The Journal of Transport and Land Use jtlu.org
Vol. 16 No. 1 [2023] pp. 437–468

JTLU

Revealing social dimensions of urban mobility with big data: A timely dialogue

Jiangyue Wu
Department of Urban Planning and Design,
University of Hong Kong
jiangyue@connect.hku.hk

Jiangping Zhou (corresponding author)
Department of Urban Planning and Design,
University of Hong Kong
zhoujp@hku.hk

Abstract: Considered a total social phenomenon, mobility is at the center of intricate social dynamics in cities and serves as a reading lens to understand the whole society. With the advent of big data, the potential for using mobility as a key social analyzer was unleashed in the past decade. The purpose of this research is to systematically review the evolution of big data's role in revealing social dimensions of urban mobility and discuss how they have contributed to various research domains from early 2010s to now. Six major research topics are detected from the selected online academic corpuses by conducting keywords-driven topic modeling techniques, reflecting diverse research interests in networked mobilities, human dynamics in spaces, event modeling, spatial underpinnings, travel behaviors and mobility patterns, and sociodemographic heterogeneity. The six topics reveal a comprehensive, research-interests, evolution pattern, and present current trends on using big data to uncover social dimensions of human mobility activities. Given these observations, we contend that big data has two contributions to revealing social dimensions of urban mobility: as an efficiency advancement and as an equity lens. Furthermore, the possible limitations and potential opportunities of big data applications in the existing scholarship are discussed. The review is intended to serve as a timely retrospective of societal-focused mobility studies, as well as a starting point for various stakeholders to collectively contribute to a desirable future in terms of mobility.

Article history:
Received: September 28, 2022
Received in revised form: May 10, 2023
Accepted: June 28, 2023
Available online: November 16, 2023

1 Introduction

Ranking top as the uppermost value and the freedom to move and reflecting a profound stratifying factor for modern time, mobility represents individuals' social capital and contributes to social differences (Cwerner et al., 2008; Elliott & Urry, 2010; Kaufmann, 2014; Larsen et al., 2006;). In modern research paradigms, scholars no longer simply consider mobility as transportation flows of humans in space and time; rather, mobility implies the evidence of change or of growth, the tool for understanding city dynamics, the pulse of the community, and essentially, the metabolism of the city itself (Burgess, 1926). In a recent review from , et al. (2022), mobility is conceptualized as "spatial movements, social phenomenon, policy tool and indicator for economic activities."

https://dx.doi.org/10.5198/jtlu.2023.2281
ISSN: 1938-7849

The *Journal of Transport and Land Use* is the official journal of the World Society for Transport and Land Use (WSTLUR) and is published and sponsored by the University of Minnesota Center for Transportation Studies.

Alongside the evolving connotations of mobility, the advent of new technologies, faster and higher capacity modes of transport, and innovative mobility tools have profoundly shaped social dimensions of urban mobility. With finer data granularity, a greater volume of data quantity, as well as more extensive coverage, human mobility big data has demonstrated great potential to reveal more sophisticated travel behaviors of different socio-economic groups, marking a leap forward when compared to small-sample travel survey data collected via traditional approaches (Hasan & Ukkusuri, 2014; Gkiotsalitis & Stathopoulos, 2016; Luo et al., 2016; Y. Wang et al., 2018; Zhou et al., 2021). Alongside these studies that based on traditional data-collection approaches, studies based on passive big data sources entered as a game-changer that allow researchers diving into more complex interactions and network mechanisms. For instance, cellular network data allow authors to capture the nearly entire population of mobile users (Poorthuis et al., 2021; Puura et al., 2018; Sun et al., 2013). This was hardly imaginable before.

Based on papers identified through a keyword-driven selection which has been conducted on two comprehensive online academic databases (Web of Science and Scopus), this review article aims to address the following research questions:

1) To what extent has the utilization of big data broadened the research scope within the field of human mobility studies? and what novel topics have been expedited as a result of its application?
2) what are the opportunities and limitations faced by the identified papers, which are a sample of the increasingly enlarging body of literature on social dimensions of urban mobility facilitated by big data?

A quick answer to the proposed question is that big data's contribution to different research arenas is heterogeneous and the focused research topics driven by big data evolved from time to time, for example, papers focusing on networked mobilities have long been a subject of enduring research interest, whereas studies related to event modeling have emerged more recently and constitute a niche research area. The ensuing text is organized as follows to elaborate this. Section 2 reviews big data's evolving roles in accelerating information richness of sophisticated human mobility activities across various disciplines, i.e., from simple toolkit advancement to bridging interdisciplinary dialogues. Section 3 presents the methods underpinning our systematic literature review, including database selection, search criteria, screening processes, and analytical method. Section 4 discusses the detected topics and topic evolution from time to time, which helps us answer the research questions. Section 5 concludes.

