Chapter 12
Social Media and Social Awareness

Xinyue Ye, Bo Zhao, Thien Huu Nguyen and Shaohua Wang

Abstract The human behaviors and interactions on social media have maintained themselves as highly dynamic real-time social systems representing individual social awareness at fine spatial, temporal, and digital resolutions. In this chapter, we introduce the opportunities and challenges that human dynamics-centered social media bring to Digital Earth. We review the information diffusion of social media, the multi-faced implications of social media, and some real-world cases. Social media, on one hand, has facilitated the prediction of human dynamics in a wide spectrum of aspects, including public health, emergency response, decision making, and social equity promotion, and will also bring unintended challenges for Digital Earth, such as rumors and location spoofing on the other. Considering the multifaceted implications, this chapter calls for GIScientists to raise their awareness of the complex impacts of social media, to model the geographies of social media, and to understand ourselves as a unique species living both on the Earth and in Digital Earth.

Keywords Social media · Human dynamics · Social awareness · Location spoofing

12.1 Introduction: Electronic Footprints on Digital Earth

Geo-positioning system-enabled instruments can record and reveal personal awareness at fine spatial, temporal, and digital resolutions (Siła-Nowicka et al. 2016; Li et al. 2017; Ye and Liu 2019). With an exponential growth, human dynamics data are retrieved from location-aware devices, leading to a revolutionary research agenda regarding what happens where and when in the everyday lives of people in both real and virtual worlds (Batty 2013; Yao et al. 2019). Many location-based social media

X. Ye (✉) · S. Wang
College of Computing, New Jersey Institute of Technology, Newark, USA
e-mail: xye@njit.edu

B. Zhao
Department of Geography, University of Washington, Seattle, USA

T. H. Nguyen
Department of Computer and Information Science, University of Oregon, Eugene, USA

H. Guo et al. (eds.), *Manual of Digital Earth*,
https://doi.org/10.1007/978-981-32-9915-3_12

(LBSM) instances have been gaining popularity, fostering the emergence of fine-grained georeferenced social media content through these personalized devices (Liu et al. 2018a, b). The proliferation of LBSM enables researchers and practitioners to efficiently track a large and growing number of human action and interaction records over time and space to develop insights and enhance decision-making process from individual to global levels. The patterns and trends produced by LBSM can identify the movement of active social media users and aid in inferring demographics and related infrastructures. The collected data on users' physical and virtual activities facilitate the in-depth understanding of human dynamics from various aspects (Barabasi 2005; Shaw et al. 2016). The large volumes of such user-generated locational and contextual information are especially beneficial to studies relevant to the evolution of population size and human settlement structure as well as highly topical subjects such as traffic and epidemiological forecasting. For instance, real-time customer shopping behaviors might be rapidly identified by searching specific keywords in tweets, which allows for urban researchers and business analysts to monitor the fine-scale dynamics of economic geography and market outcome (Ye and He 2016). This new data landscape might not directly provide an ultimate solution to long-standing social or economic issues, but can increasingly shed light on many societal characteristics that are otherwise difficult to discover using traditional questionnaires or surveys.

Human actions and interactions in the digital form as well as frequent status updates can manifest themselves as highly dynamic real-time social systems, which enable the government to formulate appropriate policies for the relevant groups and targeted communities (Shi et al. 2018; Wang and Ye 2018). The electronic footprints and perceptions left by social media users and derivatives of complicated social networks can be utilized to enhance the design of location-based services (Ye and Lee 2016). Hence, the increasing demands in mapping and analyzing social media data call for more innovative conceptual and technological advances in visual and computational methods. These research challenges and opportunities can facilitate a paradigm shift in the broader social science disciplines in this new form of data landscape. Social media messages can depict the interconnected patterns and relationships between cyberspace and physical space, and can also be distributed instantly to a large number of users globally, who may belong to different virtual communities (Shelton et al. 2015).

Geographic information has traditionally been spread by governments or industries in a top-down manner; but its broadcast is much faster through social media than official agencies. The dramatic transition towards bottom-up digital dissemination has challenged these official or professional processes. Individuals can utilize the power of volunteered geographic information to minimize the difference and/or quality between experts and nonexperts in the context of generating a large collection of user-described features and numerous georeferenced citizen observations on socio-economic phenomena. With social media platforms becoming increasingly location-enabled, users can share geo-tagged information about their own lives and, as a result, rich content about large populations can be aggregated for social and behavioral studies (Sui and Goodchild 2011). Such a practice facilitates the policy transition from long-term to short-term action with a new perspective of understanding, visualizing, and analyzing human dynamics (Batty 2013).

