearly reflects this, as their findings indicated a wealth of literature on the topic of uses and applications for LBSMD specifically from Twitter (3). Twitter will remain a popular data source for this type of research, however the inherent limitations of data sourced from this platform like user selection bias and data availability (8) should always be considered, especially with the observed over-reliance on this social network in particular.

This review found significant reliance on keyword-based search and filter strategies for data collection in the observed studies, although the more recent literature has shown an emergence of automated methods such as natural language processing (NLP) being put to use for this task. As most social media API's operate using keyword-based archiving and search strategies, this reliance is understandable. However, as this review demonstrates, researchers have started implementing post-collection processing techniques to increase the robustness of samples and eliminate false-positives to a degree beyond what is possible when reliant on API-based collection. This finding is in contrast to Steiger et al. (2015), which found a similar trend in data collection via keyword-based search strategies, and noted the weaknesses of this approach, but did not find evidence of algorithmic alternatives (3). NLP techniques allow for the algorithmic classification of large bodies of text, and are able to capture user meaning and intent far more efficiently than simple keyword-based filtering (26,53,156). Thus, as NLP and the larger field of machine learning (ML) have made considerable advancements and have been increasingly propagated to other disciplines over the last 5 years, the move by spatial researchers to include these techniques in LBSMD collection and processing protocols is understandable.

Algorithmic strategies are also playing an increasingly prevalent role in social media data analysis, and are allowing researchers to garner more and more information from larger and larger datasets (157). However, as Reich (2015) put it, big datasets do not carry with them the answers to interesting questions just by virtue of their size (158). Further, given the primacy of algorithmic solutions to big data analysis, some have called for the geographies emanating from such work to be named 'algorithm-driven geographies' as opposed to data-driven geographies (159). Thus, while algorithms have enabled deeper spatial inquiry at larger scales, they have also introduced inherent algorithmic uncertainty at every level which must be addressed by researchers in practice whenever engaging in big data enabled geographic knowledge production (159). For example, as Tu et al. (2020) demonstrated, the same phenomenon can be measured via different metrics and produce divergent results as a result of slight alterations to the data collection and generation procedures (73).

This review discussed methods of data visualization at the level of data input and visual output, and found that many of the visualization strategies centered around generating areas from discrete points. As Martin & Schuurman (2020) reminds us, data visualization is not the direct reflection of the world it seeks to represent, but rather, similar to an abstraction, it obscures certain features in order to highlight others (160). When utilizing kernel density estimations, cluster analysis, or data aggregation to a specific areal unit, the original geo-data contained within the LBSMD is being obfuscated and transformed in many ways. By flattening data to a single dimension and using it to extrapolate spatial coverages of the contained information beyond that of the original discrete points, visualization techniques, especially when automated, can lead to spatial inconsistencies.

Contributions / strengths

This review offers an extensive overview of the scope of research surrounding the interface between GIScience and social media data. The systematic search protocol combined with a specific focus on the applicable methods involved in spatial manipulation of social media data ensures readers have an understanding of the broad methods involved in this field of work. While the review does not explore any one method or element of analysis in exhaustive detail, the wealth of cited works in each section provides a strong basis of literature for any researchers interested in pursuing the finer details of a particular element discussed here.

A number of other reviews have covered content areas similar to the data considered in the present review, however the present review remains distinct in its geospatial lens and focus on operationalizing the methods necessary for GIS-social media interfacing and research production(11–14,161). For example, in 2018 Stock presented a systematic review of methods used for the extraction of location from social media datasets, which describes in great detail all of the methods presented in the related literature (10). The 2018 review was exhaustive in its approach, and while it shares many similarities to the present review's section on data collection in terms of statistics on platform popularity and the research objectives of studies utilizing social media data, it does not address methods of spatial analysis or visualization necessary for research outputs. Thus, while the present review does not cover the specifics in methods of location mining from social media to the same degree as the aforementioned review, it does cover the broader scope of this area in combination with similar coverage of subsequent required methods in the workflows of studies exploiting social media data for research purposes, namely spatial analysis and visualization.

Limitations

This review was primarily limited by incongruence between the volume of research gathered by the search protocol and the information limits on a single review. With over 200 papers included via the search and filtering process, an in-depth exploration of any specific method or paper was not possible. Therefore, this review offers a cursory examination of the patterns in trends observed in the included articles, with the expectation that readers access the cited works for greater information and detail on the specific research objectives and methods discussed.

Crowdfunding platforms as social media

Chapter 2 presents a review of methods used for GIS analysis of location-based social media data, and precedes chapter 3, which mobilizes these methods in a case study. Indeed, while chapter 3 presents an example using methods to collect, analyze, and visualize user-generated online spatial content, the particular source of this data may not be considered by some to be 'traditional' social media. The source, an online crowdfunding platform called gofundme.com, is by no means a typical social media platform, however, chapter 3 will treat it as such in its application of the previously reviewed methods.

Crowdfunding platforms are online spaces where people can post campaigns to raise money for a cause. A prime example of such a platform, gofundme.com, fits this definition, but, as established does not fit the traditional definition of social media. However, these platforms operate in a similar manner to social media platforms, and the data gathered from them have similar qualities. For example, every campaign posted contains a location from which the campaign originated, text, image and video content pertaining to the topic of the campaign, and campaigner (user) metadata. Thus, while gofundme.com is not a traditional social media, it can function as one for the purposes of gathering analyzing, and visualizing spatial user data. To this end, the following chapter wll mobilize the methods reviewed in chapter 2, using location-based data from gofundme.com as the subject of spatio-temporal and content analysis and developing visualizations from the result.

Chapter 3.

Spatial and temporal patterns in Canadian COVID-19 crowdfunding campaigns[2]

Abstract

Online charitable crowdfunding has become an increasingly prevalent way for Canadians to deal with costs that they would otherwise not be able to shoulder on their own. With the onset of COVID-19 and related lockdown measures, there is evidence of a surge in crowdfunding use relating to the pandemic. This study gathered, classified, and analysed Canadian crowdfunding campaigns created in response to COVID-19 from GoFundMe.com, a popular crowdfunding platform. Spatio-temporal analysis of classified campaigns allowed for observation of emergent trends in the distribution of pandemic-related need incidence and financial support throughout the pandemic. Campaigns raising money on behalf of established charities were the most common in the sample, and accounted for the greatest portion of funding raised, while campaigns for businesses made up a small proportion. Dense metropolitan areas accounted for the vast majority of campaign locations, and total sample funding was disproportionately raised by campaigners in Ontario and British Columbia.

Introduction

In times of crisis, charitable crowdfunding serves as a popular method to connect those willing to give with those in need. From the half-billion dollars raised by the American Red Cross on behalf those impacted by the 2010 Haiti earthquake (1), to concerts raising money to fight SARS (2) , and even to donations made at grocery store checkouts, crowdfunding initiatives funded by everyday people have been known to raise generous sums when the need arises. Even though these crowdfunding initiatives manifest in different forms, their unifying aspect is that they are charitable in nature, and on behalf of those in need as opposed to equity-based crowdfunding which raises money for the launch of a business venture. In recent years charitable crowdfunding has expanded into the online space, allowing for individual people or groups to host 'campaigns' on behalf of their particular cause in hopes of raising money directly

---

[2] Authors: Matthew K. McKitrick, Nadine Schuurman, Valorie A. Crooks, Jeremy Snyder.

from people in their personal network and beyond. GoFundMe, the largest of these platforms, has raised over $9 billion worldwide for campaigns on their platform since its launch in 2010, and in 2018 mediated roughly 90% of U.S. and 80% of global charitable crowdfunding dollars (3,4). GoFundMe and similar platforms host campaigns and allow campaigners to include text, images, videos, and periodic updates on the progress of their campaigns. This new medium of charitable crowdfunding has created a space for those with nowhere else to turn to solicit for financial assistance, and has resulted in the fulfillment of many campaign funding goals for worthy causes.

In addition to being more accessible to most people, the popularity of online crowdfunding platforms has also allowed researchers greater access to data on the topic, and thus more opportunity to analyze and investigate the manifestations of online charitable crowdfunding at large. Given the wide variety of platforms and absence of sanctioned access to this data, a diverse set of methods has been utilized in the literature to gather, analyze, and describe the various dimensions of online crowdfunding. Crowdfunding at large is investigated through the analysis of campaigns created by people seeking financial assistance associated with a need demonstrated by the content of the campaign, which are often manually searched and catalogued by researchers to develop a cohort pertaining to a specific topic from which conclusions can be drawn (5–7). In contrast to work investigating the qualitative trends in a small number of highly specific campaigns, other research has utilized automated collection mechanisms like web crawlers to gather large cohorts of campaigns from which summative statistics can be generated (8–11). Because campaigns include a mixture of both quantitative and qualitative information, research often incorporated a blend of methods and results which describe both these dimensions.

In light of the expanding body of research on the topic, many researchers have been critical of the ethical and societal implications of this new mode of donation based crowdfunding, and have increasingly found that it exacerbates inequality by awarding crowdfunding dollars not by relative need, but by ability to appeal to an audience through mastery of digital media and media literacy (12,13). Rather than alleviating societal inequities by allowing anyone to create a campaign, online donation-based crowdfunding has come under scrutiny for potentially perpetuating imbalances, replicating existing inequities in race, gender, and socio-economic status by favouring those with the means to create successful campaigns rather than those most severely in need (11,14). Geographic-oriented research in this domain similarly found evidence of spatial inequities, where communities in urban areas vastly outperformed their rural counterparts, bringing in drastically higher volumes of funding and support through crowdfunding platforms even when accounting for relative population densities (15). Nevertheless, GoFundMe is a clear leader in the charitable crowdfunding market, and nothing has demonstrated this more clearly than the incredible volume of campaign creation and charitable donations in response to the COVID-19 pandemic.

The COVID-19 pandemic has been another crisis where individual people and groups have turned to crowdfunding for help and to aid others. Between March 1 and August 31, 2020, over $625 million was raised globally across 150,000 campaigns created to support causes related to the COVID-19 pandemic; in the US, approximately 60% of pandemic-related campaigns were created to support small businesses and those dealing with unemployment (16). These figures show not only the impacts of the current crisis, the immense need it has generated, and the generosity and situational awareness of those still able to donate, but also the prevalence of GoFundMe and online charitable crowdfunding in general as a significant means with which society addresses economic hardship.

Canada has been no exception to the economic hardship caused by the COVID-19 pandemic. From February to April 2020, about 5.5 million Canadian workers had their employment situation negatively affected by the economic shutdowns, and in the proceeding period from May to September 2020, over 8.75 million unique applications were made to the Canada Emergency Response Benefit (CERB), equating to over $76 billion in funds being dispersed to Canadians in just 6 months (17,18). As a result, many Canadians have turned to crowdfunding to address the needs created by the COVID-19 pandemic and efforts to mitigate its spread.

In this article we provide a snapshot in time of COVID-19 related Canadian crowdfunding campaigns during the first 6 months of the pandemic. Specifically, the spatio-temporal characteristics of these campaigns are analyzed and visualized to show the patterns of campaign creation and success throughout Canada between January and June 2020. The aggregation and analysis of this data creates an opportunity to understand not only the needs that emerged in Canada as a result of the pandemic, but also where and when they arose relative to significant events in the pandemic timeline. This analysis is original in its spatio-temporal focus, using a map and several figures to visualize where in Canada COVID-19-related campaigns were created, when campaign creation for specific pandemic-related needs were most prevalent, and how these factors relate to campaign fundraising success.

Methods

Since campaign data are not publicly available for bulk download, it must be gathered using algorithms to 'scrape' data directly from the website if large volumes of data (too large to manually copy) are desired. This method of data collection has previously been used to target and gather campaigns raising money for a specific purpose, which are then quantitatively analyzed to examine the qualities of these campaigns. For example, Duynhoven et al. used scraped campaigns to analyze the spatial trends in Canadian crowdfunding campaigns for cancer (15), and more recently scraped campaigns have been used to examine the quantitative trends of crowdfunding for COVID-19 at large (19,20).

The focus of this research is to explore the characteristics of Canadian crowdfunding campaigns created for reasons related to the COVID-19 pandemic. We gathered campaigns from only one platform, GoFundMe, due to its popularity, broad user appeal, and because combining campaigns across platforms can bring out platform-specific effects in the data (13). The exploratory research design uses a broad analytic focus within a specific, robust sample, in search of emergent properties rather than the answer to a specific research question. The value of interpretation and discussion of campaign characteristics lies not in sample completeness, but in sample size and content. The scraped campaigns, as is common among scraped web datasets in general, do not form a probabilistic sample, and therefore are not statistically generalizable to the space of crowdfunding at large. Instead, the quantity of gathered campaigns and the specificity of their purpose allows for the examination of trends and dynamics between variables in the sample.

Crowdfunding campaigns were scraped from www.gofundme.com using a Python algorithm. The scraper gathered all of the campaigns resulting from searching "Canada" AND "COVID-19" via the GoFundMe search bar on June 30th, 2020, which amounted to 915 search results in total, each of which were saved to a local database. Four reviewers then examined a random sample of 100 of these campaigns, twice. Based on this examination and subsequent reviewer discussions, 6 content categorizations were developed to describe the underlying funding motivations present in the campaign sample. Shown in Table 1, these content categorizations were decided upon after extensive reviewer discussion surrounding the content of both the title and description of each campaign in their respective sample.