2 Big data in urban mobility studies: A systematic review

The term "big data" was originally used to describe datasets that were too large or complex to be processed by traditional analytical tools (Manovich, 2015), with volume, variety, and velocity as three main attributes. Nowadays, the term tends to appertain to sophisticated behavior analytics, predictive analytics, and other related methods. While big data's extensive usage has a rather short history, the observation, and evaluation of human mobility and derived societal investigation are long-lasting. Even without any data, whether small or big, scholars from various backgrounds have authored seminal works by observing urban street activities (Jacobs, 1961; Whyte, 1988), shooting photos or time-lapse videos (Milgram, 1977; Whyte, 1988), and/or conducting onsite field studies (Blau, 1977; Simmel, 1995). Their works laid fertile ground for researchers interested in general urban sociology as well as mobility-focused studies.

With big data's arrival, some scholars have made seminal advancements in revealing new and insightful social dimensions of urban mobility (Noulas et al., 2012; Sun et al., 2013; Q. Wang et al., 2018) whereas others have questioned concerning the collection, analysis, and ethical issues of big data applications in social science domain (Crompton, 2008; Goldthorpe, 2015). The consensus reached by both sides is that some fields are more suitable to be boosted by big data (e.g., social network analysis), and others are not (e.g., entity recognition or subjective perceptions) (Halford & Savage, 2017).

Looking beyond the advanced algorithms and improved data analytics, the societal fabric reflected by more recent scholarship on mobility differentiate broaden our horizons of mobility, which now go far beyond physical transport or movements. This is similar to how Burgess discerned the differences between money and capital (1926), which was attended and examined by Marx (1867) before. Therefore, living in a technological and data-rich era, it is an opportune time to review how the studies born in a big data era can be connected to previous paradigms. This section aims to systematically review the evolution of big data's role in revealing social dimensions of urban mobility and discuss how they have contributed to different research domains.

2.1 Empowering traditional mobility research with big data's capacity

Prior to the arrival of big data, the scholarship on mobility heavily relied on travel journals and surveys, which are typically costly, untimely, and limited to a small sample size (Stopher, 2009; Stopher & Stecher, 2006). Since the beginning of this century, rapid information technology developments have provided more cost-effective data resources, which could be exploited to facilitate innovative research approaches to complex human behaviors and travel dynamics.

The emergence of fine-granularized, time-stamped, and geo-coded data has greatly extended the studies on human mobility dynamics, which used to be confined by travel survey analysis. On the one side, an important advancement of this stream of studies is that they have validated a wide range of existing knowledge from those survey-based studies. This has demonstrated the inherent consistency of knowledge, regardless of data types, big or "small." For example, Chudyk et al. (2015) used travel journey data from 150 participants collected in 1 to 7 days to reveal senior adults' trip preferences and found the most relevant destinations were grocery stores, malls, and restaurants/cafés; later, researchers utilized millions smartcard data records in Beijing and reached comparable conclusions (Shi et al., 2021). On the other side, big data allows researchers to investigate some nuanced and subtle travel behaviors that are likely to be neglected in traditional travel journey surveys. For instance, Zhou et al. (2020) took advantage of smartcard data in Beijing metropolitan area and identified "familiar strangers" in the metro transit system who shared a portion of similar trajectories and reflected potential proximity in social status and daily routine regularity; Long et al. (2016) recognized commuters with extreme travel behaviors (defined as very early or late, or long or frequent trips) and identified their jobs-housing distribution. Due to the limited sample sizes of conventional datasets based on small samples, such nuanced travel behaviors and related social groups are very likely to be neglected, therefore, big data have enlarged the existing research scope and allowed more studies on subpopulations. This could result in more innovative clustering analyses or categorization and sophisticated and nuanced findings.

While a series of innovative works emerged, the common pitfalls in big data also led to some research deficiencies. First, most big datasets are not multivariate, especially at the individual level, which forces researchers employ multiple data sources by different organizations, for different spatiotemporal resolutions, and in different statistical units. The lack of multivariate nature pushes researchers to fuse big data from various sources or even link big data and "small" data without common keys. This could possibly lead to inferior data quality and misleading results.

Second, big data is often predominantly collected from specific social groups and for particular business purpose, which could result in overrepresented populations or biased sampling. This could neglect or even possibly harm the underrepresented social groups (Howe & Elenberg, 2020). Studies utilizing big data to investigate different research questions (senior mobility, street walkability, willingness to travel, etc.) report common research limitations that their findings suffer from biased samples and limited curation for long-term research purposes (Gkiotsalitis & Stathopoulos, 2016; Kandt & Leak, 2019; Su et al., 2019)

Third, due to the unequal information, technology, and communication (ITC) infrastructure development across the globe, the tendency of big data dominance in academia would riskily exacerbate research inequality since the most underdeveloped countries could hardly collect, process, and analyze big data to meaningful engage in related research dialogues.