The use of LBSM content represents a significant methodological advancement in social sciences and humanities research, providing rich content regarding human-environment interaction with locational estimation in ubiquitous/pervasive computing. It can efficiently assist place-based policy interventions in a timely fashion. Prompt and rigorous detection of emerging social and economic events calls for more robust algorithms to support such unprecedented research efforts in both qualitative and quantitative analyses. However, challenges and difficulties remain in processing user-generated messages to derive effective and high-quality information, considering the complex syntax and context embedded in social media messages. Additionally, if data analytics cannot be effectively conducted, the expected results could lose value for decision makers. These issues must be addressed to realize the potential of social media analytics.

Considering the above-mentioned issues, the remainder of this chapter is organized as follows. Section 12.2 describes the multifaceted implications of social media. Regarding social media, the unprecedented opportunities to predict human dynamics are introduced in Sect. 12.3, while multiple challenges are listed in Sect. 12.4. Then, the implications of these opportunities and challenges are further discussed in Sect. 12.5, followed by a conclusion in Sect. 12.6.

12.2 Multifaceted Implications of Social Media

Value systems are fundamental to anything we do. Today, the rapid proliferation of social media has greatly affected us and almost every aspect of human society. Confronted with this complicated and unstoppable interaction, we employ value structures to holistically discover the implications of social media, especially the unintended but vital ones. McLuhan's (1975) law of media is frequently utilized to capture the social consequences of various media. Tuan (2003) also proposed the psychology of power to unveil the internal logic of human's perceptions of places, and Ihde (1990) contemplated how technology mediates between human beings and the world from a phenomenological perspective.

Among these value structures, we employ Ihde's amplification-reduction structure to investigate the opportunities and challenges brought by social media. This structure reveals how technology (including social media) amplify and simultaneously reduce a certain human experience. The amplified and reduced experiences are intertwined and interrelated. More significantly, the amplified human experience is obvious whereas the reduced human experience is undiscoverable and easily ignored. Though Ihde only suggested applying this structure to the human experience, it can also be applied to understand the social implications of the investigating object. Through this structure, the opportunities and challenges of the social implications can be revealed. For example, during the 2008 Olympic Games, social media was touted as a tool of freedom to enable the general public to express their concerns about the air pollution issues in major Chinese cities. If we acknowledge the promotion of free speech as the opportunities brought about by social media, the hidden challenges can be revealed through this value structure—social media can also be used as a tool of surveillance by big brother to control the discussion on air pollution as well as a medium of advertisement by private companies to sell relevant products (e.g., masks) to prevent air pollution-related symptoms. The implications of social media are multifaceted. Therefore, the value structure can be applied to examine the impacts of social media on the rapidly evolving Digital Earth. In the following sections, we discuss the opportunities provided by social media as well as the potential challenges.

12.3 Opportunities: Human Dynamics Prediction

As a newly chartered territory for human activities, social media has resulted in tremendous electronic footprints. Such footprints represent a large number of the population and can be used to predict human dynamics on the ground via the relationships between the spread of information, user characteristics, and message contents. In this section, we discuss how social media can be used for different aspects of human society, including public health, emergency response, decision making, and social equity promotion.

12.3.1 Public Health

Social media platforms can be used to mitigate the spread of pandemics and associated anxiety. Scholars have used sentiment analysis and spatial analysis to examine how social media communication conveys information about contagious and infectious diseases and alerts the public, through identifying, tracking, and visualizing the behavioral patterns of users (Zadeh et al. 2019). For instance, Ye et al. (2018a, b) explored public health-related rumors during disease outbreaks and evaluate how such media framing sets the tone negatively, affecting the quality of disease outbreak detection and prediction, using the diffusion of Ebola rumors in social media networks as a case study. Sharma et al. (2017) find that the inaccurate Facebook posts are more popular than those with accurate and relevant information about the Zika virus. Villar and Marsh (2018) studied the impact of social media health communication of Ebola and Zika, concluding that the effect relies on users' attitudes and trust towards authorities and the media. Average citizens and ordinary social media users have very limited knowledge regarding the accuracy and relevancy of infectious diseases spreading over time and across space as well as concerning complications. As a force in health communication, social media data could be utilized to define a temporal extent of the infection and to populate a spatial database of reported occurrences of the disease. Additionally, social media data can be used to track and predict the emergence and spread of infectious diseases and distribution across various spatial and temporal scales. As a self-reported volunteered information platform and useful surveillance tool, social media feeds outperform those from official or government outlets in timeliness. They can also aid in gaining insights into the opinions and perceptions of the public.