Table 1. **Content categorizations used to describe the funding motivation present in the collected campaigns.**

In support of a(n) ___________ in relation to the COVID-19 pandemic:	Description Funding Requested to...
Canadian Business	Help a small business dealing with lockdown-related closure or general loss of income during the pandemic.
Canadian Charity	Be donated to a formal charity for charitable purposes within For example: Canada Food Bank, Canadian United Way.
International Charity	Be donated to a formal charitable organization for charitable purposes outside of Canada. For example: UNICEF.
Purchase / Manufacture of Personal Protective Equipment (PPE)	Help with the purchase or creation of Personal Protective Equipment, particularly PPE that is thought to shield against COVID-19, such as face masks (N95, Surgical, Cloth), face shields, plexiglass barriers, and other such equipment.
Family Reunification	Help with airfare and other travel costs associated with reuniting family members and pets stranded because of COVID-19 related border closures and flight cancellations.
Personal Need	Directly help people who are experiencing hardship, financial or otherwise, due to the pandemic.

Each campaign in the sample was assigned one of these content descriptors to enable quantitative analysis of the pandemic-related funding motivations in Canadian COVID-19 crowdfunding campaigns. These categorizations were developed by analysis and discussion of random campaigns from the sample.

Campaign categorizations were generated as a result of two collaborative discussion sessions held after each author reviewed a subsample of campaigns. During the discussions, authors promulgated the main themes of their subsamples and drew from extensive previous experience reviewing crowdfunding campaigns to develop a collaborative content taxonomy that included the prominent categories of campaigns in the sample while simultaneously narrowing down the campaign types to 6 specific areas that aptly described the entire sample when assigned appropriately. After the categories were developed, a single reviewer analyzed each campaign, removing those campaigns that were either not created in Canada, were created outside the prescribed study period from January - June 2020, or were not directly motivated by COVID-19 related circumstances. The start of the time period (January 2020) was chosen as a safe estimate for the earliest possible time for which COVID-19 campaigns may have been created, while the period end date (June 2020) was chosen so as to capture the campaigns from the first 6 months of the pandemic to identify emerging crowdfunding responses to COVID-19.

During this process the reviewer also assigned one of the six type descriptors to each campaign in the sample. At this stage, each campaign consisted of seven attributes, each attribute corresponding to a data column: Campaign title, Description, Date of Creation, Funding Dollars Raised, Funding Dollars Requested, Location, and Type, shown in Table 2. The location attributes were gathered in the form of place names, which were subsequently geocoded using QGIS's integrated geocode function to allow for the spatial analysis component. Table 2 serves to exemplify a standard case of a campaign scraped by the Python web scraper, where each row contains a single campaign and each column contains an a campaign attribute derived from the HTML webpage on www.gofundme.com for each individual campaign, which are the subject of the following analysis.

Table 2. Campaign dataset headers and example campaign entry.

Title	Description	Date Created	Funding Raised (CAD$)	Funding Requested	Location	Type
PPE for Frontline Workers	Donations towards the purchase of PPE for frontline workers in Toronto hospitals	05-01-2020	$51,453	$100,000	Toronto, Ontario	PPE

Campaigns relating to the pandemic were scraped from Gofundme.com in the format shown here. The 'Type' column was appended to reflect the type of need portrayed in the title and description. Location and date created values acted as space and time variables, respectively. The funding requested column shows how much money was raised by the campaign from the date it was created until June 15th, 2020.

Results

A total of 915 campaigns were initially scraped from www.gofundme.com, each of which were examined for inclusion criteria; 342 campaigns were removed during this process. The remaining 573 campaigns were analyzed spatially, temporally, monetarily, and according to content, corresponding with the variables of campaigner location, date of campaign creation, funding dollars raised per campaign, and campaign funding, respectively. The main reasons why campaigns were not included in the analysis was due to invalid campaign content (not directly related to COVID-19), invalid location (not in Canada or not in a specific Province or Territory), or because the location was unable to be geocoded. Table 3 shows a breakdown of these variables. The following 7 figures will visualize the trends seen in the scraped campaign across these 4 dimensions.

Table 3. Campaign study variables derived from scraped campaign data

Variable	Description
Campaigner Location	The listed location of the campaign, usually at the municipal level.
Campaign Date of Creation	Date that the campaign was created on.
Campaign Dollars Raised	Funding dollars raised by the campaign (CAD$).
Campaign Type	Underlying type of campaign fundraising need (See Table 1).

First, the campaigns are counted and shown by type (Figure 5). Campaigns created by, or on behalf of charities were the most numerous, accounting for 232 of the 573 total campaigns. The majority of these campaigns were created by people on behalf of established charities to raise money in their particular group, while a comparatively small number of campaigns in this category were created by the charitable organizations themselves. This latter group was clearly identifiable, as the campaigns were created by the verified user accounts of the charity and usually had much larger levels of donations. Campaigns for PPE (118), international charities (90), and general personal need (78) accounted for roughly half of the total distribution, while campaigns for private businesses and family reunification together accounted for about 10%.

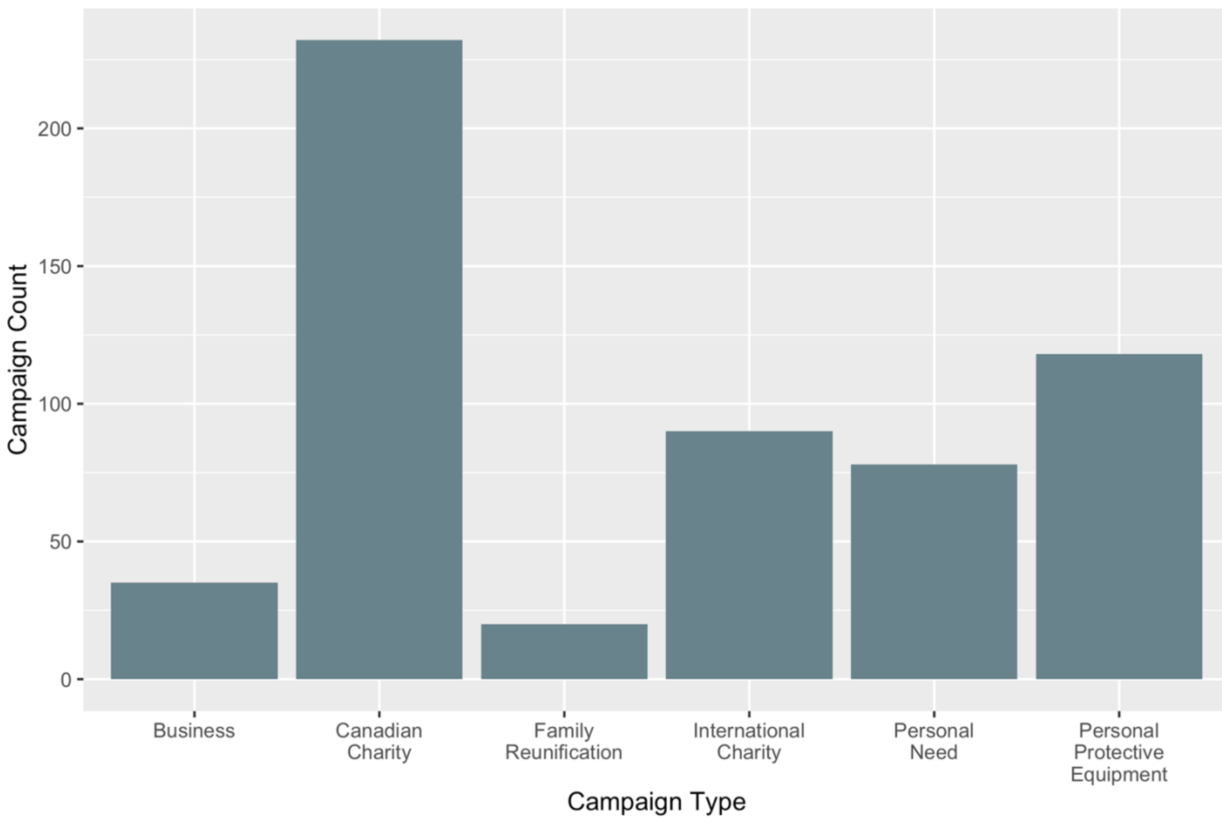

Fig. 5. Campaign count by type
A total of 573 campaigns were categorized by the type of need they addressed. The figure shows the type of need demonstrated by the gathered campaigns relating to COVID-19 from January 1st to June 15th, 2020. For a description of each 'Campaign Type', refer to Table 1.

Next, the count of COVID-19 related campaigns created per day was plotted against the reported number of daily new confirmed COVID-19 cases in Canada, with major related news events superimposed in series (Figure 6). The number of campaigns per day, shown as the grey bars, peaks in early April, as the number of daily new cases surpassed 1000 per day, but tapered off shortly after and remained low into June. The number of new cases peaked almost a month after the apex of campaign creation, and steadily declined into June. The majority of campaigns (68%) were created in the six weeks from mid-March through the end of April, in roughly the same period of time as between the first COVID-19 death in Canada and the initial easing of the lockdown measures.

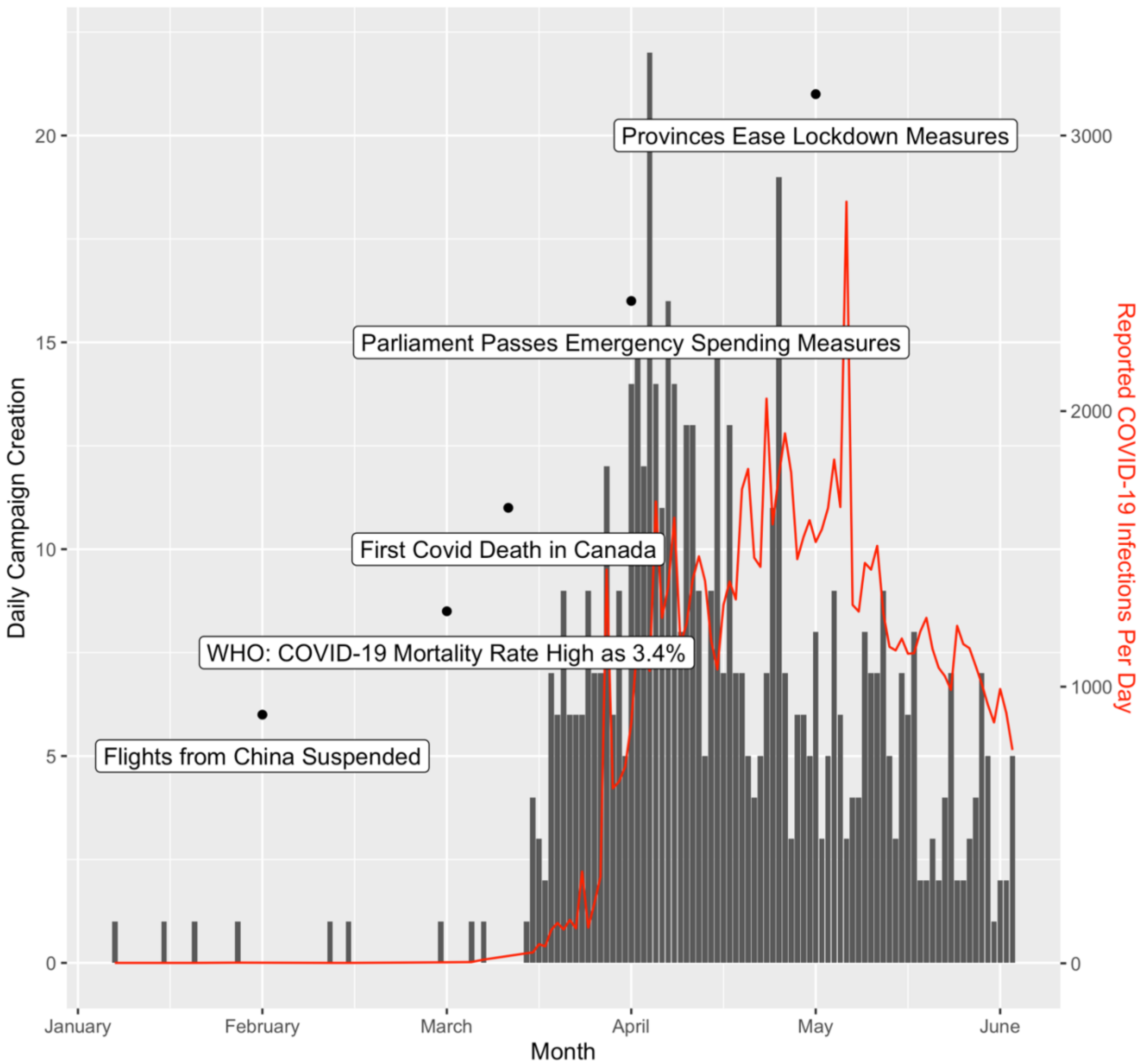

Fig. 6. Campaigns created per day vs. Canadian COVID-19 Timeline from January 1st to June 15th.

