A review of the use of geosocial media data in agent-based models for studying urban systems

Richard Wen and Songnian Li

Department of Civil Engineering, Ryerson University, Toronto, Canada

ABSTRACT
Since the rapid growth of urban populations, the study of urban systems has gained considerable attention from researchers, decision makers, governments, and organizations. Urban systems are complex and dynamic such that they produce emergent patterns such as self-organization and nonlinearity. Agent-based modelling presents an approach to simulating and abstracting urban systems to reveal and study emergent patterns from urban-related entities. However, agent-based models are difficult to effectively optimize and validate without high quality real-world data. Geosocial media data provides agent-based models with location-enabled data at high volumes and frequencies. Integrating agent-based models with geosocial media data presents opportunities in advancing and developing studies in urban systems. This paper provides a general overview of concepts, review of recent applications, and discussion of challenges and opportunities in the context of using geosocial media data in agent-based models for urban systems. We argue that ABMs focused on studying urban systems can benefit greatly from geosocial media data, given that research moves towards standard guidelines that enable the comparison and effective use of ABMs, and geosocial media data under appropriate circumstances and applications.

1. Introduction
In 2018, 55% of the world’s population lived in urban areas with a projected increase to 68% in 2050 (United Nations, 2018). The rapid growth of urban populations has brought attention to the study of urban areas and their social, environmental, and physical interactions. Urban areas contain the interactions of entities such as people, buildings, and the environment, which can form a complex whole such as a city or a country. This complex whole can be thought of as an urban system that is driven by the interactions of people and other related entities. One of these interactions can be described as social communication, which has led to major political decisions and urban planning processes that have influenced urban systems. These influences, for example, include the alteration of the climate, the change in quality of life for populations, and the transformation of landscapes on earth. Social communication is one of the drivers of urban systems, and has been difficult to describe and quantify due to scarcity of data (Batty, 2013; Bettencourt &

CONTACT Songnian Li snli@ryerson.ca Department of Civil Engineering, Ryerson University, 350 Victoria St, Toronto, ON M5B 2K3
© 2020 The Author(s). Published by Taylor & Francis Group and Science Press on behalf of the International Society for Digital Earth, supported by the CASEarth Strategic Priority Research Programme. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
West, 2010). Every individual’s choices, actions, behaviours, interactions, and decisions contribute to the overall system in a city, in which seemingly chaotic decisions become organized at the city scale (Netto, Meirelles, & Ribeiro, 2017). The capture of social-related data has since been improved by technologies such as the internet, portable devices, and particularly social media (França et al., 2016).

Social media platforms have created a way for billions of people to communicate across the Earth in near real-time (Macnamara & Zerfass, 2012; Weller, 2015). This widespread use of social media platforms has made it possible to collect massive volumes of user-generated data at highly frequent time intervals. Social media data have since gained spatial properties with the advances in location acquisition technology for portable devices. For example, mobile phones with Global Navigation Satellite System (GNSS) (e.g. Global Positioning System (GPS)) allow users of the social media platform Facebook to immediately post content related to their current location known as check-ins (Kim, 2016). Similarly, Twitter users can post 280-character content known as tweets at specific locations on the Earth. These check-ins and location-enabled Tweets are considered geosocial media data, i.e., location-based social media data, where a record is created by relating user-generated content to entities in the real-world such as buildings and places, or more precisely geographical locations defined by accurate latitude and longitude coordinates. Geosocial media data has made it possible to connect the online and offline lives of users, providing real-time geospatial data for advancing studies of social infrastructure and interactions in urban systems.

Statistical approaches, such as measures of similarity or probability, are commonly used to create location-based models from geosocial media data. A few examples include: location recommendations and comparisons (Ostermann et al., 2015; Zhang & Chow, 2015), human mobility and activity analysis (Hristova, Williams, Musolesi, Panzarasa, & Mascolo, 2016; Huang & Li, 2016; Huang, Li, Liu, & Ban, 2015), event detection and prediction (Hua, Zhao, Chen, Lu, & Ramakrishnan, 2016; Rannies et al., 2016; Xu, Li, & Huang, 2020; Xu, Li, & Wen, 2018), environmental monitoring (Jiang, Wang, Tsou, & Fu, 2016), and geographic profiling (Bao, Lian, Zhang, & Yuan, 2016). However, statistical approaches alone do not often offer the flexibility to explore the complexity and dynamics in urban systems, which involve ill-defined, non-linear, and interactive processes that lead to emergent patterns (Wu, 2002). Emergent patterns are difficult to reveal, and have often been studied using agent-based modelling.

The simulative nature and bottom-up approach of agent-based modelling has made it an effective technique for studying emergent patterns in urban systems. An Agent-based Model (ABM) is built at the individual level from interacting autonomous entities, known as agents that are simulated within a computational environment to create a system of agents. ABMs abstract the interactions of urban-related entities, such as people and buildings, to reveal emergent patterns at varying levels of detail, such as common traffic jam behaviors and urban flooding spread at the neighborhood or city level during selected hours or days. ABMs, however, are difficult to verify and optimize when applied to complex real-world problems such as simulating traffic conditions or pedestrian movement. It is common in many studies to integrate statistical approaches and available data for verifying and optimizing resulting ABMs (Bazzan & Klügl, 2013; Berger, 2001; Borshchev & Filippov, 2004; Ge & Kremers, 2015). For example, an ABM of traffic may be used to optimize the road development based on travel speed by measuring vehicle
speeds and simulating agents as vehicles. Real-world traffic data can then be compared to the simulations to verify the ABM. Similarly, geosocial media data can be integrated with statistical approaches to optimize and verify ABMs, which provides valuable information on social infrastructure and interactions among agents. The integration of numerical approaches and geosocial media data have improved ABMs to provide highly detailed simulations of urban systems in space and time. This improvement has enabled researchers to better study urban systems with ABMs.

This paper seeks to provide a review of recent agent-based modelling applications that incorporate geosocial media data in the context of urban systems. The review is done by providing background, reviewing recent research applications, and discussing benefits and limitations related to urban systems, geosocial media data, and ABMs. Research articles were selected using keyword searches (not case-sensitive) based on the combination of ABMs, urban systems, social media, and geospatial analysis, where each article must at least one of the words from each of the 4 categories:

1. **ABMs**: “agent based models”, “agent based modelling”, “ABM”
2. **Urban Systems**: “urban”, “urban system”, “urban areas”, “urbanization”, “city”, “cities”, “metropolitan”, “metropolis”, “mega-region”
3. **Social Media**: “social media”, “geosocial”
4. **Geospatial**: “location”, “geospatial”, “gis”, “geographic”, “geolocation”, “geo”, “spatial”, “spatio-temporal”

For example, if an article only had “agent based models”, and “urban system”, satisfying only two of the four categories, then it would not be included for review. However, if an article had “agent based models” (ABMs category), “urban areas” (Urban Systems category), “social media” (Social Media category), and “gis” (Geospatial category), then it would be considered in the literature review due to having at least one word from each category. The review articles were found using online databases from: Google Scholar, Web of Science, Scopus, PubMed, IEEE Xplore, and ScienceDirect. A total of 14 articles were also manually selected from the search results for reviewing recent applications that involved ABMs and geosocial media data in the context of urban systems from the years 2011 to 2019. These papers were inspected for particular applications of ABMs or geosocial media data to fields or domains, such as health, natural disaster, and traffic engineering.