12.3.2 Emergency Response

The use of massive computer-mediated communication in emergency response and disaster management has captured considerable interest from both the general public and decision makers. Social media enables fast interpersonal communication during crises through information dissemination, early warnings, environmental awareness, and public participation in disaster-affected areas, allowing for emergency workers to respond more speedily and capably (Hashimoto and Ohama 2014; Finch et al. 2016). As Yin et al. (2012) argue, "this growing use of social media during crises offers new information sources from which the right authorities can enhance emergency situation awareness. Survivors in the impacted areas can report on-the-ground information about what they are seeing, hearing, and experiencing during natural disasters. People from surrounding areas can provide nearly real-time observations about disaster scenes, such as aerial images and photos." Moreover, since social media users can access information posted by official agencies through following their accounts, organizations and agencies can leverage social media as a platform to post authoritative situational announcements and communicate with the public in emergency situations and to potentially retrieve and verify on-the-ground information using the public as the information source (Wang et al. 2016). Palen et al. (2009) examined the consequences of digital communication and information sharing on emergency response in the context of the Virginia Tech massacre. Chen et al. (2016) proposed real-time geo-tagged tweet collection and recording in a distributed geodatabase as well as real-time data redistribution using a Web GIS application. This system was applied to a hypothetical mass evacuation using tweets from Hurricane Joaquin in 2015.

12.3.3 Decision Making

As a new kind of user-generated geospatial information, social media data could be invaluable to political agenda-setting that needs to be aware of location-based topic distribution. For example, the data could help political strategists analyze the tweets of residents or voters in a given geographical area. Politicians can gauge people's reactions by monitoring the communication among Twitter accounts regarding policy issues. Ye et al. (2017) employed voting tweets regarding a water bond in California to highlight place-based situational awareness. Convention and visitors' bureaus may focus on 'hot button' issues in certain places within their cities or regions. These data could provide operational indicators about places that are most visited or preferred by visitors, which can inform the marketing strategies relevant to these locations. Local governments could analyze social media messages to determine whether a proposed construction project would be favored by the public or if other proposed projects would be perceived positively by their constituents. Ye et al. (2018a, b) examined how the Multilevel Model of Meme Diffusion (M3D) captures the debates regarding death penalty abolishment across space. At the intracity scale, Liu et al. (2018a, b) assessed the utility efficiency of subway stations in a Chinese city by matching the capacity of train services and the travel needs using social media data. Deng et al. (2018) analyzed how geotagged tweets are associated with hourly electric consumption at the building level, given the assumption that tweeting behavior is highly related to human activities.

12.3.4 Social Equity Promotion

Most social media platforms such as Twitter, Facebook, or Instagram are designed for the general public; few are dedicated to specific groups (e.g., LGBTQ, photographers, natural disaster victims, etc.). An in-depth analysis and visualization of the specific groups can promote social equity among different groups. Social awareness of where they are is the first and foremost step in enabling local residents and governments to recognize the necessity to treat these underrepresented populations equally. For example, Jack'd, a dedicated gay social networking app, enables its users to communicate online with those who are physically nearby. Through collecting online locational information from Jack'd, a 3D distribution of the gay community in Beijing were visualized (Zhao et al. 2017). By overlapping this distribution with landmarks such as major roads, university campuses, shopping malls and gay-friendly places (e.g., gay bars, gay saunas, gay-friendly gyms, gay-friendly parks and public restrooms attracting gay activities, etc.), the characteristics of this underrepresented group's distribution can be revealed. Gay people in Beijing primarily concentrate in the northwestern and eastern parts of the city. The northwestern area is the center for higher education, with several famous universities. In the eastern area of Beijing, the area from Sanlitun to Worker's stadium is acknowledged as a recreation center for LGBT people. To the south, a few famous gay-friendly residential communities are surrounded by gay saunas; to the east, there are several high-end residential communities and shopping malls in the Guomao and Sihui subdistricts. This 3D distribution reveals a hot spot of gay activities in the Tongzhou district. This may result from the relatively low house rent and convenient accessibility to Chaoyang and other local urban centers for hangouts. Through this 3D distribution of the gay community's electronic footprint, the local public health agencies can provide corresponding services for the gay community and organize more targeted activities as an effective means to promote social equity.

12.4 Challenges: Fake Electronic Footprints

In addition to those obvious opportunities in human dynamics prediction, challenges inherently in social media are often ignored. As Chun et al. (2019) argue, "uncertainty and context pose fundamental challenges in GIScience and geographic research. Geospatial data are imbued with errors (e.g., measurement and sampling) and various types of uncertainty that often obfuscate any understanding of the effects of contextual or environmental influences on human behaviors and experiences." Although social media has been touted as a platform to authentically present human trajectories and their mobilities, rumors, spoofings and privacy concerns, not limited to the physical world, are also exist on Digital Earth. In this sense, We cannot immediately treat social media messages as accurate and credible without considering the above-mentioned issues.