The volume of campaigns created per day are shown in grey bars, the daily number of new COVID-19 cases is shown with the red line, and various COVID-19 related events are shown with labels corresponding to the date of the black dot above each of them.

Campaign types were then plotted as a time series by date of creation to show the temporal distribution of need as it arose during the pandemic (Figure 7). Campaigns created for the purposes of funding a Canadian charity made up the largest proportion of newly created campaigns for the majority of weeks in the study period. Campaigns for the purchase or manufacture of PPE peaked in mid-April but tapered off and were not a significant proportion into May. Campaigns for family reunification in Canada from abroad and Canadian businesses were sporadic and minimal for most weeks, while campaigns for personal need were steady throughout. Finally, campaigns for international charities peaked in early April and occupied a consistent proportion of weekly campaigns until June.

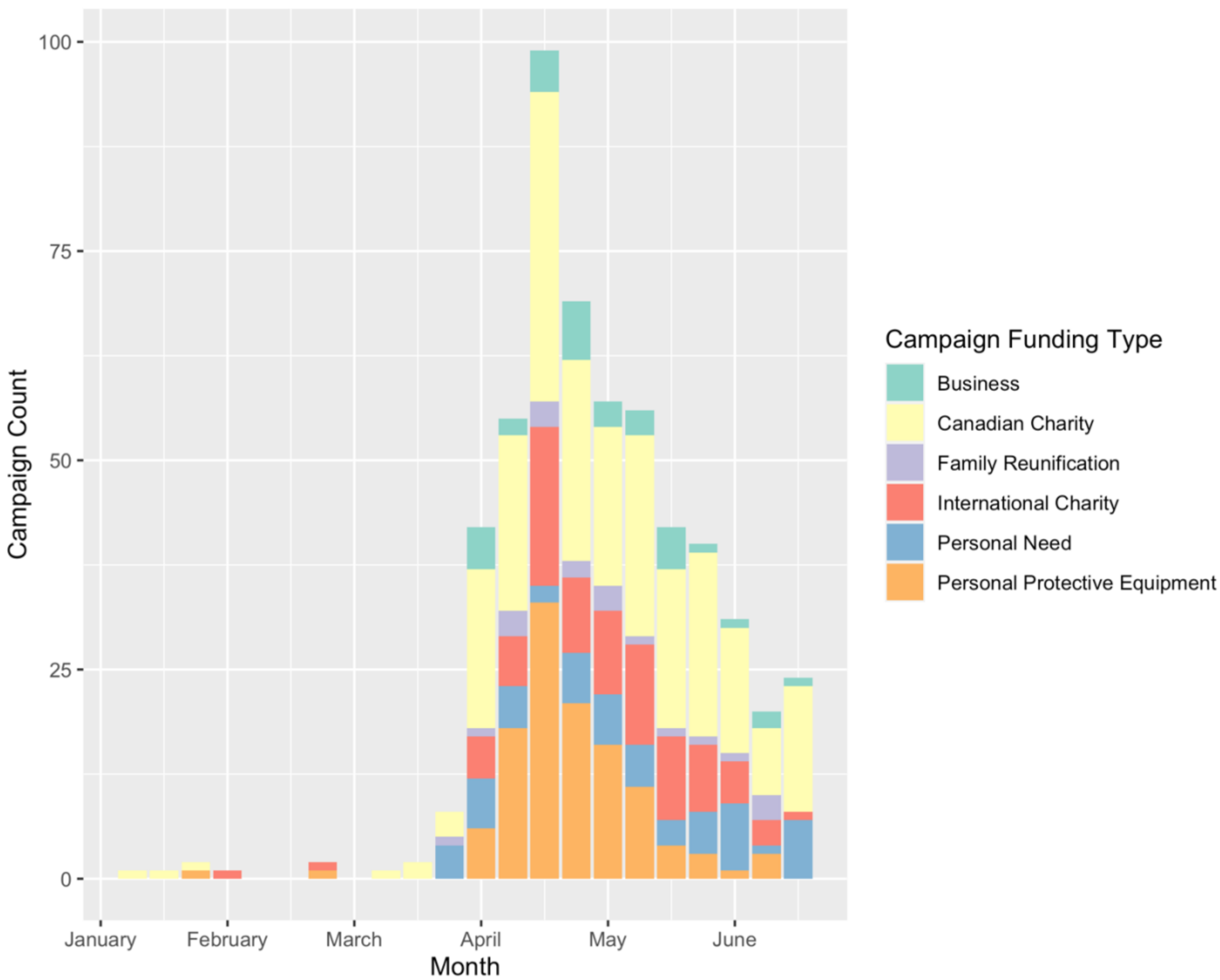

Fig. 7. Campaign class type distribution per week of the study period from January to June 2020

Shows the trends in weekly campaign creation, stratified by campaign type, across the 22 week period within which all scraped campaigns were created.

Figure 8 shows the provincial distribution of donations to campaigns in the sample, again stratified by campaign type. Ontario had by far the most donations, accounting for 64.3% of total funding in the sample. British Columbia, Alberta, and Quebec followed with 17%, 10%, and 7% of funding, respectively, with the remaining provinces accounting for the last 1%. The distribution of funding per type in each province remained similar to that breakdown at the national level, with campaigns for Canadian charities accounting for a considerable proportion in each province, and when including PPE account for the vast majority of funding dollars raised by campaigns in the sample. None of the Territories had any campaigns, and Manitoba (MB), New Brunswick (NB), Newfoundland and Labrador (NL), Nova Scotia (NS), and Saskatchewan (SK) accounted for about 1% of total funding.

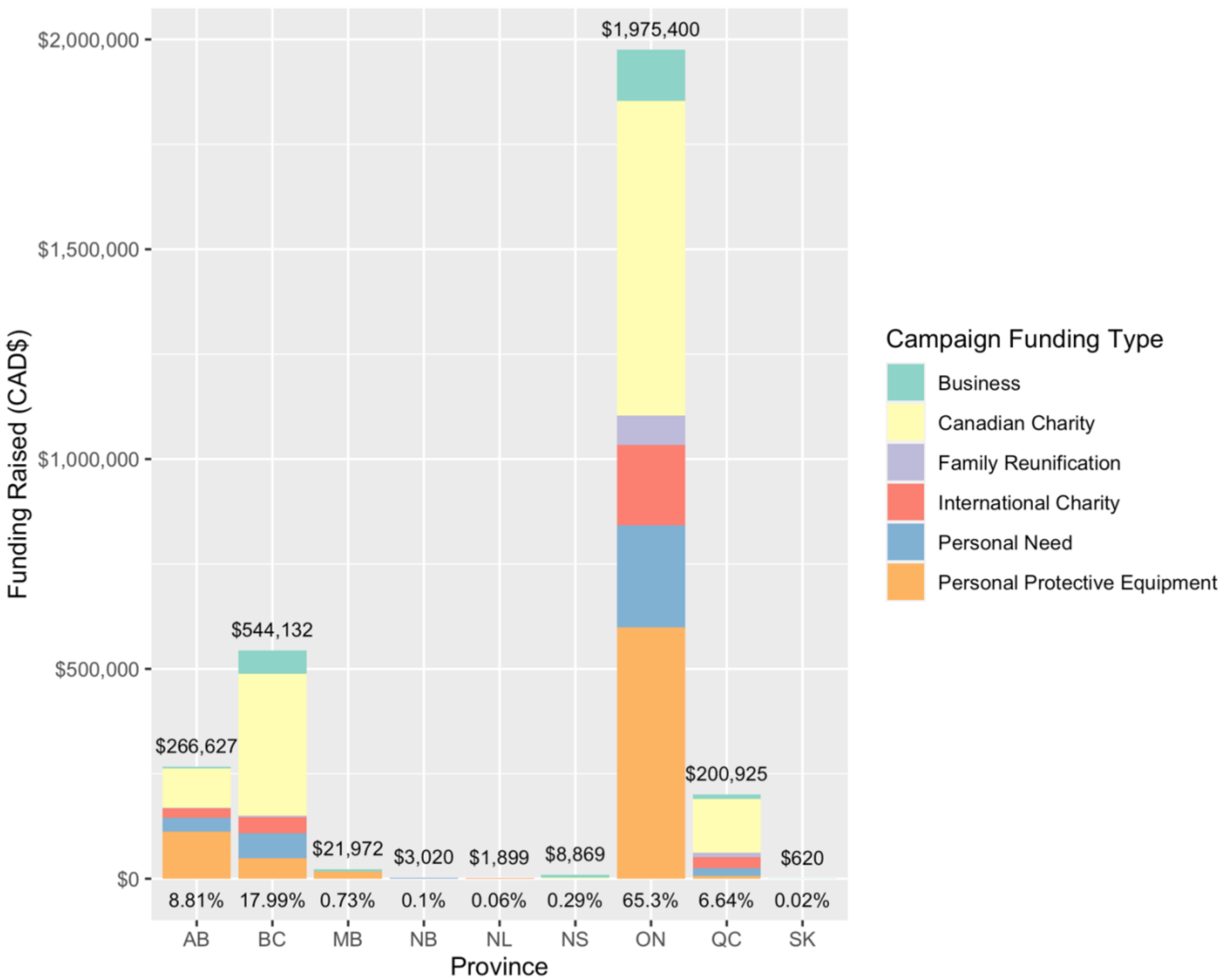

Fig. 8. Total funding dollars raised per province from the 573 campaign sample by campaign type.

Ontario received the vast majority of total funding (65.3%), followed by BC (17.99%), Alberta (8.81%), Quebec (6.64%), and the remaining provinces accounting for about 1% of total donations. Funding distributions per campaign type within the Provinces showed similar patterns to the distribution of campaigns by type at the national level.

Figure 9 shows the proportion of total dollars raised that each campaign funding group accounted for. Campaigns were grouped by how much money they had raised into bins shown along the X-axis of Figure 5. The $500 - 2500 group accounted for the most campaigns (171), but the $10,000 - $100,000 raised group accounted for the majority of funding dollars raised, with 61.83%. There were only 2 campaigns that raised more than $100,000; between them they accounted for 10.59% of all funding dollars raised by campaigns in the sample. Although there were 122 campaigns that raised between $0 - 50, these campaigns accounted for less than 0.01% of total dollars raised, and 86 campaigns raising between $50 - 250 and $250 - 500 together accounted for just 0.63% of all funding.

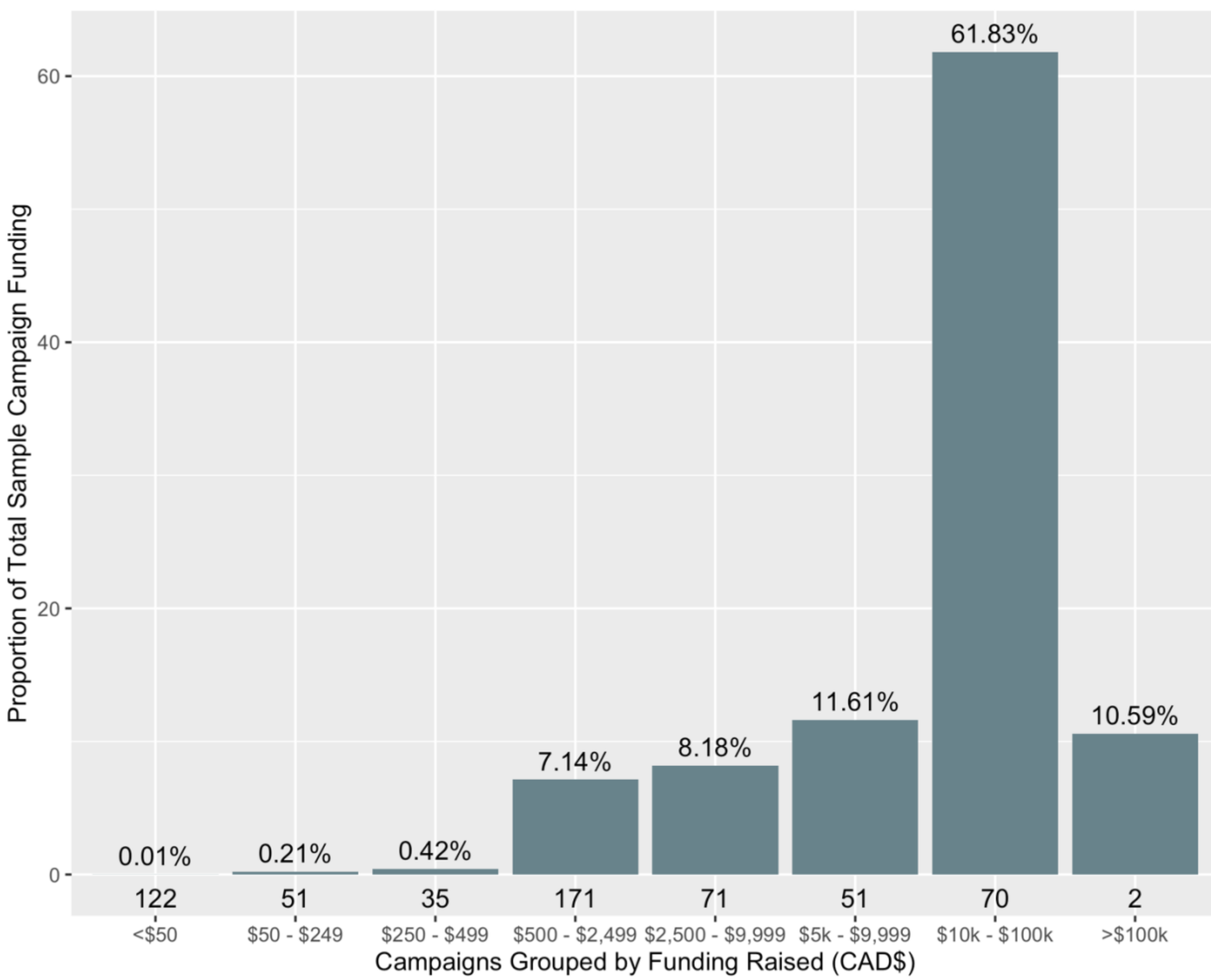

Fig. 9. Funding raised by campaigns grouped by total dollars raised.
Campaigns were assigned a grouping based on the funding raised relative to the ranges shown on the x-axis. The counts of how many campaigns were included in each group are shown underneath each corresponding bar in the chart. The proportion of funding raised by campaigns in each funding range grouping is shown on the y-axis, and is relative to the total dollars raised by all campaigns in the sample ($3,024,967).