Section 2.0 provides an overview of ABMs, urban systems, and geosocial media. Section 3.0 summarizes and organizes a short review of recent agent-based modelling applications for urban systems using selected geosocial media data. Section 4.0 discusses the benefits, limitations, challenges and opportunities of ABMs relative to urban systems and geosocial media. Section 5.0 concludes the paper with a summary and provides future prospects and recommendations.

2. **Background**

It is important to understand the general concepts of ABMs, urban systems, and geosocial media before looking into how ABMs and geosocial media data can be applied for urban systems. The terms ABM and urban system are loosely defined with various perspectives. The concept of agents and environments in ABMs span and combine multiple fields of
study such as sociology, geography, biology, engineering, computer science, and mathematics (Markovic & Zornic, 2016). An agent can represent a person, a structure, a molecule, a city, or any entity that can be defined or conceptualized. An agent can then be designed to inherit abilities such as learning, adapting, changing, and interacting. The environment which contains the agent is also flexible. An environment can represent surroundings such as structures, forests, rooms, or hypothetical surroundings that do not exist yet. Urban systems, like ABMs, also span multiple fields of study. Urban systems contain physical, environmental, and social infrastructures that can be studied from different perspectives that often require multidisciplinary knowledge (Coffey, 1998). It is also important to understand the general structure and availability of popular sources of geosocial media data, which enhances ABMs for urban systems. This section seeks to provide a generalized perspective of ABMs, urban systems, and geosocial media through research article reviews, such that it can be applied to various fields of study.

2.1. Agent-based models

An ABM consists of simulated entities called agents that interact inside of an environment. Wooldridge and Jennings (1995) specified a weak notion of describing agents to have the following properties:

- **Autonomy.** Agents act individually without direct control from humans or external forces.
- **Social ability.** Agents interact among each other with a communication language.
- **Reactivity.** Agents are aware of and respond to their environment.
- **Proactiveness.** Agents are driven to satisfy their objectives.

The construction of an ABM requires that the following be defined (Crooks & Heppenstall, 2012):

- **Agents.** The characteristics and properties of the different entities to be created.
- **Environments.** The surroundings shared among all or select agents.
- **Interactions.** The processes that occur between agents with other agents and their shared environment.

Ideally, ABMs should be optimized for lower computational complexity and response times, and higher modularity and objective achievement quality (Barbati, Bruno, & Genovese, 2012). ABMs are appropriate for the exploration of systems as systems consist of interacting entities that can be modeled in a bottom-up approach to discover emergent patterns.

Emergent patterns occur in complex systems when entities, such as people and buildings, interact to form a larger entity, such as a city, with properties that do not exist in the individual entities (O’Connor & Wong, 2002). A notable example of emergent patterns is Conway’s Game of Life, which involves a grid of cells with rules that define the interactions between cells to mimic simplistic behaviors of population growth and decay (Gardner, 1970). The rules change the state of a cell to alive or dead based on the states of its eight adjacent cells. These cell interactions eventually reveal emergent patterns in the
overall grid by applying the rules to every cell repetitively. This repetitive application of rules is considered a simulation, an approximation or abstraction of a real-world system (Davis, Eisenhardt, & Bigham, 2007; Ingalls, 2008; Law, 2008). Simulations in ABMs enable the exploration of systems with emergent patterns by allowing users to develop and test theories that could be used in the real-world (Davis et al., 2007).

At a software level, there are different tools for a variety of programming languages to help create ABMs that range from small-scale (hundreds) to extreme-scale (millions of agents), which can require a cluster of computers to run the models. For example, AgentScript in Java is a simple tool built for small-scale ABMs, while Swarm in Objective-C/Java is a much more complex tool for extreme-scale ABMs that leverage computing clusters to run models at scale. Abar, Theodoropoulos, Lemarinier, and O’Hare (2017) has a comprehensive list of ABM software tools, and an analysis of their advantages and disadvantages when choosing a software framework or platform for ABMs.

### 2.2. Urban systems

Urban systems are complex and dynamic with interacting social, ecological, and technical entities. These entities can be seen as components or parts that form a whole system, otherwise known as a complex adaptive system, in which emergent properties exist (Bretagnolle, Daudé, & Pumain, 2006; Mitchell, 2009). Urban systems are often characterized by non-linear self-organization and adaptation which emerge from the interactions of urban-related entities (Levy, Martens, & van der Heijden, 2016; Storper, Van Marrewijk, & Van Oort, 2012). A few examples of urban-related entities include, but are not limited to, particles, people, animals, man-made structures, vegetation, atmospheres, landscapes, districts, cities, and countries. Urban systems may also be formed from entities that are smaller systems, which can be described as a system of systems (Cocks, 2006; Johnson & Hernandez, 2016). For example, a system such as a city can contain multiple districts, which are viewed as smaller systems, but seen as entities in the context of the larger city system. A common trait in urban systems is that people are the core entities that drive the system. As urban populations increase, urban systems tend to create hierarchical organizations, such as cities that contain neighborhoods and neighborhoods that contain smaller communities, which also become central places that provide services to surrounding areas (Fujita, Krugman, & Mori, 1999; Fujita & Mori, 1997). The development and characteristics of urban systems are influenced by the interacting urban entities that form the system, which exhibits emergent properties such as nonlinearity and self-organization.

### 2.3. Geosocial media data

Geosocial media data has become increasingly available as portable devices, such as smartphones and tablets, and social media platforms are incorporating location acquisition technologies. By the end of 2019, Facebook, a platform offering social networking services such as content posting and instant messaging, had an average of ~1.59 billion monthly active users where ~90% were mobile users (Facebook Inc, 2019a). By 2016, Twitter, a platform centered around online communication with 280-character user posts called tweets, had an average of ~330 million monthly active users where ~80% were mobile users (Twitter Inc, 2019a). Examples of other popular social media platforms
include, but are not limited to, Sina Weibo, Tencent, Flickr, Instagram, Snapchat, Foursquare, and Baidu (Ebrahimpour, Wan, Velázquez García, Cervantes, & Hou, 2020; Qi, Li, Wang, & Gao, 2019; Wu, Zhi, Sui, & Liu, 2014). Geosocial media data is collected in massive quantities and near real-time at internet-accessible locations across the earth.