12.4.1 Rumors

The unmoderated nature of social media user's posting behavior might lead to the accumulation of invalidated and unverified information and news involving speculation and uncertainty regarding social events (Ye et al. 2018a, b). Jones et al. (2017) found those who relying on social media for updates of a campus lockdown tend to suffer from greater distress due to their increased exposure to conflicting content in social media channels. Rumors are considered messages or forms of interaction among people about certain events that may not be true. As a nonprofessional medium, social media platforms can spread rumors. However, some information from reliable sources can minimize rumor propagation, lowering the level of anxiety in the virtual community. Zubiaga et al. (2018) noted that the openness of social media platforms also enables the study of user behavior on sharing and discussing both long-standing and newly emerging rumors based on natural language processing and data mining methods, especially for four components: rumor detection, rumor tracking, rumor stance classification, and rumor veracity classification.

12.4.2 Location Spoofing

Location spoofing is a deliberate geographic practice to disguise one's actual location with inconsistent locational information (Zhao and Sui 2017). It facilitates the spoofer to virtually travel to places of interest for various purposes. For smartphones, the spoofing mechanism can be divided into three steps, (1) blocking the positioning service of a smartphone to acquire the actual locational information, (2) generating inconsistent locational information, and (3) transmitting it to an operating LBSM app (e.g., Twitter, Facebook, Pokémon Go, etc.) on a smartphone (Zhao and Zhang

2018). As a result, the LBSM app mistakes the fake location as where the operating smartphone really is. Specifically, the positioning service relies on a hybrid approach that integrates three major positioning techniques: built-in GPS, surrounding WiFi network triangulation, and cellular tower network triangulation. For these three techniques, the more accurate, the higher priority in deciding the final result. In practice, the fundamental function of location spoofing is to downgrade the accuracy of the positioning technique or totally block the positioning function. There are three common location spoofing techniques, in terms of falsifying the MAC addresses of surrounding WiFi routers, spoofing GPS signals in the environment, and mocking in-transit locational information. The last method is predominantly adopted by dedicated mobile android apps for location spoofing. Such apps enable users to virtually visit a place other than the actual location. An example is presented below to clarify this.

In reaction to the 2009 presidential election in Iran, the government of Iran regularly monitors all activities on social networks (Ansari 2012). During the campaign, social network sites were suddenly blocked, and online political activity became the target of harsh criticism and reprisals from the government. To prevent this surveillance and protect online protestors, many internationally based Green Movement supporters spread disinformation over Twitter to mislead local police. Foreign supporters who were not in Iran decided to set their online locations to Tehran to protect those who were tweeting from Tehran. This strategy may have helped some Iranian opposition leaders avoid persecution, but also made it impossible to understand the real impacts of Twitter on the protest.

12.4.3 Privacy Abuse

When users share content and their data on social media, there is a risk that such content and data are collected and exploited in a way that is not expected by the users. This poses a serious challenge in terms of privacy for user data and calls for the responsibility of the network administrators, researchers and users to preserve privacy in social networks. Two broad classes of privacy issues in social networks—user-user privacy and user-third party privacy—are discussed below.

In social media, one user might share content about another user or party. Although this mechanism helps spread the content over the networks efficiently, it inherently presents a tremendous risk for privacy violation. For instance, your friends might share a picture you posted, showing you were in a restaurant with another friend. The picture sharing might be done without your consent and accidentally reveal your location, private information that you do not want to share beyond your friend list. To prevent such privacy breaches, social media administrators have implemented mechanisms for users to make complaints and request that the content be removed from the networks. However, before the content can be reviewed and revoked, it might have caused some detrimental consequences for the users. It would be more effective if such content dissemination was validated at the very beginning. Addressing the user-user privacy issue requires collaboration among scientists from different disciplines, including computer scientists, GIScientists, and psychologists. For example, Kekulluoglu et al. (2017) studied a hybrid negotiation architecture with a reciprocity mechanism to mimic the social responsibility in reality, and a credit system was used to encourage agents/users to respect other's privacy in social media.

Regarding user-third party privacy, content and data generated by social media users might be collected by different third parties for various purposes, potentially causing serious data leaks and violating the privacy of users. A retailer might retrieve user profiles and posts to deliver appropriate ads to the users or an upstart voter-profiling company could exploit such information to characterize the personalities of users and influence their voting decisions (e.g., the recent Facebook privacy crisis and data leak with Cambridge Analytica on American elections described in Rosenberg et al. (2018)). Another example is researchers who query user data to infer various user characteristics (e.g., depression, drug abuse) (Choudhury et al. 2013). While such inferences can provide valuable insights into different social problems and support monitoring systems for social issues, the leaks of such inferred information for specific users can cause biases and affect the users' ability to participate in social activities (e.g., jobs, school admission). Consequently, it is important to develop technological strategies to ensure privacy in user data-related activities in social media. The Future of Privacy Forum and DataGuidance (2018) delivered the report "Comparing privacy laws: GDPR v. CCPA." This report compares the European Union's General Data Protection Regulation (GDPR) effective on May 25, 2018, and the California Consumer Privacy Act of 2018 (CCPA) scheduled to be in effect on January 1, 2020. Both laws would also fundamentally influence social media platforms in collecting/sharing/employing users' data online and offline.