Figure 10 shows funding raised by campaign type. Again, campaigns for Canadian Charities and PPE were representative of the majority (>70%) of funding, over $2 million (CAD). The distributions here are similar to those shown stratified by week and count of campaign type, with campaigns for businesses and family reunification accounting for about 10%, and those for Personal Need and International Charities the remaining 20%.

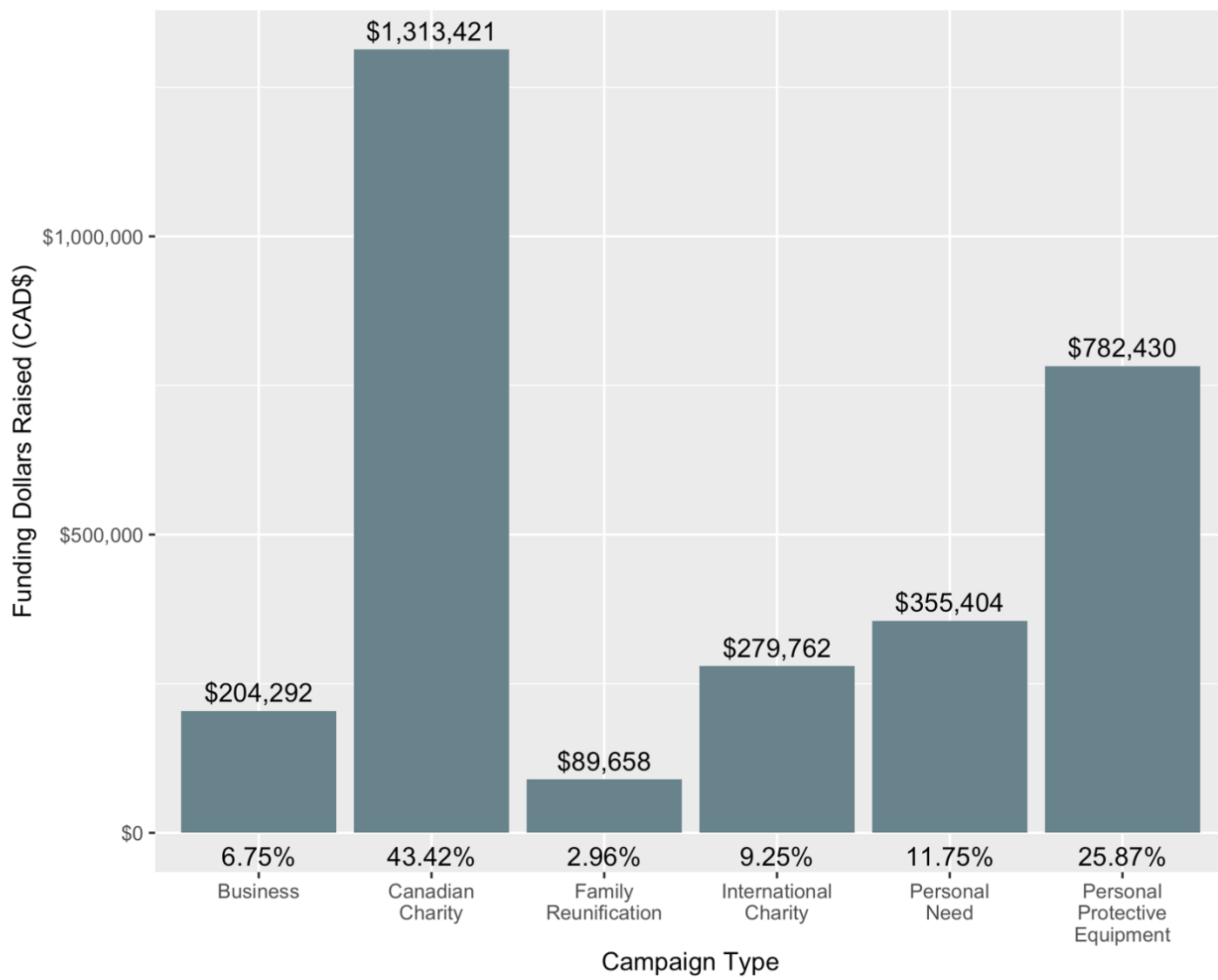

Fig. 10. Dollars and proportion of total funding raised by campaign type.
The dollar value atop each bar indicates how many dollars were raised by campaigns of the corresponding type, cumulatively. The percentage figure below the bar shows what proportion of total dollars raised by campaigns in the sample each dollar amount corresponds to (N dollars raised = $3,024,967).

Finally, figure 11 shows the spatial distributions of campaigners across Canada superimposed over color-coded provinces depicting the average level of funding received by campaigns in each province. The dense, urban areas of Canada like Toronto, Vancouver, Montreal, Ottawa, and Calgary show high concentrations of campaigners, with few campaigners shown in predominately rural or Northern areas of the country. Ontario and British Columbia (BC) showed the highest average campaign funding raised, followed by Alberta and Quebec. No campaigns created in the Territories or Prince Edward Island were included in the study sample; Saskatchewan, Newfoundland and Labrador, and New Brunswick showed fewer than 5 campaigns each, Manitoba and Nova Scotia had less than 10 each.

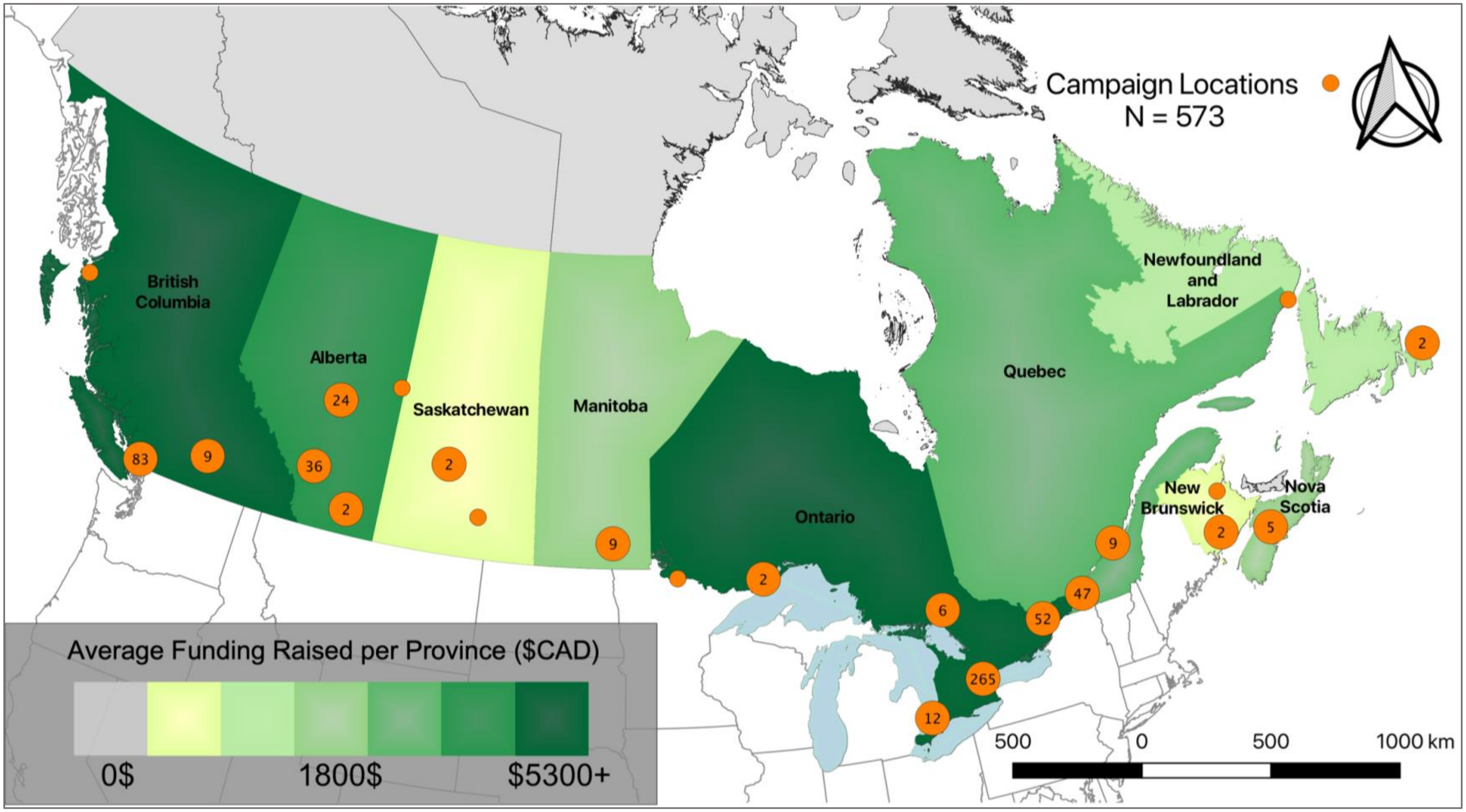

Fig. 11. Spatial distribution of campaigns in Canada and average campaign funding per province.

Discussion

Distributions of campaigns according to campaign funding objective showed similarities with other analyses of COVID-19 campaigns, specifically the prevalence of campaigns for charities (international or local), PPE, businesses, and general areas of personal need (Elmer et al. 2020). However, the campaigns in the present study showed a much higher proportion of campaigns for charitable organizations (accounting for over 50% of campaigns in the sample), and more particularly campaigns for charities within the same country/community (as opposed to charities working internationally). Campaigns for PPE were also very common in both studies (19) likely due to the widespread shortages experienced worldwide at the onset of the pandemic (21).

Campaigns for businesses struggling as a result of the pandemic and related lockdown measures were surprisingly few in number and in dollars raised, accounting for only about 6% of both campaigns in the sample and funding dollars raised despite widespread coverage of the fact that significant numbers of small businesses were and are facing bankruptcy and closure (22,23). This could be due to the time period within which these campaigns were gathered, as businesses may have been able to stay solvent during the early days of the pandemic, surviving on cash reserves and government aid. A future analysis of campaigns from later periods in the pandemic could reveal a greater prevalence of business-related campaigns, as dampened revenues and lockdown measures persist in Canada.

Campaigns for personal need in Canada made up about 14% of campaigns in the sample and almost 12% of dollars raised ($355,404), which, while a significant amount of money, was lower than expected relative to the other categories given the broad definition of 'personal need' used here and the wide variety of potential personal needs that can arise during a pandemic. For example, a previous campaign analysis found that campaigns for funerals, family/friend support, and food/supplies (all included in the present 'personal need' category) together made up 31% of sample COVID-19 related campaigns (19). Given the extensive aid offered by the Canadian government to Canadians in the amount of $2,000 monthly cheques (17), it could be that the demand for crowdfunding dollars on behalf of individual Canadians in need as a result of the pandemic is lower in the present study due to the specifically Canadian sample, as opposed to the non-country specific sample of campaigns gathered by Elmer et al. (2020).

Campaigns for family reunification, although relatively few in comparison to other categories, still raised almost $90,000 in total, and were included to demonstrate this distinct type of need. As opposed to campaigns for businesses, it seems likely that campaigns for family reunification were prevalent at the beginning of the pandemic due to the abrupt nature of border closures and orders for Canadians to return home from abroad, but tapered off as everyone who needed to arrive home had already done so, demonstrated by the absence of campaigns in this category from the last week of the study period (Figure 6).

Temporal trends in campaign creation showed that the initial spike in campaigns coincided precisely with the first surge in COVID-19 cases in Canada (Figure 2). Campaigns for charities and PPE dominated these early weeks. This shows the reactionary crowdfunding response of Canadians as it became clear that COVID-19 was to become a serious societal crisis, and how digital donation-based crowdfunding acted as a primary medium for Canadians to manage the crisis, whether by donating, or by campaigning for themselves or on behalf of others. Similar trends were seen in American crowdfunding, where creation of COVID-19 related campaigns coincided temporally with increases in detected COVID-19 cases (20).