Geosocial media data often consists of points, taken from GPS coordinates of mobile devices, with text representing user posts. The data is often free with some restrictions, where commercial options offer less flexible restrictions. Access is usually provided by social media platform companies using an Application Programming Interface (API) to query and request data (Rama & Kak, 2015). For example, Twitter provided a Representational State Transfer (REST) API that permits users with developer accounts to read and write Twitter data for free with some restrictions such as request and Tweet limits (Twitter Inc, 2019b). Facebook provided a Hypertext Transfer Protocol (HTTP) based API where users with developer accounts can read and write content by sending requests to a graph-based structure consisting of nodes, edges, and fields (Facebook Inc, 2019b). The structure and access to geosocial media data provided by social media platform companies vary and may change over time, which requires user adjustments when combining various incompatible data formats from different sources.

Geosocial media data or social media data in general are often controlled by the platforms in which the data is generated. This could pose some data representation issues stemming from the proportion of data or data fields released for access that each platform allows. For example, although Twitter streaming data can be accessed through their API, only percentage of that data is available for free or publicly accessible. This can lead to these small samples of data not being representative of the whole – depending on what the use of the data is. However, in some cases, the small samples of data are enough for particular applications such as event detection due to the volume and velocity of data, or hold similar results to using the entire dataset (Li, Shah, Thomas, Anderson, & Liu, 2016b).

3. Applications

ABMs have been combined with statistical approaches and geosocial media data to simulate and optimize urban system models that involve social interactions. ABMs have been used particularly when emergent phenomena exist or in cases where high-quality real-world data is unobtainable or lacking. The rapid advancement and increasing availability of geosocial media has made it important to consider the integration of ABMs with geosocial media data, where the studies reviewed in this paper have focused on the integration of large amounts of locatable textual information (user posts) extractable from geosocial media data. Geosocial media data provides highly detailed empirical data for validation and calibration techniques for the justification and practicality of ABMs for real-world applications (Darvishi & Ahmadi, 2014). The location aspect of geosocial media data can be incorporated to include space dynamics, which provides improved semantic details (from the social media text, video, and sound) at the local geographic level. Thus, agent-based modelling of social infrastructures in urban systems can be improved by incorporating geosocial media data to calibrate and validate ABMs in time and space (Heppenstall, Malleson, & Crooks, 2016) as seen in Table 1. The reviewed ABM studies are organized into several summarized sub-sections that group studies into broader
Table 1. Geosocial media data usage in ABMs and urban systems.

| Application                  | Data Source                          | Use                                                                                                                                                                                                 | Authors                        |
|------------------------------|--------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------|
| Human Mobility               | Flickr                               | Calibrate models of human mobility and interactions in urban space                                                                                                                              | Prager & Wiegand, 2014         |
| Human Mobility               | Aggregate Chinese social media check-ins | Verify model of human movement and activity in the city                                                                                                                                          | Wu et al., 2014                |
| Human Mobility               | Twitter                              | Integrate sentiment and mobility to enhance models and analyses                                                                                                                             | Wang & Taylor, 2016            |
| Human Mobility               | Twitter                              | Improve accuracy of model to predict citywide human movement patterns                                                                                                                          | Goh, Koh, & Zhang, 2019        |
| Population Health            | Twitter                              | Surveillance and near real-time models of Dengue of cities                                                                                                                                 | Gomide et al., 2011            |
| Disaster/                     | Twitter                              | Fit models for social information diffusion with different diffusion patterns                                                                                                                      | Rand, Hermann, Schein, and     |
| Emergency                    |                                      |                                                                                                                                                                                                 | Vodopivec (2015)               |
| Disaster/                    | Twitter                              | Enhance models of flood simulations to provide greater simulated detail                                                                                                                         | Smith, Liang, James, and Lin (2017) |
| Emergency                    |                                      |                                                                                                                                                                                                 |                                 |
| Land Use/                    | Twitter                              | Correlating demographic and socioeconomic factors to social media for enhanced modelling and analysis                                                                                          | Jiang, Li, and Ye (2019)       |
| Sustainability               |                                      |                                                                                                                                                                                                 |                                 |
| Land Use/                    | Twitter, Tencent, Flickr, Sina Weibo, Foursquare, Baidu | Enhance urban planning and land monitoring models                                                                                                                                                | Qi et al. (2019)               |
applications. The reviewed studies demonstrated that ABMs are particularly powerful when they are combined with other approaches and when geosocial media data were used to validate/calibrate/optimize model results.

3.1. Human mobility

Human mobility refers to the movement patterns of people. Prager and Wiegand (2014) used a random walk ABM and geosocial media data from a platform called Flickr to explore factors influencing the use of space in urbanized areas. Wu et al. (2014) used geosocial check-in data and an ABM to reproduce human mobility patterns to verify a movement and activity-based model. Wang and Taylor (2016) used Twitter data to study human mobility patterns during natural disasters, which were found to be resilient for particular disasters, but less resilient when more powerful natural disasters occur. Goh et al. (2019) used Twitter data to predict the citywide movement of crowds with a deep-neural-network-based approach – noting a slight improvement in prediction accuracy when incorporating Twitter data. Human mobility patterns differ based on many influential factors such as time, place, situation, weather, and the surrounding environment. The understanding of human mobility supplements many other important urban systems applications such as disease spread, urban planning, traffic management, and market forecasting.

3.2. Population health

Health refers to the well-being of humans in the context of physical, mental, and social well-being states (Huber et al., 2011). Frias-Martinez, Williamson, and Frias-Martinez (2011) used ABMs with cell phone records defining agent behavior to simulate H1N1 virus spread for the support of government decisions, which enabled time and space dynamics to be modeled as opposed to using traditional census or survey data. Gomide et al. (2011) used geo-tagged Tweets for the surveillance of Dengue in Brazil based on the dimensions of volume, location, time, and public perception. Luo, Gao, and Cassels (2018) implemented spatially-explicit agent-based epidemic models to simulate geosocial interactions of disease spread through population movement. This enabled improved decisions for effective influenza vaccination strategies based on identified containment areas and emerging interaction patterns. Luo, Gao, and Cassels (2018), however, mentioned that the addition of social media data would provide improved estimates of origin and destination behaviors of populations in urban areas. Geosocial media data provides valuable insights into population health as the communication between people can be analyzed and recorded at greater frequencies, volumes, and locations than traditional forms of data such as surveys.

3.3. Disaster and emergency

ABMs have been used with geosocial media data, particularly Twitter, to manage and mitigate disasters and emergencies. Rand et al. (2015) used ABM and social media data from Twitter to explore the diffusion of information in time during crisis events, where models were fit similarly, but were shown to produce different diffusion patterns. Durak
and Till (2015) used social agents to simulate artificial data of urban population evacuation to optimize a genetic algorithm for traffic light operation in the case of disasters. Haer, Botzen, and Aerts (2016) utilized an ABM to evaluate the effectiveness of social networks for flood risk communication, but emphasized that empirical data such as social media data are needed to calibrate the model. Smith et al. (2017) compared floods identified by Graphics Processing Unit (GPU) accelerated hydrodynamic modelling simulations to floods identified by geosocial media, and found that the results were relatively similar – noting that geosocial media, even with a small sample size, has the potential to enhance existing models by providing higher levels of detail in real-time. Geosocial media can create improved ABMs that are more representative of the current state of real-world emergencies and disasters by providing timely location-enabled data at low costs, which can improve, speed-up, or rival existing state-of-the-art models.