12.5 From Awareness to Action

A close scrutiny of the opportunities and challenges would raise our awareness of the potential capacity of social media in understanding human dynamics. As Yang et al. (2016, p. 61) argued, "the convergence of social media and GIS provides an opportunity to reconcile space-based GIS and place-based social media." Driven by this awareness, GIScientists should take actions to model the geographies of social media, propose innovative approaches to location spoofing screening and connect the virtual world in social media with the real word to better explain social media phenomena.

12.5.1 Modeling the Geographies of Social Media

Tracking and predicting the diffusion of social media information from a neighborhood to a global scale raises a series of questions such as where and when certain topics will be discussed and become popular. Sui and Goodchild (2011) suggested two hypotheses to test the nature of social media message diffusion such as geo-tagged hashtags spread through Twitter. The spatial influence model states that the spatially nearby locations tend to be impacted in the near future, and the community affinity influence model asserts that such dissemination would occur between functionally connected places. However, the reality is usually a combination of these two models. Such predictions will be useful for policymakers to estimate the spatial and functional influence of economic downturns facilitated by supply-chain networks. The community affinities are expected to enhance the prediction power of purely spatial models.

12.5.2 Detecting Location Spoofing Through Geographic Knowledge

If we examine location spoofing from the traditional standard of scientific data, it is highly unlikely that such "fake" information is generated by environmental uncertainties, measurement uncertainties, or limited knowledge about measurement (Zhang and Goodchild 2002). Today, location spoofing cannot simply be treated as fake data, as these data are associated with complicated generative motivations from different stakeholders, governments, local business or average social media users. To identify location spoofing, it is necessary to determine the motivations why the author produces that location, and then judge whether it is spoofed or not.

Therefore, we must seek appropriate solutions to the positioning inconsistency and the motivations for spoofing. Usually, self-reporting (e.g., survey, questionnaire) or observations can qualitatively collect and interpret human motivations that trigger the generation of positional inconsistency. However, in practice, it is difficult to measure the real motivation: admittedly, the survey or questionnaire participants might not report their true intentions of location spoofing due to the fear of being recognized as location spoofers or rumormongers.

The positioning inconsistency in spoofing can be quantitatively detected. Theoretically, any spoofing detection is supposed to unveil a certain underlying positioning inconsistency. As Goodchild (2013) indicated, a geospatial accuracy model interprets how a world is constructed geographically. In this sense, spoofing detection is meant to detect scenarios that do not follow the way in which the world is geographically constructed. One crucial theoretical framework to build up the geographic truth is Hägerstrand's Time Geography (1970). This analytical framework conceptualizes the trajectory of each individual as a life path, which is restricted by several predefined human behavioral constraints in space and time. Meanwhile, a series of analytical tools to measure human dynamics are provided by Time Geography, such as space-time path, prism, and cube. Zhao and Sui (2017) provided a Bayesian time geographic estimation approach to determine the places that an examined user is unlikely to appear. Time-geographic density estimation (TGDE) was used to model the human appearance in a region over time. TGDE can convert trajectories (e.g., a time sequence of historical geo-tags from an individual) to a visiting probability distribution of spatial positions over time. This model can effectively convey the behavioral constraints and describe where and when an individual is more likely to visit. A location with a lower probability value is more likely to be spoofed. Moreover, with the rise of deep learning such as long short-term memory, LSTM (Greff et al. 2017), it is worth investigating application of deep learning techniques in detecting fake location information. For example, given a set of sequenced historical geo-tags of an individual, LSTM can be used to model the sequential information and build deep learning-based classification methods.

12.5.3 Connecting Social Media with the Real World

In social media, people share their thoughts and emotions about events in the real world. Such events might be explicitly mentioned or implicitly referred to in their posts. For instance, some social media posts might explicitly include a link to a news article they would like to discuss whereas other posts might express the users' attitude on some events without citing those events. In many inference problems for social media data (e.g., sentiment analysis, opinion mining), it is crucial to determine the corresponding realistic events to fully understand and explain the trends and phenomena in social media (i.e., connecting social media with the real world). One example is that social media posts concerning implicit events where the absence of the implied events would clearly impede accurate analysis of the posts.