Spatial analysis of campaigns showed a clear urban-rural divide in Canada, where the vast majority of campaigns were created in densely populated, urban areas. This is undoubtedly due to asymmetries in Canada's population distribution, as demonstrated by the vast majority of funding dollars also being raised in the most populous Provinces, however it could also be due to lower levels of access to communication technologies in rural communities or to the strength of voluntary or informal care sectors in rural communities were residents look inwards to friends and neighbours for support rather than outwards to society at large (15).

This study focused on analyzing the spatio-temporal patterns in crowdsourced data from GoFundMe.com. As such, it is differentiated from studies using similar data with a focus on inferential statistics (11). Studies like Kenworthy et al. (2020) used a randomized sample of campaign data derived from an extremely large dataset (>165,000 campaigns) to draw and extrapolate inferential conclusions about crowdfunding users, campaign characteristics, and crowdfunding success at the national level (11). Analyses such as this, which favour statistical significance and inference instead of considering spatio-temporal variables, are therefore distinct from the purview of the present paper because conclusions are drawn from aspatial factors. However, clearly both approaches reveal patterns in the data and are complementary.

Limitations

The present study acknowledges limitations to the methods used in gathering and analyzing crowdfunding campaigns. First, it is entirely possible that the search protocol over/under exaggerated certain types of campaigns, or entirely missed campaigns that would have otherwise met inclusion criteria. For example, the large number of campaigns on behalf of Canadian charities included in the sample could be due in part to the utilization of 'Canada' in the GoFundMe search, which would favour campaigns for organizations like the 'Canada Food Bank' or 'Canadian United Way' as the word 'Canada' is explicitly used in the title and campaign description. In contrast, campaigns for an individual person looking for donations from their local network would perhaps not include the word 'Canada', even if the campaign was in Canada because it would be self-evident to those donating. Given the nature of the campaign data and that the only way to access it is through the front-end of the website (as opposed to having access to the back-end which would allow for more options when querying), it is impossible to say whether or not all the relevant campaigns were included. However, since all the campaigns included in the final analysis were individually examined and vetted for inclusion criteria, we can say that the sample represents a robust, though likely incomplete, collection of relevant campaigns.

Another limiting factor is that the sample here represents a snapshot of campaigns at the time they were harvested. This means that, although campaigns as far back as January 2020 were gathered, it is possible that many campaigns were created and then deleted before having a chance to be gathered. This could lead to an over representation of successful campaigns, as unsuccessful campaigns are rather more likely to be deleted after a short time (15). This snapshot also only gathered information on the first 6 months of the pandemic; future studies could investigate the characteristics of crowdfunding campaigns during the latter portions of the pandemic to observe whether the same types of campaign characteristics persisted.

Conclusion

Charitable crowdfunding has become an increasingly popular way for people to deal with unmanageable expenses, and the expenses incurred by many millions of Canadians over the last twelve months due to COVID-19 are no exception. Collection and analysis of COVID-19 related crowdfunding campaigns from GoFundMe revealed a significant number of Canadians who had turned to crowdfunding to help themselves personally, but more so on behalf of charitable organizations in Canada and abroad. The spatio-temporal approach allowed for identification of trends across Canada and across time with respect to COVID-19 and how the pandemic developed from its beginnings in February to the lull in cases and relaxing of restrictions in June. Funding and campaign need type were the primary dimensions of analysis, both of which showed clear patterns when analyzed across time and space in Canadian provinces, with the most pronounced numbers of campaigns and funding dollars raised being in BC and Ontario during the months of April and May. Campaign creation was also greatest in the period between when the initial lockdown measures were imposed and when the Canada Emergency Response Benefit was announced.

The study design in terms of campaign collection, classification, and analysis allowed for campaigns on the charitable crowdfunding website GoFundMe to be used as a proxy for understanding how these types of platforms facilitate need in times of crisis, as well as the specific areas of need in Canada that were revealed as a result of the pandemic. While other studies on COVID-19-crowdfunding have been published, our results demonstrate that spatio-temporal specificity allows for a nuanced study in how pandemic-related needs and crowdfunding change over time and react to events and policies in specific contexts. Future studies utilizing a similar collection protocol should aim to minimize the campaign selection bias in the collection process by performing multiple searches using utilizing multiple search terms across a period of weeks or months, to ensure the greatest possible number of relevant campaigns are collected.

Chapter 4.

Conclusion

Social media continues to impact society in many different areas, one of which remains the area of spatial research. The ways in which massive amounts of spatial user data are changing the means by which spatial patterns can be studied are numerous, and this thesis sought to illustrate the methods that make these impacts possible. To this end, the roots of spatial analysis were traced in chapter 1 to establish the technological and theoretical bases from which GIS and GIScience emerged. Chapter 2 explored the next iteration of the co-evolution of GIS and technology by reviewing the ways in which spatial social media data have been exploited for GIS research purposes. Chapter 3 illustrated the reviewed methods in a concrete case study of spatial and temporal crowdfunding patterns in Canada during the COVID-19 pandemic. Together, these chapters illuminate the research capabilities enabled over the past decade by the massive uptake of social media and by extension the immense volumes of spatial user data generated each day.

Chapter 2 Contributions

Chapter 2 focused on methods of collecting, analyzing, and visualizing location-based social media data. The purpose of this review was to establish a practical array of methods used in social media analysis and to demonstrate a catalogue of LBSMD applications in GIS research.

Information on this topic was gathered and analyzed using a scoping review protocol established by Arksey and O'Malley (2005) (15). A total of 222 articles were included for full-text review during this process, after approximately 900 papers were eliminated through abstract and title review. Articles were selected based on their inclusion of LBSMD as a primary data resource, and were analyzed to detect trends in collection, analysis, and visualization of location-based data from social media platforms.

Methods of collection were found to be dominated by use of the Twitter and Flickr API's, while a minority of researchers opted to use manual or automated custom search strategies for sources that do not offer API services. Data queries were found to the main drivers of collected sample results, with different inputs for locational, keyword content, and user ID parameters being the main variables of collection.

Methods of analysis were split into those analyzing the ST patterns contained in the metadata, or those analyzing the text or image content within the posts themselves. ST analysis consisted mainly of detecting anomalous activity with reference to specific times or locations, while content analysis sought to derived semantic meaning or sentiment from post text. In the majority of cases, both types of analysis were utilized in unison to both categorize posts by their semantic or sentiment information and then analyze the ST patterns of those posts to identify patterns.

Methods of visualization were commonly derived from the standard GIS toolkit, and aside from tables and graphs mainly included cartographic maps that displayed visual information derived from collections of LBSM-point data. Some studies simply plotted the data as ST points to demonstrate patterns, while others transformed the data into raster surfaces, thematic area-based maps, qualitative content diagrams like spatial topic-models, transmission diagrams, and web maps.

The chapter closed with a discussion on the increasing prevalence of algorithms in geography, especially in big-data driven geographies like those generated through use of LBSMD. While algorithms enable geographers to generate knowledge at rates and scales never before possible, they also come with inherent limitations that must always be considered when working with large datasets.

Chapter 3 Contributions

Chapter 3 mobilizes the methods reviewed in chapter 2 in a case study of spatial-temporal patterns in Canadian online crowdfunding activity during the first 6 months of the COVID-19 pandemic. First, a custom-built automated data collection strategy was employed to gather a large sample of pandemic-related campaigns. Next, the content of the gathered campaigns was analyzed manually to generate classifiers based on the purpose of each campaign. Finally, the spatio-temporal patterns in the data were analyzed to illuminate differences in campaign success and creation as a function spatial location within Canada and the groups of identified campaign purposes.

Meaningful differentiation in these factors was found across Canada, with provinces like BC and Ontario accounting for the vast majority of funding dollars raised. The patterns in campaign creation also strongly reflected events in the pandemic-related news cycle, indicating the influences of events on crowdfunding behaviours. Campaign funding purposes also showed differentials across time, with campaigns for PPE falling off substantially towards the start of June 2020.

Closing remarks

From the outset the purpose of this work was to explore the implications of social media data to spatial research. However, the onset of the COVID-19 pandemic in 2020 provided an opportunity to utilize these GIS-social media methods for a topic of analysis that remains somewhat distant from traditional definitions of social media. As discussed though, the incongruence between crowdfunding platforms and social media platforms remains a minor one, and therefore did not stand in the way of the meaningful application of the discussed methods of LBSMD utilization as the characteristics of data from both sources are similar. While additional work is certainly required to further establish a concrete array of methods for LBSMD use, as well as in uncovering spatio-temporal patterns of crowdfunding behaviours using these methods, this thesis' contributions to the geographies of big-data analysis and in particular social media analysis will ideally be used to inform other spatially-oriented researchers in similar analytic pursuits.

References

Chapter 1 references

1. Goodchild MF. Part 1 Spatial analysts and GIS practitioners. Journal of Geographical Systems. 2000 Mar 9;2(1):5–10.

2. Longley Paul A, Maguire David J, Rhind David W, Goodchild Michael F. Geographic Information: Science, Systems, and Society. 4th Edition. 2015;1–1.

3. Goodchild MF. Reimagining the history of GIS. Annals of GIS. 2018 Jan 2;24(1):1–8.

4. Goodchild MF. Geographical information science. International Journal of Geographical Information Systems. 1992 Jan 1;6(1):31–45.

5. Wright DJ, Goodchild MF, Proctor JD. GIS: Tool or Science? Demystifying the Persistent Ambiguity of GIS as “Tool” Versus “Science.” Annals of the Association of American Geographers. 1997;87(2):346–62.

6. Most used social media 2021 [Internet]. Statista. [cited 2021 May 5]. Available from: https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/

7. U.S. mobile social media traffic share 2019 [Internet]. Statista. [cited 2021 May 5]. Available from: https://www.statista.com/statistics/477368/us-social-media-visits-share/

8. Pew Research. Demographics of Mobile Device Ownership and Adoption in the United States [Internet]. Pew Research Center: Internet, Science & Tech. [cited 2021 Mar 19]. Available from: https://www.pewresearch.org/internet/fact-sheet/mobile/

9. Sui D, Goodchild M. The convergence of GIS and social media: Challenges for GIScience. International Journal of Geographical Information Science. 2011;25(11):1737–48.

10. Crooks A, Croitoru A, Stefanidis A, Radzikowski J. #Earthquake: Twitter as a Distributed Sensor System. Transactions in GIS. 2013;17(1):124–47.

11. Padmanabhan A, Wang S, Cao G, Hwang M, Zhang Z, Gao Y, et al. FluMapper: A cyberGIS application for interactive analysis of massive location-based social media. Concurrency and Computation-Practice & Experience. 2014 Sep 10;26(13):2253–65.

12. Ristea A, Andresen MA, Leitner M. Using tweets to understand changes in the spatial crime distribution for hockey events in Vancouver. Canadian Geographer-Geographe Canadien. 2018 FAL;62(3):338–51.

13. Huang X, Wang C, Li Z. A near real-time flood-mapping approach by integrating social media and post-event satellite imagery. Annals of GIS. 2018;24(2):113–23.

Chapter 2 references

1. Paul A. Longley. Geographic Information: Science, Systems, and Society. In: Geographic information science & systems. 4th ed. Wiley; 2015. p. 20, 21.

2. Sui D, Goodchild M. The convergence of GIS and social media: Challenges for GIScience. International Journal of Geographical Information Science. 2011;25(11):1737–48.

3. Steiger E, Albuquerque JP de, Zipf A. An Advanced Systematic Literature Review on Spatiotemporal Analyses of Twitter Data. Transactions in GIS. 2015;19(6):809–34.

4. Goodchild MF. Citizens as sensors: the world of volunteered geography. GeoJournal. 2007 Nov 30;69(4):211–21.

5. Sloan L, Morgan J. Who Tweets with Their Location? Understanding the Relationship between Demographic Characteristics and the Use of Geoservices and Geotagging on Twitter. PLoS One [Internet]. 2015 Nov 6 [cited 2021 Feb 5];10(11). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4636345/

6. Li L, Goodchild MF, Xu B. Spatial, temporal, and socioeconomic patterns in the use of Twitter and Flickr. Cartography and Geographic Information Science. 2013 Mar 1;40(2):61–77.

7. boyd danah, Crawford K. CRITICAL QUESTIONS FOR BIG DATA: Provocations for a cultural, technological, and scholarly phenomenon. Information, Communication & Society. 2012 Jun;15(5):662–79.

8. Jensen EA. Putting the methodological brakes on claims to measure national happiness through Twitter: Methodological limitations in social media analytics. Plos One. 2017 Sep 7;12(9):e0180080.

9. Huang Q, Xu C. A data-driven framework for archiving and exploring social media data. Annals of GIS. 2014;20(4):265–77.

10. Stock K. Mining location from social media: A systematic review. Computers Environment and Urban Systems. 2018 Sep;71:209–40.

11. Wiegmann M, Kersten J, Senaratne H, Potthast M, Klan F, Stein B. Opportunities and Risks of Disaster Data from Social Media: A Systematic Review of Incident Information [Internet]. Databases, GIS, Remote Sensing, Early Warning Systems and Monitoring Technologies; 2020 Jul [cited 2021 Apr 5]. Available from: https://nhess.copernicus.org/preprints/nhess-2020-221/

12. Mirzaalian F, Halpenny E. Social media analytics in hospitality and tourism: A systematic literature review and future trends. Journal of Hospitality and Tourism Technology. 2019 Jan 1;10(4):764–90.