3.4. Land use and sustainability

Land use and sustainability have major impacts in the development and well-being of urban systems. Land use scenarios have been simulated using negotiation agents, having speaker and listener roles, that propose land use plans to understand the changes from decision making policies before they are implemented (Ghavami, Taleai, & Arentze, 2016). Zhang, Vorobeychik, Letchford, and Lakkaraju (2016) constructed data-driven agents by automatic learning of characteristics and behaviors from existing rooftop solar adoption data to forecast and optimize solar adoption decisions. In a study by Jiang et al. (2019), socioeconomic and demographic factors were correlated with social media data to understand the socioeconomic drivers and variations across different urbanized regions. This could be used to further understand public opinion and biases in land use planning and change. Many urban monitoring and land use analyses rely on remote sensed imagery data from satellites, which could be very expensive and do not provide fine enough temporal quality. Qi et al. (2019) examined low-cost methods that use semantic information in geosocial media data with remotely sensed imagery to enhance or enable urban planning and the monitoring of urban areas. Political and planning decisions that affect the land use and sustainability in urban systems can be supported by ABMs that incorporate low cost, timely, and large-scale geosocial media data to evaluate hypothetical scenarios.

4. Discussion

ABMs have been recently studied and applied to urban systems simulation. However, ABMs are often situational, where an understanding of the benefits and limitations of ABMs for urban systems is required for appropriate and justified use. In addition, ABMs pose challenges and opportunities that can lead to the development and advancement of urban systems research. This section seeks to acknowledge the appropriate use and development of ABMs for urban systems with geosocial media data.

4.1. Benefits and limitations

The benefits of ABMs are focused on their ability to represent complex and dynamic urban systems in time and space. ABMs are particularly useful when non-linear and heterogeneous
entities, which exhibit complex behaviors such as learning, adaptation, and self-organization, are involved (Bonabeau, 2002). Entity modularity is also another benefit of ABMs, where the model can be easily altered to adjust, include, and exclude urban entities. Finally, ABMs can be easily visualized to promote the inclusive understanding of the resulting models and emergent patterns without requiring expert knowledge of agent-based modelling (Hall & Virrantaus, 2016). ABMs offer the flexibility of incorporating a mixture of approaches from fields not limited to the ones traditionally seen in the study of urban systems.

The limitations of ABMs are centered on the validity and practicality in real-world urban systems applications. ABMs built to simulate real-world phenomena must often have a specific purpose and are commonly difficult to quantify and verify, in addition to being computationally intensive (Bonabeau, 2002). ABMs are also situational, which are often only appropriate if certain criteria such as the clear existence of interacting entities, or the possibility of emergent and dynamic properties are met (Macal & North, 2008). For example, researchers or developers of ABM systems may require interdisciplinary knowledge (technical and domain knowledge) to carefully create agent behaviours and interactions that are reasonable according to expert knowledge in a field or domain application. ABMs for urban systems are difficult to define and are computationally intensive even with the recent advances in computing hardware, which makes it problematic to create when extremely complex or large urban systems are to be modeled (Abar et al., 2017). For example, multiple computers may be required to run complex large-scale models, and results must then be combined from the different computers, which requires that models be parallelizable.

Geosocial media data help address some of the limitations of ABMs for urban systems by providing empirical data to verify and calibrate the results of models. ABMs can be verified, depending on the level of detail or scale required for the study, using geosocial media data. For example, modelled agent positions can be compared with the positions of user posts from geosocial media data to provide a measure of how accurate an ABM is to real-world data (Prager & Wiegand, 2014). The ABM can then simulate positions that are difficult (positions every seconds) or impossible (future positions) to obtain with verifiable evidence. This is particularly useful due to the large amount of geosocial media data and the individual level of detail that the data provide. Another benefit of geosocial media data for ABMs for urban systems is that geosocial media data provide information to calibrate ABMs so that particular parameters or behaviours (depending on the study or model design) can be optimized. For example, using large samples of geosocial media data, model parameters (such as the selection of agent population sizes) can be estimated, and behaviours (such as the probability of interactions between agents or responses to environmental changes) can be adjusted to better reflect real-world conditions parameters (Bohensky, Smajgl, & Herr, 2007; Haer et al., 2016). The main advantage of geosocial media data is that it provides high velocity and high volume data at the individual level, which are difficult to obtain with traditional methods such as surveys or target-group studies, to calibrate, verify, and improve ABMs for modelling urban systems.

However, geosocial media presents limitations relative to representativeness, privacy concerns, and data quality. Since geosocial media data are controlled by the social media platform provider, only a limited sample of data is available, which poses questionable data representativeness issues. Depending on a study’s objectives, obtaining data will require careful selection of samples such that they abide by the regulations
and restrictions set by the provider (Li et al., 2016b). For example, limiting extraction of data to a particular area/location to lower data rate limits, or only extracting data for certain keywords, phrases, or user accounts. An additional consideration to collaborate and apply for access to less restrictive geosocial media data may also be needed based on a study’s design (Morstatter, Pfeffer, & Liu, 2014). In this case, certain organizations or research groups may be limited due to cost, reputation, timing, and other factors that could potentially pose a bias between studies (one study’s model performing better than the other due to data access). Since geosocial media data contains individual level information, there is also a concern of privacy with respect to identification and misuse of data (Smith, Szongott, Henne, & Von Voigt, 2012). For example, geosocial media data can be used for public defamation, revoking rights of free speech, or identifying user addresses for internet stalking, to name a few examples. Given the privacy issues, researchers will also need to carefully consider anonymizing data or reducing the level of detail in the data for their study – especially in cases where the data is to be shared to the public or to other organizations. Lastly, geosocial media data quality is dependent on the culture and community of the social media platform’s user base. Depending on a study’s objectives, there needs to be consideration into the validity, believability, relevancy, and consistency of user posts (Immonen, Pääkkönen, & Ovaska, 2015). This is due to users having the freedom to create multiple identities online anonymously, and having the ability to also create automated bots that have the potential to spread misinformation or false user posts (Shao, Ciampaglia, Varol, Flammini, & Menczer, 2017). If ABMs model after erroneous or misrepresented samples, then the model may not reflect real-world conditions well, or in worse cases draw upon false conclusions controlled by misinformation or bot-generated posts. Although geosocial media provides a data source that was not traditionally possible in the past, allowing large-scale behavioural and geolocated data to be collected, it poses many limitations in validity, privacy, and representativeness for developing ABMs for urban systems.