To model the real world, we can resort to public information resources such as news articles and public knowledge resources such as Wikipedia and Freebase (Bollacker et al. 2008). These resources cover a wide range of events across various aspects of life. They are also updated with new events in almost real time due to the recent advances in publication technologies, promoting these public information resources as a digital counterpart of the real world. Consequently, we can connect social media data with the real word via the reflected world of public information resources. The major technical challenges to accomplish this connection involve the ability to autonomously extract events from those public resources (e.g., news articles) and the capacity to link the information in social media to the appropriate detected events. Such challenges would require a deep analysis of the semantics of the information presented in both (e.g., the posts in social media and the events in public resources) to identify the events and connections with high accuracy. Fortunately, deep semantic understanding of such information is being actively investigated in artificial intelligence research, including natural language processing, computer vision, graph modeling and machine learning. For instance, many recent studies have shown that events in news articles can be effectively curated using deep learning techniques, a branch of machine learning that is capable of automatically inducing the underlying representations for data to achieve high extraction performance (Nguyen et al. 2016; Nguyen and Grishman 2018; Nguyen and Nguyen 2019). As these event extraction techniques can also recognize the time and locations at which the events occur, they can be beneficial for GIScientists in geographical research of populations on social media and the real world. In addition, deep learning might also present effective solutions for the problem of linking social media data with realistic events due to its recently demonstrated capacity for embedding and representation learning for various problems. Once converged, such advances in these fields of computer science might eventually offer an opportunity to connect the virtual world and real world by solving the aforementioned technical challenges.

Finally, the realistic events from public information resources enable novel semantic-based solutions to combat the problems of rumors or fake news in social media. An important property of the public resources discussed in this section is that they generally capture trustful information/events, as such information is verified by the media administrators for accuracy and correctness. This is one reason why news articles are usually slower than social media in presenting the information to the public. Consequently, if the social media information can be accurately linked and compared with the information/events in the trustful information sources, novel detection and tracking techniques can be proposed to prevent rumors and fact-check the information spread over social media. Artificial intelligence research can provide the fundamental technologies to tackle these problems, as demonstrated in recent research in natural language processing and deep learning (i.e., Yin and Roth 2018).

12.6 Conclusion

As a crucial platform for human dynamics and activities, social media content can be mined in multiple approaches to determine how individuals connect and share information as well as purposefully move across scales and resolutions (Croitoru et al. 2013; Miller et al. 2019). When social media activities are attached with locational information, these online human activities can generate tremendous electronic footprints on Digital Earth. Especially when merging with other digital overlays of authoritative data through multisource data fusion, such as land use, urban planning, and natural resources data, powerful interoperation and prediction that require both electronic footprints and digital overlays on Digital Earth become feasible with the optimal weights for combination (De Albuquerque et al. 2015; Lin et al. 2019). These digital overlays serve as the socioenvironmental context within which the geosocial media dynamics and events occur and evolve, calling for scientific cross-fertilization of many separate domains toward an integrated science of human dynamics.

In this chapter, we introduce the opportunities and challenges that human dynamics-centered social media bring to Digital Earth. We review the information diffusion of social media, the multifaced implications of social media, and some real-world cases. Social media will facilitate the prediction of human dynamics in a wide spectrum of aspects, including public health, emergency response, decision making and social equity promotion. Social media will also bring unintended challenges for Digital Earth, such as rumors and fake location spoofing. Considering the multifaceted implications, this chapter calls for GIScientists to raise their awareness of the complex impacts of social media and urges them to model the geographies of social media as well as filter fake locations through geographic knowledge, targeting a more robust geosocial knowledge discovery. Social media will continue to evolve, along with the development of human society. Social media has become a crucial part of human activities on Digital Earth. Any effort that ignores the importance of social media will bring the effort into question. Therefore, the study of social media provides new data sources and data collection methods about real-world activities and happenings, and social media help us in profoundly understanding ourselves as a unique species living both on the Earth and in Digital Earth.

Acknowledgements This material is partially based upon work supported by the National Science Foundation under Grant Nos. 1416509, 1535031, and 1739491. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

References

Ansari A (2012) The role of social media in Iran's green movement (2009-2012). Glob Media J Aust Ed 12:1–6

Barabási A-L (2005) The origin of bursts and heavy tails in human dynamics. Nature 435(7039):207–211

Batty M (2013) Big data, smart cities and city planning. Dialogues Hum Geogr 3(3):274–279

Bollacker K, Evans C, Paritosh P et al (2008) Freebase: a collaboratively created graph database for structuring human knowledge. In: Proceedings of the 2008 ACM SIGMOD international conference on management of data - SIGMOD '08. ACM Press, New York, NY, p 1247–1250

Chen X, Elmes G, Ye X et al (2016) Implementing a real-time twitter-based system for resource dispatch in disaster management. GeoJournal 81(6):863–873