13. Muniz-Rodriguez K, Ofori SK, Bayliss LC, Schwind JS, Diallo K, Liu M, et al. Social Media Use in Emergency Response to Natural Disasters: A Systematic Review With a Public Health Perspective. Disaster med public health prep. 2020 Feb;14(1):139–49.

14. Wilkins EJ, Wood SA, Smith JW. Uses and Limitations of Social Media to Inform Visitor Use Management in Parks and Protected Areas: A Systematic Review. Environmental Management. 2021 Jan 1;67(1):120–32.

15. Stefanidis A, Crooks A, Radzikowski J. Harvesting ambient geospatial information from social media feeds. GeoJournal. 2013;78(2):319–38.

16. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. International Journal of Social Research Methodology. 2005 Feb 1;8(1):19–32.

17. Pham MT, Rajić A, Greig JD, Sargeant JM, Papadopoulos A, McEwen SA. A scoping review of scoping reviews: advancing the approach and enhancing the consistency. Research Synthesis Methods. 2014;5(4):371–85.

18. Crooks A, Croitoru A, Stefanidis A, Radzikowski J. #Earthquake: Twitter as a Distributed Sensor System. Transactions in GIS. 2013;17(1):124–47.

19. Resch B, Uslander F, Havas C. Combining machine-learning topic models and spatiotemporal analysis of social media data for disaster footprint and damage assessment. Cartography and Geographic Information Science. 2018;45(4):362–76.

20. Cao K, Huang Q. Geo-sensor(s) for potential prediction of earthquakes: can earthquake be predicted by abnormal animal phenomena? Annals of GIS. 2018;24(2):125–38.

21. Smith L, Liang Q, James P, Lin W. Assessing the utility of social media as a data source for flood risk management using a real-time modelling framework. Journal of Flood Risk Management. 2017;10(3):370–80.

22. Lwin KK, Zettsu K, Sugiura K. Geovisualization and Correlation Analysis between Geotagged Twitter and JMA Rainfall Data: Case of Heavy Rain Disaster in Hiroshima. Leung Y, Guo DS, Chen CC, editors. 2015.

23. Huang X, Wang C, Li Z. A near real-time flood-mapping approach by integrating social media and post-event satellite imagery. Annals of GIS. 2018;24(2):113–23.

24. Allen C, Tsou M-H, Aslam A, Nagel A, Gawron J-M. Applying GIS and Machine Learning Methods to Twitter Data for Multiscale Surveillance of Influenza. Plos One. 2016 Jul 25;11(7):e0157734.

25. Robertson C, Yee L. Avian Influenza Risk Surveillance in North America with Online Media. PLoS One. 2016;11(11):e0165688.

26. Padmanabhan A, Wang S, Cao G, Hwang M, Zhang Z, Gao Y, et al. FluMapper: A cyberGIS application for interactive analysis of massive location-based social media. Concurrency and Computation-Practice & Experience. 2014 Sep 10;26(13):2253–65.

27. Wakamiya S, Kawai Y, Aramaki E. Twitter-Based Influenza Detection After Flu Peak via Tweets With Indirect Information: Text Mining Study. JMIR Public Health Surveill. 2018 Sep 25;4(3):e65.

28. Knox JA, Rackley JA, Black AW, Gensini VA, Butler M, Dunn C, et al. TORNADO DEBRIS CHARACTERISTICS AND TRAJECTORIES DURING THE 27 APRIL 2011 SUPER OUTBREAK AS DETERMINED USING SOCIAL MEDIA DATA. Bulletin of the American Meteorological Society. 2013 Sep;94(9):1371–80.

29. Ford B, Burke M, Lassman W, Pfister G, Pierce JR. Status update: Is smoke on your mind? Using social media to assess smoke exposure. Atmospheric Chemistry and Physics. 2017;17(12):7541–54.

30. Basic Privacy Settings & Tools | Facebook Help Center [Internet]. [cited 2021 Jan 30]. Available from: https://www.facebook.com/help/325807937506242

31. About public and protected Tweets [Internet]. [cited 2021 Jan 30]. Available from: https://help.twitter.com/en/safety-and-security/public-and-protected-tweets

32. Sun Y, Bakillah M. Analysis of tourist activity based on the tracking data collected by flickr. gisScience - Die Zeitschrift fur Geoinformatik. 2013;(3):87–92.

33. Zaccomer GP, Grassetti L. Flickr as a Tool for the Analysis of Photographic Tourism: The Estimation of Geotagging Rate and Its Use for Mapping the World. In: Gervasi O, Murgante B, Misra S, Borruso G, Torre CM, Rocha A, et al., editors. Computational Science and Its Applications - Iccsa 2017, Pt Iii. 2017. p. 266–81.

34. Lee JY, Tsou M-H. Mapping Spatiotemporal Tourist Behaviors and Hotspots Through Location-Based Photo-Sharing Service (Flickr) Data. In: Kiefer P, Huang H, VanDeWeghe N, Raubal M, editors. Progress in Location Based Services 2018. 2018. p. 315–34.

35. ElQadi MM, Dorin A, Dyer A, Burd M, Bukovac Z, Shrestha M. Mapping species distributions with social media geo-tagged images: Case studies of bees and flowering plants in Australia. Ecological Informatics. 2017;39:23–31.

36. Sinclair M, Ghermandi A, Sheela AM. A crowdsourced valuation of recreational ecosystem services using social media data: An application to a tropical wetland in India. Science of the Total Environment. 2018 Nov 15;642:356–65.

37. Levin N, Lechner AM, Brown G. An evaluation of crowdsourced information for assessing the visitation and perceived importance of protected areas. Applied Geography. 2017 Feb;79:115–26.

38. Tenerelli P, Demar U, Luque S. Crowdsourcing indicators for cultural ecosystem services: A geographically weighted approach for mountain landscapes. Ecological Indicators. 2016;64:237–48.

39. Gschwend C, Purves RS. Exploring Geomorphometry through User Generated Content: Comparing an Unsupervised Geomorphometric Classification with Terms Attached to Georeferenced Images in Great Britain. Transactions in Gis. 2012 Aug;16(4):499–522.

40. Chen J, Shaw S-L. Representing the Spatial Extent of Places Based on Flickr Photos with a Representativeness-Weighted Kernel Density Estimation. In: Miller JA, OSullivan D, Wiegand N, editors. Geographic Information Science, (GIScience 2016). 2016. p. 130–44.

41. Oteros-Rozas E, Martin-Lopez B, Fagerholm N, Bieling C, Plieninger T. Using social media photos to explore the relation between cultural ecosystem services and landscape features across five European sites. Ecological Indicators. 2018;94:74–86.

42. Donahue ML, Keeler BL, Wood SA, Fisher DM, Hamstead ZA, McPhearson T. Using social media to understand drivers of urban park visitation in the Twin Cities, MN. Landscape and Urban Planning. 2018 Jul;175:1–10.

43. Wu C, Ye X, Ren F, Wan Y, Ning P, Du Q. Spatial and Social Media Data Analytics of Housing Prices in Shenzhen, China. PLoS One. 2016;11(10):e0164553.

44. Yao Y, Zhang J, Hong Y, Liang H, He J. Mapping fine-scale urban housing prices by fusing remotely sensed imagery and social media data. Transactions in GIS. 2018;22(2):561–81.

45. Wu C, Ye X, Ren F, Du Q. Check-in behaviour and spatio-temporal vibrancy: An exploratory analysis in Shenzhen, China. Cities. 2018;77:104–16.

46. Chen Y, Parkins JR, Sherren K. Using geo-tagged Instagram posts to reveal landscape values around current and proposed hydroelectric dams and their reservoirs. Landscape and Urban Planning. 2018 Feb;170:283–92.

47. Nummi P. Crowdsourcing Local Knowledge with PPGIS and Social Media for Urban Planning to Reveal Intangible Cultural Heritage. Urban Planning. 2018;3(1):100–15.

48. Ahmouda A, Hochmair HH, Cvetojevic S. Analyzing the effect of earthquakes on OpenStreetMap contribution patterns and tweeting activities. Geo-Spatial Information Science. 2018;21(3):195–212.

49. hiQ Labs, Inc. v. LinkedIn Corp. Vol. 273, F. Supp. 3d. 2017. p. 1099.

50. Yu S, Sun W, Jia M. A Dynamic Proxy Based Crawler Strategy for Data Collection on CyberGIS. In: 2018 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (cyberc 2018). 2018. p. 483–7.

51. Basch CH, Blankenship EB, Fung IC-H, Yarborough CC, Rustin RC, Yin J. Eastern Equine Encephalitis and YouTube videos: A content analysis. Infection Disease & Health. 2018 Dec;23(4):197–202.

52. Giovos I, Keramidas I, Antoniou C, Deidun A, Font T, Kleitou P, et al. Identifying recreational fisheries in the Mediterranean Sea through social media. Fisheries Management and Ecology. 2018;25(4):287–95.

53. Lewis QW, Park E. Volunteered Geographic Videos in Physical Geography: Data Mining from YouTube. Annals of the American Association of Geographers. 2018;108(1):52–70.

54. Martin ME, Schuurman N. Area-Based Topic Modeling and Visualization of Social Media for Qualitative GIS. Annals of the American Association of Geographers. 2017 Sep;107(5):1028–39.

55. Yang W, Mu L. GIS analysis of depression among Twitter users. Applied Geography. 2015 Jun;60:217–23.

56. Cvetojevic S, Hochmair HH. Analyzing the spread of tweets in response to Paris attacks. Computers Environment and Urban Systems. 2018 Sep;71:14–26.

57. Tweet geospatial metadata [Internet]. 2020 [cited 2020 Jan 22]. Available from: https://developer.twitter.com/en/docs/tutorials/tweet-geo-metadata

58. Li Z, Wang C, Emrich CT, Guo D. A novel approach to leveraging social media for rapid flood mapping: a case study of the 2015 South Carolina floods. Cartography and Geographic Information Science. 2018 Mar 4;45(2):97–110.

59. Sakaki T, Okazaki M, Matsuo Y. Earthquake shakes Twitter users: real-time event detection by social sensors. 2010;10.

60. Acampora G, Anastasio P, Risi M, Tortora G, Vitiello A. Automatic Event Geo-Location in Twitter. IEEE ACCESS. 2020;8:128213–23.

61. Cheng T, Wicks T. Event detection using Twitter: a spatio-temporal approach. PLoS One. 2014;9(6):e97807.

62. Farnaghi M, Ghaemi Z, Mansourian A. Dynamic Spatio-Temporal Tweet Mining for Event Detection: A Case Study of Hurricane Florence. International Journal of Disaster Risk Science. 2020;11(3):378–93.

63. Xu S, Li S, Huang W. A spatial-temporal-semantic approach for detecting local events using geo-social media data. Transactions in GIS. 2020;24(1):142–73.

64. Xu Z, Liu Y, Yen NY, Mei L, Luo X, Wei X, et al. Crowdsourcing Based Description of Urban Emergency Events Using Social Media Big Data. IEEE TRANSACTIONS ON CLOUD COMPUTING. 2020 Jun;8(2):387–97.

65. Gao Y, Wang S, Padmanabhan A, Yin J, Cao G. Mapping spatiotemporal patterns of events using social media: a case study of influenza trends. International Journal of Geographical Information Science. 2018;32(3):425–49.

66. See L, Schepaschenko D, Lesiv M, McCallum I, Fritz S, Comber A, et al. Building a hybrid land cover map with crowdsourcing and geographically weighted regression. Isprs Journal of Photogrammetry and Remote Sensing. 2015 May;103:48–56.

67. Lwin KK, Sugiura K, Zettsu K. Space-time multiple regression model for grid-based population estimation in urban areas. International Journal of Geographical Information Science. 2016;30(8):1579–93.

68. Ristea A, Andresen MA, Leitner M. Using tweets to understand changes in the spatial crime distribution for hockey events in Vancouver. Canadian Geographer-Geographe Canadien. 2018 FAL;62(3):338–51.

69. Rybarczyk G, Banerjee S, Starking-Szymanski MD, Shaker RR. Travel and us: the impact of mode share on sentiment using geo-social media and GIS. Journal of Location Based Services. 2018;12(1):40–62.

70. Chen Y, Liu X, Gao W, Wang RY, Li Y, Tu W. Emerging social media data on measuring urban park use. Urban Forestry & Urban Greening. 2018 Apr;31:130–41.

71. Cao Z, Wu Z, Li S, Guo G, Song S, Deng Y, et al. Explicit Spatializing Heat-Exposure Risk and Local Associated Factors by coupling social media data and automatic meteorological station data. Environmental Research [Internet]. 2020;188. Available from: http://dx.doi.org/10.1016/j.envres.2020.109813

72. Karmegam D, Mappillairaju B. Spatio-temporal distribution of negative emotions on Twitter during floods in Chennai, India, in 2015: a post hoc analysis. INTERNATIONAL JOURNAL OF HEALTH GEOGRAPHICS. 2020 May 28;19(1).