4.2. Challenges and opportunities

The challenges of ABMs for urban systems using geosocial media data are focused on effective management and standardization of data and models. The enormous volume and rapid frequency of geosocial media data creates issues in the handling and processing of the data for use in ABMs (Li et al., 2016a). Geosocial media data is often generated as unstructured text data for human communication, which requires that it be efficiently processed and assessed for quality before analysis. For example, location-enabled Tweets require Natural Language Processing (NLP) methods (tokenization, stemming, chunking, and part-of-speech tagging) to filter for hashtags and clean stop words before it is ready for building models such as topic models or named entity recognizers (Pinto, Gonçalo Oliveira, & Oliveira Alves, 2016). The volume and spatial dimension of the Tweets affect the processing time required for the filter and data cleaning, which signifies the importance of selecting or developing algorithms that are scalable and efficient without drastically degrading model quality. In addition to the management of geosocial media data, the appropriate abstraction of urban systems to build ABMs can effectively lower the complexity and processing time of resulting models, which poses another challenge in
selecting a simplification of agents such that computational efficiency and correctness are optimal. Optimal efficiency and correctness are based on a study’s conditions and vary based on a study’s data, objective, and knowledge of the study’s phenomena. For example, ABM simulations may reach an optimal state given particular reasons – such as finding a repeating pattern, finding the best ratio of computing time and accuracy, and/or finding the lowest error within a set of study-defined constraints. Perhaps the most important challenge for both geosocial media data and ABMs in the context of urban systems is the development and adoption of standards. Both geosocial media data and ABMs have inconsistency in structures due to multiple platforms and flexible design principles, respectively. The inconsistent structure makes it difficult to compare various sources of geosocial media data and ABMs. The widespread use of geosocial media data and maturity of ABMs make it difficult to popularize standards due to the adoption of already available data and modelling approaches.

The challenges of ABMs and geosocial media data lead to opportunities to develop and advance ABMs and geosocial media data for urban systems studies. As the volume of geosocial media is expected to increase, opportunities arise to develop scalable ABMs that can make use of the data. To standardize agent-based modelling structures, protocols, such as the Overview Design Concepts and Details (ODD) (Grimm et al., 2010; Müller et al., 2013) and the Foundation of Intelligent Physical Agents (FIPA) (O’Brien & Nicol, 1998; Bellifemine, Poggi, & Rimassa, 1999; Poslad et al., 2000), have been developed in an effort to allow easy communication and comparison of ABMs. However, these protocols will require timely improvements that adapt to the needs of researchers and users over time. For example, the addition of newly developed comparison measures, such as metrics that evaluate emerging ABM patterns (Parker & Meretsky, 2004), and frameworks/guidelines that account for advancements in technology, such as cloud infrastructure (Fortino, Guerrieri, Russo, & Savaglio, 2014; Singh & Malhotra, 2012), big data (Kavak, Padilla, Lynch, & Diallo, 2018; Scheutz & Mayer, 2016), and mobile computing (Anagnostopoulos et al., 2007; Qi, Xu, & Wang, 2003). Similarly, geosocial media data has different APIs that standardize within platform data structures, but are often inconsistent outside specified platforms.

Geosocial media data availability also vary due to different platform usage (users using one platform for a particular purpose and another for another purpose), user behaviour (users choosing to enable geolocation or not), socio-economic status (56% of low income households use social media in 2015), and age (90% of young American adults aged 18 to 29 use social media in 2015), to name a few (Morstatter & Liu, 2017; Perrin, 2015). These can result in data biases depending on a researcher’s study objectives. For example, North American countries may tend to have more Twitter/Facebook/Instagram users, and China has restricted Twitter/Facebook/Instagram platforms for its citizens, which may cause larger usage of other alternatives such as WeChat and Sina Weibo (Men & Tsai, 2013). In this case, using Twitter/Facebook/Instagram to conduct social related studies in China may not be representative due to the availability of social media platforms. Another example may be to use social media data to study the political views of older or elderly adults. In this example, the portion of older or elderly adults using social media may be very low, and not representative of that target population for the study.

Privacy is another challenge of geosocial media data. Issues with fine-grained user data such as coordinate data, name, and profile picture can lead to misuse of geosocial media
5. Conclusion

This paper presented a review of agent-based modelling approaches in the context of urban systems using select geosocial media data. Urban systems have been described as complex and dynamic systems that involve self-organizing, interacting, and adapting entities that form emergent patterns. The complex and dynamic nature of urban systems made agent-based modelling an appropriate and effective approach for studying the social interactions in urban systems. People are the driving entities that have heavily affected urban systems through actions such as political decisions and urban planning. The social interactions among people were described as a form of communication, which has been advanced by geosocial media. The incorporation of geosocial media data, agent-based modelling, and statistical approaches have optimized and verified ABMs for various real-world applications involving social infrastructures in urban systems. However, agent-based modelling remains as a computationally intensive approach, especially combined with geosocial media data, that is commonly difficult to quantify and verify without involving subjective adjustments and decisions. Thus, it is important to develop and standardize ABMs that use geosocial media data such that they are efficient and modular without drastically compromising correctness.

The rapid urban growth across the world has made it important to study urban systems for the benefit of the future. McPhearson, Haase, Kabisch, and Gren (2016) suggested the examination and exploration of urban form, sustainability, land use, and functionality to advance understanding of urban systems. These examinations and explorations lead to improved management of social, environmental, and physical issues that arise in urban systems – particularly in developing areas (Cohen, 2006). The growing availability of geosocial media data (Facebook Inc, 2019a; Twitter Inc, 2019a) and increasing interest in ABMs (Markovic & Zornic, 2016) have encouraged multidisciplinary research and the use of location-enabled social media data, which can potentially lead to innovative studies in urban systems.

Disclosure statement

No potential conflict of interest was reported by the authors.
Notes on contributors

Richard Wen received the B.E.S. degree in geomatics with a computer science minor from the University of Waterloo, Waterloo, Ontario, Canada in 2014, and the M.S.A. degree in spatial analysis from Ryerson University, Toronto, Ontario, Canada in 2016. Currently, he is pursuing the Ph.D. degree in civil engineering at Ryerson University, Toronto, Canada. From 2014 to 2016, he was a Research Assistant with Geographic Information Science and Systems Group (GIS2) at Ryerson University, Toronto, Canada. Since 2016, he has been a Research Visitor at the Centre for Global Health Research, St. Michael’s Hospital, Toronto, Canada, and a Research Assistant at the Geographic Information Systems and GeoCollaboration Lab (GGL) at Ryerson University, Toronto, Canada. His research interests include geographic information systems and science, artificial intelligence and machine learning, spatial decision support systems, and systems and software engineering.

Songnian Li received the B.Eng. degree in surveying engineering from Wuhan Technical University of Surveying and Mapping, Wuhan, China, in 1983 and the Ph.D. degree in geodesy and geomatics engineering from the University of New Brunswick, Fredericton, Canada, in 2002. From 1983 to 1997, he was a Lecturer, Associate Professor and Visiting Scholar in universities in China and Canada. Since 2001, he has been an Assistant Professor, Associate Professor and Professor in the Department of Civil Engineering, Ryerson University, Canada. His current research interests include geospatial big data, spatio-temporal analysis, moving object and LBS data, geocollaboration and geospatial web and web services. Prof. Li has served as a council member of the International Society for Photogrammetry and Remote Sensing (ISPRS) (2016-2020), among other professional societies, was an invited researcher by Japan Society for the Promotion of Science in 2009 and 2012, and was a recipient of ISPRS U. V. Helava Best Paper Award 2016.