Choudhury MD, Gamon M, Counts S et al (2013) Predicting depression via social media. In: Proceedings of the seventh international aaai conference on weblogs and social media. p 128–137

Chun Y, Kwan M-P, Griffith DA (2019) Uncertainty and context in GIScience and geography: challenges in the era of geospatial big data. Int J Geogr Inf Sci 33(6):1131–1134

Croitoru A, Crooks A, Radzikowski J et al (2013) Geosocial gauge: a system prototype for knowledge discovery from social media. Int J Geogr Inf Sci 27(12):2483–2508

de Albuquerque JP, Herfort B, Brenning A et al (2015) A geographic approach for combining social media and authoritative data towards identifying useful information for disaster management. Int J Geogr Inf Sci 29(4):667–689

Deng C, Lin W, Ye X et al (2018) Social media data as a proxy for hourly fine-scale electric power consumption estimation. Environ Plan A Econ Space 50(8):1553–1557

Finch KC, Snook KR, Duke CH et al (2016) Public health implications of social media use during natural disasters, environmental disasters, and other environmental concerns. Nat Hazards 83(1):729–760

Future of Privacy Forum and DataGuidance (2018) Comparing privacy laws: GDPR v. CCPA. https://fpf.org/wp-content/uploads/2018/11/GDPR_CCPA_Comparison-Guide.pdf. Accessed 12 May 2019

Goodchild MF (2013) The quality of big (geo)data. Dialogues Hum Geogr 3(3):280–284

Greff K, Srivastava RK, Koutnik J et al (2017) LSTM: a search space odyssey. IEEE Trans Neural Netw Learn Syst 28(10):2222–2232

Hägerstraand T (1970) What about people in regional science? Pap Region Sci 24(1):7–24

Hashimoto Y, Ohama A (2014) The role of social media in emergency response: the case of the great East Japan earthquake. NIDS J Def Secur 15:99–126

Ihde D (1990) Technology and the lifeworld: from garden to earth (No. 560). Indiana University Press, Bloomington

Jones NM, Thompson RR, Schetter CD et al (2017) Distress and rumor exposure on social media during a campus lockdown. Proc Natl Acad Sci USA 114(44):11663–11668

Kekulluoglu D, Kokciyan N, Yolum P (2017) Preserving privacy as social responsibility in online social networks. ACM Trans Internet Technol 18(4):1–22

Li M, Dong L, Shen Z et al (2017) Examining the interaction of taxi and subway ridership for sustainable urbanization. Sustainability 9(2):242

Lin J, Wu Z, Li X (2019) Measuring inter-city connectivity in an urban agglomeration based on multi-source data. Int J Geogr Inf Sci 33(5):1062–1081

Liu Q, Wang Z, Ye X (2018a) Comparing mobility patterns between residents and visitors using geo-tagged social media data. Trans GIS 22(6):1372–1389

Liu X, Macedo J, Zhou T et al (2018b) Evaluation of the utility efficiency of subway stations based on spatial information from public social media. Habitat Int 79:10–17

McLuhan M (1975) McLuhan's laws of the media. Technol Cult 16(1):74–78

Miller HJ, Dodge S, Miller J et al (2019) Towards an integrated science of movement: converging research on animal movement ecology and human mobility science. Int J Geogr Inf Sci 33(5):855–876

Nguyen TH, Grishman R (2018) Graph convolutional networks with argument-aware pooling for event detection. In: The association for the advancement of artificial intelligence (AAAI). AAAI Press, Menlo Park, California, p 5900–5907

Nguyen TM, Nguyen TH (2019) One for all: neural joint modeling of entities and events. In: The association for the advancement of artificial intelligence (AAAI), arXiv.org > cs > arXiv:1812.00195

Nguyen TH, Cho K, Grishman R (2016) Joint event extraction via recurrent neural networks. In: Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, p. 300–309

Palen L, Vieweg S, Liu SB et al (2009) Crisis in a networked world. Soc Sci Comput Rev 27(4):467–480

Rosenberg M, Confessore N, Cadwalladr C (2018) How trump consultants exploited the facebook data of millions, New York Times. https://www.nytimes.com/2018/03/17/us/politics/cambridge-analytica-trump-campaign.html?module=inline. Accessed 12 May 2019

Sharma M, Yadav K, Yadav N et al (2017) Zika virus pandemic—analysis of facebook as a social media health information platform. Am J Infect Control 45(3):301–302

Shaw S-L, Tsou M-H, Ye X (2016) Editorial: human dynamics in the mobile and big data era. Int J Geogr Inf Sci 30(9):1687–1693

Shelton T, Poorthuis A, Zook M (2015) Social media and the city: rethinking urban socio-spatial inequality using user-generated geographic information. Landsc Urban Plan 142:198–211