73. Li S, Lyu D, Huang G, Zhang X, Gao F, Chen Y, et al. Spatially varying impacts of built environment factors on rail transit ridership at station level: A case study in Guangzhou, China. Journal of Transport Geography [Internet]. 2020;82. Available from: http://dx.doi.org/10.1016/j.jtrangeo.2019.102631

74. Tu W, Zhu T, Xia J, Zhou Y, Lai Y, Jiang J, et al. Portraying the spatial dynamics of urban vibrancy using multisource urban big data. Computers, Environment and Urban Systems [Internet]. 2020;80. Available from: http://dx.doi.org/10.1016/j.compenvurbsys.2019.101428

75. Fan T, Guo N, Ren Y. Consumer clusters detection with geo-tagged social network data using DBSCAN algorithm: a case study of the Pearl River Delta in China. GeoJournal. 2021;86(1):317–37.

76. Sun Y, Wang S, Zhang X, Chan TO, Wu W. Estimating local-scale domestic electricity energy consumption using demographic, nighttime light imagery and Twitter data. Energy [Internet]. 2021;226. Available from: http://dx.doi.org/10.1016/j.energy.2021.120351

77. Luini LP, Cardellicchio D, Felletti F, Marucci FS. Socio-Spatial Intelligence: social media and spatial cognition for territorial behavioral analysis. Cognitive Processing. 2015 Sep;16:S299–303.

78. Huang Q, Wong DWS. Modeling and Visualizing Regular Human Mobility Patterns with Uncertainty: An Example Using Twitter Data. Annals of the Association of American Geographers. 2015 Nov 2;105(6):1179–97.

79. D'Silva K, Noulas A, Musolesi M, Mascolo C, Sklar M. Predicting the temporal activity patterns of new venues. EPJ Data Sci. 2018;7(1):13.

80. Liu X, Macedo J, Zhou T, Shen L, Liao Y, Zhou Y. Evaluation of the utility efficiency of subway stations based on spatial information from public social media. Habitat International. 2018 Sep;79:10–7.

81. Jin C, Nara A, Yang J-A, Tsou M-H. Similarity measurement on human mobility data with spatially weighted structural similarity index(SpSSIM). Transactions in GIS. 2020;24(1):104–22.

82. Wu R, Luo G, Jin Q, Shao J, Lu C-T. Learning evolving users behaviors on location-based social networks. GeoInformatica. 2020;24(3):713–43.

83. Dietz LW, Sen A, Roy R, Woerndl W. Mining trips from location-based social networks for clustering travelers and destinations. INFORMATION TECHNOLOGY & TOURISM. 2020 Mar;22(1, SI):131–66.

84. Tang J, Wan L, Nochta T, Schooling J, Yang T. Exploring Resilient Observability in Traffic-Monitoring Sensor Networks: A Study of Spatial-Temporal Vehicle Patterns. ISPRS INTERNATIONAL JOURNAL OF GEO-INFORMATION. 2020 Apr;9(4).

85. Brito DEF, Assuncao RM, Souza RCSNP, Meira W. SCPP: A Point Process-based Clustering of Spatial Visiting Patterns. ACM Transactions on Spatial Algorithms and Systems [Internet]. 2021;7(1). Available from: http://dx.doi.org/10.1145/3423405

86. Calafiore A, Palmer G, Comber S, Arribas-Bel D, Singleton A. A geographic data science framework for the functional and contextual analysis of human dynamics within global cities. Computers, Environment and Urban Systems [Internet]. 2021;85. Available from: http://dx.doi.org/10.1016/j.compenvurbsys.2020.101539

87. Liu X, Huang Q, Gao S, Xia J. Activity knowledge discovery: Detecting collective and individual activities with digital footprints and open source geographic data. Computers, Environment and Urban Systems [Internet]. 2021;85. Available from: http://dx.doi.org/10.1016/j.compenvurbsys.2020.101551

88. Jestico B, Nelson T, Winters M. Mapping ridership using crowdsourced cycling data. Journal of Transport Geography. 2016 Apr;52:90–7.

89. Ferster C, Nelson T, Laberee K, Winters M. Mapping bicycling exposure and safety risk using Strava Metro. Applied Geography [Internet]. 2021;127. Available from: http://dx.doi.org/10.1016/j.apgeog.2021.102388

90. Nelson T, Ferster C, Laberee K, Fuller D, Winters M. Crowdsourced data for bicycling research and practice. Transport Reviews. 2021 Jan 2;41(1):97–114.

91. Nelson T, Roy A, Ferster C, Fischer J, Brum-Bastos V, Laberee K, et al. Generalized model for mapping bicycle ridership with crowdsourced data. Transportation Research Part C: Emerging Technologies. 2021 Apr;125:102981.

92. Wang Y, Li H, Wang T, Zhu J. The mining and analysis of emergency information in sudden events based on social media. Wuhan Daxue Xuebao (Xinxi Kexue Ban)/Geomatics and Information Science of Wuhan University. 2016;41(3):290–7.

93. Wang Y, Wang T, Ye X, Zhu J, Lee J. Using social media for emergency response and urban sustainability: A case study of the 2012 Beijing rainstorm. Sustainability (Switzerland). 2016;8(1):1–17.

94. Xu Z, Zhang H, Sugumaran V, Choo K-KR, Mei L, Zhu Y. Participatory sensing-based semantic and spatial analysis of urban emergency events using mobile social media. Eurasip Journal on Wireless Communications and Networking. 2016 Feb 9;44.

95. Marquet RL, Bartelds AIM, van Noort SP, Koppeschaar CE, Paget J, Schellevis FG, et al. Internet-based monitoring of influenza-like illness (ILI) in the general population of the Netherlands during the 2003-2004 influenza season. BMC Public Health. 2006 Oct 4;6:242.

96. Widener MJ, Li W. Using geolocated Twitter data to monitor the prevalence of healthy and unhealthy food references across the US. Applied Geography. 2014 Oct;54:189–97.

97. Wang Y, Xu K, Kang Y, Wang H, Wang F, Avram A. Regional influenza prediction with sampling twitter data and PDE model. International Journal of Environmental Research and Public Health [Internet]. 2020;17(3). Available from: http://dx.doi.org/10.3390/ijerph17030678

98. Pradeepa S, Manjula KR. Epidemic zone of COVID-19 from social media using hypergraph with weighting factor (HWF). JOURNAL OF SUPERCOMPUTING. 2021;

99. Li Z, Li X, Porter D, Zhang J, Jiang Y, Olatosi B, et al. Monitoring the Spatial Spread of COVID-19 and Effectiveness of Control Measures Through Human Movement Data: Proposal for a Predictive Model Using Big Data Analytics. JMIR RESEARCH PROTOCOLS. 2020 Dec;9(12).

100. Antoniou V, Vassilakis E, Hatzaki M. Is Crowdsourcing a Reliable Method for Mass Data Acquisition? The Case of COVID-19 Spread in Greece During Spring 2020. ISPRS INTERNATIONAL JOURNAL OF GEO-INFORMATION. 2020 Oct;9(10).

101. Issa E, Tsou M-H, Nara A, Spitzberg B. Understanding the spatio-temporal characteristics of Twitter data with geotagged and non-geotagged content: two case studies with the topic of flu and Ted (movie). Annals of Gis. 2017;23(3):219–35.

102. Stevens R, Bonett S, Bannon J, Chittamuru D, Slaff B, Browne SK, et al. Association Between HIV-Related Tweets and HIV Incidence in the United States: Infodemiology Study. JOURNAL OF MEDICAL INTERNET RESEARCH. 2020 Jun 24;22(6).

103. Zhang Y, Wu W, Wang Q, Su F. A geo-event-based geospatial information service: A case study of typhoon hazard. Sustainability (Switzerland) [Internet]. 2017;9(4). Available from: http://dx.doi.org/10.3390/su9040534

104. Rosser JF, Leibovici DG, Jackson MJ. Rapid flood inundation mapping using social media, remote sensing and topographic data. Natural Hazards. 2017 May;87(1):103–20.

105. Feng Y, Brenner C, Sester M. Flood severity mapping from Volunteered Geographic Information by interpreting water level from images containing people: A case study of Hurricane Harvey. ISPRS Journal of Photogrammetry and Remote Sensing. 2020;169:301–19.

106. Ghermandi A. Analysis of intensity and spatial patterns of public use in natural treatment systems using geotagged photos from social media. Water Research. 2016 Nov 15;105:297–304.

107. Domenech A, Mohino I, Moya-Gomez B. Using Flickr Geotagged Photos to Estimate Visitor Trajectories in World Heritage Cities. ISPRS INTERNATIONAL JOURNAL OF GEO-INFORMATION. 2020 Nov;9(11).

108. Wong DWS, Huang Q. “Voting with Their Feet”: Delineating the Sphere of Influence Using Social Media Data. Isprs International Journal of Geo-Information. 2017 Nov;6(11):325.

109. Toepke S. Structure Occupancy Curve Generation using Geospatially Enabled Social Media Data. Rocha JG, Grueau C, editors. 2016.

110. Haidery SA, Ullah H, Khan NU, Fatima K, Rizvi SS, Kwon SJ. Role of Big Data in the Development of Smart City by Analyzing the Density of Residents in Shanghai. ELECTRONICS. 2020 May;9(5).

111. Hoepken W, Mueller M, Fuchs M, Lexhagen M. Flickr data for analysing tourists’ spatial behaviour and movement patterns A comparison of clustering techniques. JOURNAL OF HOSPITALITY AND TOURISM TECHNOLOGY. 2020 Feb 26;11(1):69–82.

112. Jing C, Dong M, Du M, Zhu Y, Fu J. Fine-Grained Spatiotemporal Dynamics of Inbound Tourists Based on Geotagged Photos: A Case Study in Beijing, China. IEEE ACCESS. 2020;8:28735–45.

113. Hasnat MM, Hasan S. Identifying tourists and analyzing spatial patterns of their destinations from location-based social media data. Transportation Research Part C: Emerging Technologies. 2018;96:38–54.

114. Kovacs Z, Vida G, Elekes abel, Kovalcsik T. Combining social media and mobile positioning data in the analysis of tourist flows: A case study from Szeged, Hungary. Sustainability (Switzerland) [Internet]. 2021;13(5). Available from: http://dx.doi.org/10.3390/su13052926

115. Wang L-C, Yan C-X, Jing-Wang. The Research on Spatial-temporal Characteristics of Tourist Flow in Lanzhou Based on Sina Microblog Big Data. 2016.

116. Duan J, Zhai W, Cheng C. Crowd detection in mass gatherings based on social media data: A case study of the 2014 shanghai new years eve stampede. International Journal of Environmental Research and Public Health. 2020;17(22):1–14.

117. Miao R, Wang Y, Li S. Analyzing urban spatial patterns and functional zones using sina weibo poi data: A case study of Beijing. Sustainability (Switzerland). 2021;13(2):1–15.

118. Hu Y, Gao S, Janowicz K, Yu B, Li W, Prasad S. Extracting and understanding urban areas of interest using geotagged photos. Computers, Environment and Urban Systems. 2015;54:240–54.

119. Belcastro L, Kechadi MT, Marozzo F, Pastore L, Talia D, Trunfio P. Parallel extraction of Regions-of-Interest from social media data. CONCURRENCY AND COMPUTATION-PRACTICE & EXPERIENCE. 2021 Apr 25;33(8, SI).

120. Liu Q, Ullah H, Wan W, Peng Z, Hou L, Qu T, et al. Analysis of Green Spaces by Utilizing Big Data to Support Smart Cities and Environment: A Case Study About the City Center of Shanghai. ISPRS INTERNATIONAL JOURNAL OF GEO-INFORMATION. 2020 Jun;9(6).

121. Tsou M-H, Yang J-A, Lusher D, Han S, Spitzberg B, Gawron JM, et al. Mapping social activities and concepts with social media (Twitter) and web search engines (Yahoo and Bing): a case study in 2012 US Presidential Election. Cartography and Geographic Information Science. 2013;40(4):337–48.

122. Iranmanesh A, Comert NZ, Hoskara SO. Reading urban land use through spatio-temporal and content analysis of geotagged Twitter data. GEOJOURNAL. 2021;

123. Wei X, Yao XA. Constructing and analyzing spatial-social networks from location-based social media data. CARTOGRAPHY AND GEOGRAPHIC INFORMATION SCIENCE. 2021 May 4;48(3):258–74.

124. Li S, Liu Z, Li Y. Temporal and spatial evolution of online public sentiment on emergencies. INFORMATION PROCESSING & MANAGEMENT. 2020 Mar;57(2).

125. Baeza JL, Bley J, Hartkopf K, Niggemann M, Arias J, Wiedenhofer A. Evaluating cultural impact in discursive space through digital footprints. Sustainability (Switzerland) [Internet]. 2021;13(7). Available from: http://dx.doi.org/10.3390/su13074043

126. Jiang H, Qiang M, Lin P. Assessment of online public opinions on large infrastructure projects: A case study of the Three Gorges Project in China. Environmental Impact Assessment Review. 2016 Nov;61:38–51.