ORCID

Richard Wen  http://orcid.org/0000-0002-8904-8152
Songnian Li  http://orcid.org/0000-0002-8244-5681

References

Abar, S., Theodoropoulos, G. K., Lemarinier, P., & O’Hare, G. M. (2017). Agent based modelling and simulation tools: A review of the state-of-art software. Computer Science Review, 24, 13–33.
Anagnostopoulos, C. B., Tsounis, A., & Hadjiefthymiades, S. (2007). Context awareness in mobile computing environments. Wireless Personal Communications, 42(3), 445–464.
Bao, J., Lian, D., Zhang, F., & Yuan, N. J. (2016). Geosocial media data analytic for user modeling and location-based services. The SIGSPATIAL Special, 7(3), 11–18.
Barbati, M., Bruno, G., & Genovese, A. (2012). Applications of agent-based models for optimization problems: A literature review. Expert Systems with Applications, 39(5), 6020–6028.
Batty, M. (2013). The new science of cities. Cambridge, Massachusetts: MIT press.
Bazzan, A. L. C., & Klügl, F. (2013). A review on agent-based technology for traffic and transportation. The Knowledge Engineering Review, 29, 1–29.
Beigi, G., & Liu, H. (2020). A survey on privacy in social media: Identification, mitigation, and applications. ACM Transactions on Data Science, 1(1), 1–38.
Bellifemine, F., Poggi, A., & Rimassa, G. (1999). JADE–A FIPA-compliant agent framework. Proceedings of PAAM, 99(97–108), 33.
Berger, T. (2001). Agent-based spatial models applied to agriculture: A simulation tool for technology diffusion, resource use changes and policy analysis. Agricultural Economics, 25(2–3), 245–260.
Bettencourt, L. M., Lobo, J., Strumsky, D., & West, G. B. (2010). Urban scaling and its deviations: Revealing the structure of wealth, innovation and crime across cities. *PloS One*, 5(11), e13541.

Bohensky, E., Smajgl, A., & Herr, A. (2007). Calibrating behavioural variables in agent-based models: Insights from a case study in East Kalimantan, Indonesia. In *MODSIM 2007 International Congress on Modelling and Simulation*. Christchurch, New Zealand: Modelling and Simulation Society of Australia and New Zealand.

Bonabeau, E. (2002). Agent-based Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences of the USA*, 99(10), 7280–7287.

Borschchev, A., & Filippov, A. (2004). From system dynamics and discrete event to practical agent based modeling: Reasons, techniques, tools. In M. Kennedy, G. W. Winch, R. S. Langer, J. I. Rowe, & J. M. Yanni (Eds.), *Proceedings of the 22nd international conference of the system dynamics society* (Vol. 22, p. 45). Oxford.

Bretagnolle, A., Daudé, E., & Pumain, D. (2006). From theory to modelling: Urban systems as complex systems. *Cybergeo: European Journal of Geography*, 1–23. doi:10.4000/cybergeo.2420

Cocks, D. (2006). How Should We Use the Term "System of Systems" and Why Should We Care? Introduction to the Problem. *INCOSE International Symposium*, 16(1), 427–438.

Coffey, W. J. (1998). Urban systems research: An overview. *Canadian Journal of Regional Science*, 21(3), 327–364.

Cohen, B. (2006). Urbanization in developing countries: Current trends, future projections, and key challenges for sustainability. *Technology in Society*, 28(1–2), 63–80.

Crooks, A. T., & Heppenstell, A. J. (2012). Introduction to agent-based modelling. In *Agent-Based Models of Geographical Systems*, 164, 85–105.

Darvishi, M., & Ahmadi, G. (2014). Validation techniques of agent based modelling for geospatial simulations. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. 40(2/W3), 91–95. 10.5194/isprsarchives-XL-2-W3-91-2014

Davis, J. P., Eisenhardt, K. M., & Bigham, C. B. (2007). Developing theory through simulation methods. *The Academy of Management Review*, 32(2), 480–499.

Durak, M., & Till, R. (2015). Optimizing an agent-based traffic evacuation model using genetic algorithms. In L. Yilmaz, W. K. V. Chan, I. Moon, T. M. K. Roeder, C. Macal, & M. D. Rossetti (Eds.), *Proceedings of the 2015 Winter Simulation Conference* (pp. 288–299). Huntington Beach: IEEE Xplore.

Ebrahimpour, Z., Wan, W., Velázquez García, J. L., Cervantes, O., & Hou, L. (2020). Analyzing social-geographic human mobility patterns using large-scale social media data. *ISPRS International Journal of Geo-Information*, 9(2), 125.

Facebook Inc. (2019a). Facebook annual report. Retrieved June 19, 2020, from https://investor.fb.com/financials/?section=annualreports

Facebook Inc. (2019b). The graph API. Retrieved June 19, 2020, from https://developers.facebook.com/docs/graph-api

Fortino, G., Guerrieri, A., Russo, W., & Savaglio, C. (2014). Integration of agent-based and cloud computing for the smart objects-oriented IoT. In *Proceedings of the 2014 IEEE 18th international conference on computer supported cooperative work in design (CSCWD)* (pp. 493–498). Hsinchu, Taiwan, China: IEEE.

Français, U., Sayama, H., McSwiggen, C., Daneshvar, R., & Bar-Yam, Y. (2016). Visualizing the “heartbeat” of a city with tweets. *Complexity*, 21(6), 280–287.

Frias-Martinez, E., Williamson, G., & Frias-Martinez, V. (2011). An agent-based model of epidemic spread using human mobility and social network information. *2011 IEEE Third International Conference on and 2011 IEEE Third International Conference on Social Computing (SocialCom)*, Boston, MA, 49–56. https://doi.org/10.1109/PASSAT/SocialCom.2011.142

Fujita, M., Krugman, P., & Mori, T. (1999). On the evolution of hierarchical urban systems. *European Economic Review*, 43(2), 209–251.

Fujita, M., & Mori, T. (1997). Structural stability and evolution of urban systems. *Regional Science and Urban Economics*, 27(4–5), 399–442.

Gardner, M. (1970). Mathematical games: The fantastic combinations of John Conway’s new solitaire game “life”. *Scientific American*, 223(4), 120–123.
Ge, X., & Kremers, E. (2015). Optimization applied with agent based modelling in the context of urban energy planning. 2015 Winter Simulation Conference (WSC), 3096–3097. 10.1109/WSC.2015.7408417

Ghavami, S. M., Taleai, M., & Arentze, T. (2016). Socially rational agents in spatial land use planning: A heuristic proposal based negotiation mechanism. Computers, Environment and Urban Systems, 60, 67–78.