Shi X, Xue B, Tsou M-H et al (2018) Detecting events from the social media through exemplar-enhanced supervised learning. Int J Digit Earth 12(9):1083–1097

Siła-Nowicka K, Vandrol J, Oshan T et al (2016) Analysis of human mobility patterns from GPS trajectories and contextual information. Int J Geogr Inf Sci 30(5):881–906

Sui D, Goodchild M (2011) The convergence of GIS and social media: challenges for GIScience. Int J Geogr Inf Sci 25(11):1737–1748

Tuan Y-F (2003) On human geography. Daedalus 132(2):134–137

Villar ME, Marsh E (2018) Social media and infectious disease perceptions in a multicultural society. In: Villar ME, Marsh E (eds) Reconceptualizing new media and intercultural communication in a networked society. IGI Global, Pennsylvania, US, p 328–350

Wang Z, Ye X (2018) Space, time, and situational awareness in natural hazards: a case study of hurricane sandy with social media data. Cartogr Geogr Inf Sci 46(4):334–346
Wang Z, Ye X, Tsou M-H (2016) Spatial, temporal, and content analysis of Twitter for wildfire hazards. Nat Hazards 83(1):523–540
Yang X, Ye X, Sui DZ (2016) We know where you are. Int J Appl Geospat Res 7(2):61–75
Yao XA, Huang H, Jiang B et al (2019) Representation and analytical models for location-based big data. Int J Geogr Inf Sci 33(4):707–713
Ye X, He C (2016) The new data landscape for regional and urban analysis. GeoJournal 81(6):811–815
Ye X, Lee J (2016) Integrating geographic activity space and social network space to promote healthy lifestyles. SIGSPATIAL Spec 8(1):20–33
Ye X, Liu X (2019) Introduction: cities as social and spatial networks. In: Ye X, Liu X (eds) Cities as spatial and social networks. Springer, Cham, p 1–8
Ye X, Li S, Sharag-Eldin A et al (2017) Geography of social media in public response to policy-based topics. In: Ye X, Li S, Sharag-Eldin A et al (eds) Geospatial data science techniques and applications. CRC Press, Boca Raton, US, p 221–232
Ye X, Li S, Yang X et al (2018a) The fear of ebola: a tale of two cities in China. In: Ye X, Li S, Yang X et al (eds) Big data support of urban planning and management. Springer, Cham, p 113–132
Ye X, Sharag-Eldin A, Spitzberg B et al (2018b) Analyzing public opinions on death penalty abolishment. Chin Sociol Dialogue 3(1):53–75
Yin W, Roth D (2018) TwoWingOS: a two-wing optimization strategy for evidential claim verification. In: Proceedings of the 2018 conference on empirical methods in natural language processing. Association for Computational Linguistics, Brussels, Belgium, p 105–114
Yin J, Lampert A, Cameron M et al (2012) Using social media to enhance emergency situation awareness. IEEE Intell Syst 27(6):52–59
Zadeh AH, Zolbanin HM, Sharda R et al (2019) Social media for nowcasting flu activity: spatio-temporal big data analysis. Inf Syst Front 21(4):743–760
Zhang J, Goodchild MF (2002) Uncertainty in geographical information. CRC Press, Boca Raton, FL
Zhao B, Sui DZ (2017) True lies in geospatial big data: detecting location spoofing in social media. Ann GIS 23(1):1–14
Zhao B, Sui DZ, Li Z (2017) Visualizing the gay community in Beijing with location-based social media. Environ Plan A 49(5):977–979
Zhao B, Zhang S (2018) Rethinking spatial data quality: pokémon go as a case study of location spoofing. Prof Geogr 71(1):96–108
Zubiaga A, Aker A, Bontcheva K et al (2018) Detection and resolution of rumours in social media. ACM Comput Surv 51(2):1–36

Xinyue Ye is an Associate Professor in the College of Computing at New Jersey Institute of Technology where he directs Urban Informatics & Spatial Computing lab. He develops and implements new methods on spatiotemporal-social network analysis/modelling/simulation for different application domains such as economic development, disaster response, land use, public health, and urban crime.

Bo Zhao is an Assistant Professor in the Department of Geography at University of Washington. His research interests are on the social implications of geospatial technologies, the study of different spoofing phenomena in geography towards a post-truth reflection on the value of geospatial technologies, the geovisualization and geo-narrative of the coupled human and natural system.

Thien Huu Nguyen is an Assistant Professor in the Department of Computer and Information Science at University of Oregon. His research explores mechanisms to understand human languages for computers so that computers can perform cognitive language-related tasks for us.

Shaohua Wang is an Assistant Professor in the College of Computing at New Jersey Institute of Technology where he directs Software Engineering lab. His research spans the study of software engineering from system, empirical, and machine learning perspectives.