127. Kovacs-Gyoeri A, Ristea A, Kolcsar R, Resch B, Crivellari A, Blaschke T. Beyond Spatial Proximity-Classifying Parks and Their Visitors in London Based on Spatiotemporal and Sentiment Analysis of Twitter Data. Isprs International Journal of Geo-Information. 2018 Sep;7(9):378.

128. Seltzer EK, Horst-Martz E, Lu M, Merchant RM. Public sentiment and discourse about Zika virus on Instagram. Public Health. 2017 Sep;150:170–5.

129. Nguyen TT, Meng H-W, Sandeepe S, McCullough M, Yu W, Lau Y, et al. Twitter-derived measures of sentiment towards minorities (2015-2016) and associations with low birth weight and preterm birth in the United States. Computers in Human Behavior. 2018 Dec;89:308–15.

130. Camacho K, Portelli R, Shortridge A, Takahashi B. Sentiment mapping: point pattern analysis of sentiment classified Twitter data. Cartography and Geographic Information Science. 2021;48(3):241–57.

131. Zivanovic S, Martinez J, Verplanke J. Capturing and mapping quality of life using Twitter data. GeoJournal. 2020;85(1):237–55.

132. Arif F, Dulhare UN. A Machine Learning Based Approach for Opinion Mining on Social Network Data. In: Satapathy SC, Bhateja V, Raju KS, Janakiramaiah B, editors. Computer Communication, Networking and Internet Security. 2017. p. 135–47.

133. Ghosh D, Guha R. What are we "tweeting" about obesity? Mapping tweets with topic modeling and Geographic Information System. Cartography and Geographic Information Science. 2013;40(2):90–102.

134. Lai J, Cheng T, Lansley G. Improved targeted outdoor advertising based on geotagged social media data. Annals of GIS. 2017;23(4):237–50.

135. Yao F, Wang Y. Tracking urban geo-topics based on dynamic topic model. COMPUTERS ENVIRONMENT AND URBAN SYSTEMS. 2020 Jan;79.

136. Nishimura T, Nishida K, Toda H, Sawada H. Social media knows what road it is: quantifying road characteristics with geo-tagged posts. Social Network Analysis and Mining. 2017 Nov 9;7(1):57.

137. Lock O, Pettit C. Social media as passive geo-participation in transportation planninghow effective are topic modeling & sentiment analysis in comparison with citizen surveys? Geo-Spatial Information Science. 2020;23(4):275–92.

138. Yan Y, Chen J, Wang Z. Mining public sentiments and perspectives from geotagged social media data for appraising the post-earthquake recovery of tourism destinations. Applied Geography [Internet]. 2020;123. Available from: http://dx.doi.org/10.1016/j.apgeog.2020.102306

139. Blei DM, Ng AY, Jordan MI. Latent Dirichlet Allocation. :30.

140. Di Minin E, Fink C, Hiippala T, Tenkanen H. A framework for investigating illegal wildlife trade on social media with machine learning. Conserv Biol. 2019 Feb;33(1):210–3.

141. Wang Y, Fu X, Jiang W, Wang T, Tsou M-H, Ye X. Inferring urban air quality based on social media. Computers, Environment and Urban Systems. 2017;66:110–6.

142. Bhalla P, Bhattacharya P, Areendran G, Raj K. Ecotourism spatio-temporal models to identify visitation patterns across the Indian Himalayan Region. GEOJOURNAL. 2020;

143. Mitas L, Mitasova H. Spatial Interpolation. In: Geographical Information Systems: Principles, Techniques, Management and Applications. Wiley; 1999. p. 481–92.

144. Liu X, He J, Yao Y, Zhang J, Liang H, Wang H, et al. Classifying urban land use by integrating remote sensing and social media data. International Journal of Geographical Information Science. 2017;31(8):1675–96.

145. Shao H, Zhang Y, Li W. Extraction and analysis of city's tourism districts based on social media data. Computers, Environment and Urban Systems. 2017;65:66–78.

146. Delaney KP, Kramer MR, Waller LA, Flanders WD, Sullivan PS. Using a geolocation social networking application to calculate the population density of sex-seeking gay men for research and prevention services. J Med Internet Res. 2014 Nov 18;16(11):e249.

147. McKenzie G, Janowicz K, Gao S, Yang J-A, Hu Y. POI Pulse: A multi-granular, semantic signature-based information observatory for the interactive visualization of big geosocial data. Cartographica. 2015;50(2):71–85.

148. Jung J-K. Code clouds: Qualitative geovisualization of geotweets. Canadian Geographer-Geographe Canadien. 2015 SPR;59(1):52–68.

149. Dunkel A. Visualizing the perceived environment using crowdsourced photo geodata. Landscape and Urban Planning. 2015;142:173–86.

150. Xing H, Meng Y, Shi Y. A dynamic human activity-driven model for mixed land use evaluation using social media data. Transactions in GIS. 2018;22(5):1130–51.

151. Williamson W, Ruming K. Using Social Network Analysis to Visualize the Social-Media Networks of Community Groups: Two Case Studies from Sydney. Journal of Urban Technology. 2016;23(3):69–89.

152. Wang Z, Ye X, Lee J, Chang X, Liu H, Li Q. A spatial econometric modeling of online social interactions using microblogs. Computers, Environment and Urban Systems. 2018;70:53–8.

153. Huang Q, Cervone G, Zhang G. A cloud-enabled automatic disaster analysis system of multi-sourced data streams: An example synthesizing social media, remote sensing and Wikipedia data. Computers, Environment and Urban Systems. 2017;66:23–37.

154. Chen X, Elmes G, Ye X, Chang J. Implementing a real-time Twitter-based system for resource dispatch in disaster management. Geojournal. 2016 Dec;81(6):863–73.

155. Shook E, Turner VK. The socio-environmental data explorer (SEDE): a social media-enhanced decision support system to explore risk perception to hazard events. Cartography and Geographic Information Science. 2016 Nov;43(5):427–41.

156. Werts JD, Mikhailova EA, Post CJ, Sharp JL. An Integrated WebGIS Framework for Volunteered Geographic Information and Social Media in Soil and Water Conservation. Environmental Management. 2012 Apr;49(4):816–32.

157. Pezanowski S, MacEachren AM, Savelyev A, Robinson AC. SensePlace3: a geovisual framework to analyze place-time-attribute information in social media. Cartography and Geographic Information Science. 2018;45(5):420–37.

158. Xu C, Wong DW, Yang C. Evaluating the “geographical awareness” of individuals: an exploratory analysis of twitter data. Cartography and Geographic Information Science. 2013;40(2):103–15.

159. Liao C, Brown D, Fei D, Long X, Chen D, Che S. Big data-enabled social sensing in spatial analysis: Potentials and pitfalls. Transactions in Gis. 2018 Dec;22(6):1351–71.

160. Reich J. Rebooting MOOC Research. Science. 2015 Jan 2;347(6217):34–5.

161. Kwan M-P. Algorithmic Geographies: Big Data, Algorithmic Uncertainty, and the Production of Geographic Knowledge. In: Kwan M-P, Schwanen T, editors. Geographies of Mobility [Internet]. 1st ed. Routledge; 2018 [cited 2021 May 27]. p. 32–40. Available from:

https://www.taylorfrancis.com/books/9781351969819/chapters/10.4324/9781315266336-4

162. Martin ME, Schuurman N. Social Media Big Data Acquisition and Analysis for Qualitative GIScience: Challenges and Opportunities. ANNALS OF THE AMERICAN ASSOCIATION OF GEOGRAPHERS. 2020 Sep 2;110(5):1335–52.

163. Bazzaz Abkenar S, Haghi Kashani M, Mahdipour E, Jameii SM. Big data analytics meets social media: A systematic review of techniques, open issues, and future directions. Telematics and Informatics. 2021 Mar 1;57:101517.

Chapter 3 references

1. Red Cross. Haiti Earthquake: 10-Year Update from the Red Cross [Internet]. 2020 [cited 2021 Jan 8]. Available from: https://www.redcross.org/about-us/news-and-events/news/2020/haiti-earthquake-10-year-update-from-the-red-cross.html

2. Jul 30 CA· P, July 30 2020 8:30 AM ET | Last Updated:, 2020. The 2003 concert that rocked Toronto after SARS | CBC Archives [Internet]. CBC. 2020 [cited 2021 Jan 8]. Available from: https://www.cbc.ca/archives/the-2003-concert-that-rocked-toronto-after-sars-1.5650768

3. About GoFundMe [Internet]. GoFundMe. [cited 2020 Dec 6]. Available from: https://www.gofundme.com/c/about-us

4. Harris A. GoFundMe keeps gobbling up competitors, says it's "very good for the market" [Internet]. Fast Company. 2018 [cited 2020 Dec 6]. Available from: https://www.fastcompany.com/40554199/gofundme-keeps-gobbling-up-competitors-says-its-very-good-for-the-market

5. Palad V, Snyder J. "We don't want him worrying about how he will pay to save his life": Using medical crowdfunding to explore lived experiences with addiction services in Canada. Int J Drug Policy. 2019 Mar 1;65:73–7.

6. Snyder J, Turner L, Crooks VA. Crowdfunding for Unproven Stem Cell–Based Interventions. JAMA. 2018 May 8;319(18):1935–6.

7. Farnel M. Kickstarting trans*: The crowdfunding of gender/sexual reassignment surgeries. New Media Soc. 2015 Feb 1;17(2):215–30.

8. Vox F, Folkers KM, Turi A, Caplan AL. Medical Crowdfunding for Scientifically Unsupported or Potentially Dangerous Treatments. JAMA. 2018 Oct 23;320(16):1705–6.

9. Majumdar A, Bose I. My words for your pizza: An analysis of persuasive narratives in online crowdfunding. Inf Manage. 2018 Sep;55(6):781–94.

10. Ren J, Raghupathi V, Raghupathi W. Understanding the Dimensions of Medical Crowdfunding: A Visual Analytics Approach. J Med Internet Res. 2020;22(7):e18813.

11. Kenworthy N, Dong Z, Montgomery A, Fuller E, Berliner L. A cross-sectional study of social inequities in medical crowdfunding campaigns in the United States. PLOS ONE. 2020 Mar 5;15(3):e0229760.

12. Berliner LS, Kenworthy NJ. Producing a worthy illness: Personal crowdfunding amidst financial crisis. Soc Sci Med. 2017 Aug 1;187:233–42.

13. Lee S, Lehdonvirta V. New Digital Safety Net or Just More ‘Friendfunding’? Institutional Analysis of Medical Crowdfunding in the United States [Internet]. Rochester, NY: Social Science Research Network; 2020 Mar [cited 2020 Aug 26]. Report No.: ID 3559901. Available from: https://papers.ssrn.com/abstract=3559901

14. Snyder J. Crowdfunding FOR MEDICAL CARE: Ethical Issues in an Emerging Health Care Funding Practice. Hastings Cent Rep. 2016;46(6):36–42.

15. Duynhoven A van, Lee A, Michel R, Snyder J, Crooks V, Chow-White P, et al. Spatially exploring the intersection of socioeconomic status and Canadian cancer-related medical crowdfunding campaigns. BMJ Open. 2019 Jun 1;9(6):e026365.

16. GoFundMe. The Data Behind Donations During the COVID-19 Pandemic [Internet]. Medium. 2020 [cited 2020 Dec 6]. Available from: https://medium.com/gofundme-stories/the-data-behind-donations-during-the-covid-19-pandemic-c40e0f690bfa

17. Canada Emergency Response Benefit statistics - Canada.ca [Internet]. [cited 2020 Sep 13]. Available from: https://www.canada.ca/en/services/benefits/ei/claims-report.html

18. Government of Canada SC. The Daily — Labour Force Survey, July 2020 [Internet]. 2020 [cited 2020 Dec 6]. Available from: https://www150.statcan.gc.ca/n1/daily-quotidien/200807/dq200807a-eng.htm

19. Elmer G, Ward-Kimola S, Burton AG. Crowdfunding during COVID-19: An international comparison of online fundraising. First Monday [Internet]. 2020 Oct 16 [cited 2020 Dec 6]; Available from: https://journals.uic.edu/ojs/index.php/fm/article/view/10869

20. Rajwa P, Hopen P, Mu L, Paradysz A, Wojnarowicz J, Gross CP, et al. Online Crowdfunding Response to Coronavirus Disease 2019. J Gen Intern Med. 2020 May 27;1–3.

21. Shortage of personal protective equipment endangering health workers worldwide [Internet]. 2020 [cited 2021 Jan 6]. Available from: https://www.who.int/news/item/03-03-2020-shortage-of-personal-protective-equipment-endangering-health-workers-worldwide

22. Foran P. 1 of 7 stores may not survive COVID-19 pandemic, Canadian business group says | CTV News [Internet]. 2020 [cited 2021 Jan 6]. Available from: https://toronto.ctvnews.ca/1-of-7-stores-may-not-survive-covid-19-pandemic-canadian-business-group-says-1.5155104

23. Slaughter G. New closures will be “fatal” for thousands of small businesses, advocate warns [Internet]. Coronavirus. 2020 [cited 2021 Jan 6]. Available from: https://www.ctvnews.ca/health/coronavirus/new-closures-will-be-fatal-for-thousands-of-small-businesses-advocate-warns-1.5184987