Goh, G., Koh, J. Y., & Zhang, Y. (2019). Twitter-informed crowd flow prediction. IEEE International Conference on Data Mining Workshops, ICDMW, 2018-November, 624–631. 10.1109/ICDMW.2018.00097

Gomide, J., Veloso, A., Meira, W., Almeida, V., Benevenuto, F., Ferraz, F., & Teixeira, M. (2011). Dengue surveillance based on a computational model of spatio-temporal locality of Twitter. Proceedings of the 3rd International Web Science Conference (1–8). 10.1145/2527031.2527049

Grimm, V., Berger, U., DeAngelis, D. L., Polhill, J. G., Giske, J., & Railsback, S. F. (2010). The ODD protocol: A review and first update. Ecological Modelling, 221(23), 2760–2768.

Haer, T., Botzen, W. J. W., & Aerts, J. C. J. H. (2016). The effectiveness of flood risk communication strategies and the influence of social networks—Insights from an agent-based model. Environmental Science & Policy, 60, 44–52.

Hall, A., & Virrantaus, K. (2016). Visualizing the workings of agent-based models: Diagrams as a tool for communication and knowledge acquisition. Computers, Environment and Urban Systems, 58, 1–11.

Heppenstall, A., Malleson, N., & Crooks, A. (2016). “Space, the final frontier”: How good are agent-based models at simulating individuals and space in cities? Systems, 4(1), 9.

Hristova, D., Williams, M. J., Musolesi, M., Panzarasa, P., & Mascolo, C. (2016). Measuring urban social diversity using interconnected geosocial networks. Proceedings of the 25th International Conference on World Wide Web, 21–30. 10.1145/2872427.2883065

Hua, T., Zhao, L., Chen, F., Lu, C.-T., & Ramakrishnan, N. (2016, November). How events unfold: Spatiotemporal mining in social media. The SIGSPATIAL Special, 7(3), 19–25.

Huang, W., & Li, S. (2016). Understanding human activity patterns based on space-time-semantics. ISPRS Journal of Photogrammetry and Remote Sensing, 121, 1–10.

Huang, W., Li, S., Liu, X., & Ban, Y. (2015). Predicting human mobility with activity changes. International Journal of Geographical Information Science, 29(9), 1569–1587.

Huber, M., Knottnerus, J. A., Green, L., van der Horst, H., Jadad, A. R., Kromhout, D., … Schnabel, P. (2011). How should we define health? BMJ, 343, 1–3.

Immonen, A., Pääkkönen, P., & Ovaska, E. (2015). Evaluating the quality of social media data in big data architecture. IEEE Access, 3, 2028–2043.

Ingalls, R. G. (2008). Introduction to Simulation. In S. Mason, R. Hill, M. Lars, & O. Rose (Eds.), Proceedings of the 40th Conference on Winter Simulation (pp. 17–26). Retrieved from http://www.informs-sim.org/wsc08papers/263.pdf

Jiang, W., Wang, Y., Tsou, M. H., & Fu, X. (2016). Using geo-targeted social media data to detect outdoor air pollution. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, 41, 553–554.

Jiang, Y., Li, Z., & Ye, X. (2019). Understanding demographic and socioeconomic biases of geotagged Twitter users at the county level. Cartography and Geographic Information Science, 46(3), 228–242.

Johnson, B., & Hernandez, A. (2016). Exploring Engineered Complex Adaptive Systems of Systems. In C. H. Dagli (Ed.). Procedia Computer Science, 95, 58–65.

Kavak, H., Padilla, J. J., Lynch, C. J., & Diallo, S. Y. (2018, April). Big data, agents, and machine learning: Towards a data-driven agent-based modeling approach. In Proceedings of the Annual Simulation Symposium (pp. 1–12). Baltimore, Maryland.

Kim, H. S. (2016). What drives you to check in on Facebook? Motivations, privacy concerns, and mobile phone involvement for location-based information sharing. Computers in Human Behavior, 54, 397–406.
Law, A. M. (2008). How to build valid and credible simulation models. In S. Mason, R. Hill, L. Monch, & O. Rose (Eds.), Proceedings of the 40th Conference on Winter Simulation (pp. 39–47). Retrieved from http://dl.acm.org/citation.cfm?id=1516758

Levy, S., Martens, K., & van der Heijden, R. (2016). Agent-based models and self-organisation: Addressing common criticisms and the role of agent-based modelling in urban planning. Town Planning Review, 87(3), 321–338.

Li, Q., Shah, S., Thomas, M., Anderson, K., & Liu, X. (2016b). How much data do you need? Twitter decahose data analysis. In The 9th International Conference on Social Computing, Behavioral-Cultural Modeling, & Prediction. Washington, DC, USA.

Li, S., Dragicevic, S., Castro, F. A., Sester, M., Winter, S., Coltekin, A., … Cheng, T. (2016). Geospatial big data handling theory and methods: A review and research challenges. ISPRS Journal of Photogrammetry and Remote Sensing, 115, 119–133.

Lu, R., Zhu, H., Liu, X., Liu, J. K., & Shao, J. (2014). Toward efficient and privacy-preserving computing in big data era. IEEE Network, 28(4), 46–50.

Luo, W., Gao, P., & Cassels, S. (2018, July). A large-scale location-based social network to understanding the impact of human geosocial interaction patterns on vaccination strategies in an urbanized area. Computers, Environment and Urban Systems, 72, 78–87.

Macal, C. M., & North, M. J. (2008). Agent-Based Modeling and Simulation: AMBS Examples. In S. J. Mason, R. R. Hill, L. Mönch, O. Rose, T. Jefferson, & J. W. Fowler (Eds.), Proceedings of the 2009 Winter Simulation Conference (pp. 101–112). 10.1109/WSC.2008.4736060

Macnamara, J., & Zerfass, A. (2012). Social Media Communication in Organizations: The Challenges of Balancing Openness, Strategy, and Management. International Journal of Strategic Communication, 6(4), 287–308.

Markovic, A., & Zornic, N. (2016). Trends in the application of agent-based modeling and simulation. Central European Conference on Information and Intelligent Systems, 65–70. Varazdin, Croatia.

Masoumzadeh, A., & Joshi, J. (2011, November). Anonymizing geo-social network datasets. In Proceedings of the 4th ACM SIGSPATIAL International Workshop on Security and Privacy in GIS and LBS (pp. 25–32). Chicago, IL.

McPhearson, T., Haase, D., Kabisch, N., & Gren, Å. (2016). Advancing understanding of the complex nature of urban systems. Ecological Indicators, 70, 566–573.

Men, L. R., & Tsai, W. H. S. (2013). Beyond liking or following: Understanding public engagement on social networking sites in China. Public Relations Review, 39(1), 13–22.

Mitchell, M. (2009). Complexity: A Guided Tour. New York, NY: Oxford University Press. doi:10.1063/1.3326990

Morstatter, F., & Liu, H. (2017). Discovering, assessing, and mitigating data bias in social media. Online Social Networks and Media, 1, 1–13.

Morstatter, F., Pfeffer, J., & Liu, H. (2014). When is it biased? Assessing the representativeness of twitter’s streaming API. In Proceedings of the 23rd international conference on world wide web (pp. 555–556). Seoul, Korea.

Müller, B., Bohn, F., Dreßler, G., Groeneveld, J., Klassert, C., Martin, R., … Schwarz, N. (2013). Describing human decisions in agent-based models - ODD+D, an extension of the ODD protocol. Environmental Modelling and Software, 48, 37–48.

Netto, V. M., Meirelles, J., & Ribeiro, F. L. (2017). Social interaction and the city: The effect of space on the reduction of entropy. Complexity, 2017.

O’Brien, P. D., & Nicol, R. C. (1998). FIPA—towards a standard for software agents. BT Technology Journal, 16(3), 51–59.

O’Connor, T., & Wong, H. Y. (2002). Emergent Properties. In The Stanford Encyclopedia of Philosophy (pp. 1–23). Retrieved from http://plato.stanford.edu/archives/fall2012/entries/social-institutions/

Ostermann, F. O., Huang, H., Andrienko, G., Andrienko, N., Capineri, C., Farkas, K., & Purves, R. S. (2015). Extracting and Comparing Places Using geosocial Media. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, II-3/WS, 311–316. doi:10.5194/isprsannals-ii-3-ws-311-2015
Parker, D. C., & Meretsky, V. (2004). Measuring pattern outcomes in an agent-based model of edge-effect externalities using spatial metrics. Agriculture. *Ecosystems & Environment, 101*(2–3), 233–250.

Perrin, A. (2015). Social media usage. *Pew Research Center, 52–68.*

Pinto, A., Goncalo Oliveira, H., & Oliveira Alves, A. (2016). Comparing the performance of different NLP toolkits in formal and social media text. In 5th Symposium on Languages, Applications and Technologies (SLATE’16). Maribor, Slovenia: Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik.

Poslad, S., Buckle, P., and Hadingham, R. 2000. The FIPA-OS agent platform: Open source for open standards. Proceedings of the 5th international conference and exhibition on the practical application of intelligent agents and multi-agents, (Vol. 355, No. 20).

Prager, S. D., & Wiegand, R. P. (2014). Modeling Use of Space from Social Media Data Using a Biased Random Walker. *Transactions in GIS, 18*(6), 817–833.

Qi, H., Xu, Y., & Wang, X. (2003). Mobile-agent-based collaborative signal and information processing in sensor networks. *Proceedings of the IEEE, 91*(8), 1172–1183.

Qi, L., Li, J., Wang, Y., & Gao, X. (2019). Urban Observation: Integration of Remote Sensing and Social Media Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 1–13.* doi:10.1109/jstars.2019.2908515

Rama, G. M., & Kak, A. (2015). Some structural measures of API usability. *Software - Practice and Experience, 45*(1), 75–110.

Rand, W., Herrmann, J., Schein, B., & Vodopivec, N. (2015). An Agent-Based Model of Urgent Diffusion in Social Media. *Journal of Artificial Societies & Social Simulation, 18*(2), 1.

Ranneries, S. B., Kalar, M. E., Nielsen, S. A., Dalgaard, L. N., Christensen, L. D., & Kanhabua, N. (2016). Wisdom of the local crowd: Detecting Local Events Using Social Media Data. In W. Nejdl, W. Hall, P. Parigi, & S. Staab (Eds.), *Proceedings of the 8th ACM Conference on Web Science - WebSci ’16* (pp. 352–354). Hannover, Germany. doi:10.1145/2908131.2908197

Schuetz, M., & Mayer, T. (2016). Combining agent-based modeling with big data methods to support architectural and urban design. In *Understanding Complex Urban Systems* (pp. 15–31). Cham: Springer.

Shao, C., Ciampaglia, G. L., Varol, O., Flammini, A., & Menczer, F. (2017). The spread of fake news by social bots. *arXiv Preprint arXiv:1707.07592, 96, 104.*

Singh, A., & Malhotra, M. (2012). Agent based framework for scalability in cloud computing. *International Journal of Computer Science & Engineering Technology (IJCSET), 3*(4), 41–45.

Smith, L., Liang, Q., James, P., & Lin, W. (2017). 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, 10*(3), 370–380.

Smith, M., Szonott, C., Henne, B., & Von Voigt, G. (2012). Big data privacy issues in public social media. In 2012 6th IEEE international conference on digital ecosystems and technologies (DEST) (pp. 1–6). Campione d’Italia, Italy: IEEE.

Storper, M., Van Marrewijk, C., & Van Oort, F. G. (2012). Introduction: Processes of change in urban systems. *Journal of Regional Science, 52*(1), 1–9.

Twitter Inc. (2019a). *Annual Report 2019.* Retrieved from https://investor.twitterinc.com/annuals-proxies.cfm

Twitter Inc. (2019b). Twitter Developer Documentation. Retrieved June 19, 2020, from. https://dev.twitter.com/docs

United Nations. (2018). *World Urbanization Prospects: The 2018 Revision.* Retrieved from https://population.un.org/wup/Publications/Files/WUP2018-KeyFacts.pdf

Wang, Q., & Taylor, J. E. (2016). Patterns and limitations of urban human mobility resilience under the influence of multiple types of natural disaster. *PLoS ONE, 11*(1), 1–14.

Weller, K. (2015). Trying to understand social media users and usage: The forgotten features of social media platforms. *Online Information Review, 40*(2), 256–264.

Wooldridge, M., & Jennings, N. R. (1995). Intelligent agents: Theory and practice. *The Knowledge Engineering Review, 10*(2), 115–152.

Wu, F. (2002). Complexity and urban simulation: Towards a computational laboratory. *Geography Research Forum, 22, 22–40.*
Wu, L., Zhi, Y., Sui, Z., & Liu, Y. (2014). Intra-urban human mobility and activity transition: Evidence from social media check-in data. *PLoS ONE, 9*(5), 1–13.

Xu, S., Li, S., & Huang, W. (2020). A Spatial-temporal-semantic Approach for Detecting Local Events using Geosocial Media Data. *Transactions in GIS, 24*(1), 142–173.

Xu, S., Li, S., & Wen, R. (2018). Sensing and Detecting Traffic Events using Geosocial Media Data: A Review. *Computers. Environment and Urban Systems, 72*(2018), 146–160.

Zhang, H., Vorobeychik, Y., Letchford, J., & Lakkaraju, K. (2016). Data-driven agent-based modeling, with application to rooftop solar adoption. *Autonomous Agents and Multi-Agent Systems, 30*(6), 1023–1049.

Zhang, J.-D., & Chow, C.-Y. (2015). Point-of-Interest Recommendations in Location-Based Social Networks. *The SIGSPATIAL Special, 7*, 26–33.