In sum, the first two decades of this century have witnessed rapid ITC advancements, which had a profound impact on mobility research. The emergence of multi-format, fine-granularized, large-sample, and nearly real-time big data sources greatly extended the research scope and depth of traditional mobility studies, allowing scholars to re-validate existing knowledge as well as investigate more social issues and topics concerning mobility. This stream of studies also reflects scholars' dilemma in using "small" data and/or big data. This remains a hot and overarching debatable topic hitherto.

2.2 Embracing uncertainty and complexity of big data

Human mobility data could be categorized into two types: one is recorded and stored on a massive level to depict population's characteristics, while another is a vector format and recorded on an individual basis to reflect human movement dynamics (Haraguchi, 2022). Traditional datasets (e.g., travel surveys and censuses) enable scholars to understand population distribution and its sociodemographic features; however, with the advent of multidimensional big data, researchers are able to understand more dynamic human movement and activity patterns (Birenboim et al., 2015; Etter et al., 2013; González et al., 2008; Huang et al., 2015; Rashidi et al., 2017; Shoval et al., 2015; Song et al., 2010; Xu et al., 2018). Since most emerging (big) data are not collected purposedly for the scholarship on human mobility, there remain many uncertainties and complexities in terms of their appropriate role in advancing the scholarship.

Emerging data's sources could be categorized as follows: a) financial activity logs to reveal individual's movement and consumption activities (Brockmann et al., 2006; Singh et al., 2015; Sobolevsky et al., 2014), b) call detail records (CDRs) to reflect both individual and interpersonal movement pattern (Gonzalez et al., 2008), c) vehicle trajectory data to demonstrate fine-granularized spatio-temporal dynamics of a specific type vehicles, for example, taxi trajectories as a popular data source (Yuan et al., 2013), d) public transit smart-card data (SCD) which records massive detailed boarding and alighting activities of huge number of transit users (Hasan, Schneider et al., 2013), and e) location-based social media platform (e.g., twitter posts, Flicker photos, Weibo tags, etc.) to characterize geographic places as well as recognize human mobility pattern (Cheng et al., 2011; Cho et al., 2011; Hasan, Zhan et al., 2013). Besides, some studies also investigated innovative data sources, for example, wireless network visits (Liang et al., 2016; Traunmueller et al., 2018), internet browsing (Y. Wang et al., 2018), and linguistic data (Wu et al., 2016). Table 1 summarizes the characteristics of different emerging data that have been exploited in the existing scholarship on human mobility.

As Table 1 presents, big data firstly emerged in the latter half of the first decade of the 21st century and were widely exploited in the scholarship. They have threefold impacts on the scholarship. To begin, more connotations of human mobility patterns are defined. Mobility is no longer narrowly defined by common daily travel or occasional long trips but is also linked to credit card transactions, social media checked-ins, phone calls with friends, or a Wi-Fi network visits. The linkage disclosed new issues that have not been examined before. Theoretically, some new relationships related to human movements and the environment have been hypothesized and tested, for example, the interplay between travel behaviors and weather conditions or air-pollution degrees (Chen, Chen et al., 2021; Wei, 2022; Xu et al., 2020; Zhao et al., 2018) Methodologically, interdisciplinary dialogues have been introduced, participants of these dialogues including scholars from fields such as computer sciences, geography, computational social science, and urban studies.

In the dialogues, participants also faced the inevitable complexities underlying emerging big data and their linkage to existing "small" data (see Table 1). As noted by Chen et al. (2016), a vast number of studies utilized the aggregated data, which therefore resulted in a risk of ecological fallacy. The missing or weak linkages between the aggregated mobility data and individual-level socio-economic data sources remain a long-lasting research challenge. Furthermore, data enrichment to complement lacking information in big data sources has gained increasing attention. For example, many transit systems only record the boarding or alighting activities of travellers rather instead of both. This forced developments of tools to tackle related issues through deep-learning architecture (Jung & Sohn, 2017), supervised machine learning (Shalit et al., 2020), and probabilistic model (Cheng et al., 2021). Eventually, it could be better off for the scholarship advancements if transit systems collected more information regarding passengers' boarding or alighting.

Table 1. Advantages and complexities of different types of human mobility data

Type	Definition	Advantages	Complexities	Studies
Financial activity logs	Data collected when purchases or transactions made by users	1) Richer socio-economic attributes (e.g., citizenship) 2) Longer-term coverage 3) Unique attribute related to consumption activities	1) Privacy concerns related to individual financial status 2) Often used in transcontinental mobility analysis, samples are biased 3) Often limited sample size	Brockmann et al., 2006; Singh et al., 2015; Sobolevsky et al., 2014.
CDRs	Data collected when cell phones communicate with mobile phone networks	1) Reports cell timestamps and the geolocations of users 2) Provides a large amount of data 3) Spatial coverage is large	1) Cell tower density can affect the accuracy of locations 2) Privacy concerns if data are not anonymized 3) Biases in mobile phone ownership	Gonzalez et al., 2008; Huang et al., 2018; Jiang et al., 2016.
Vehicle trajectories	Data collected when vehicles with GPS tracker are in motion	1) Consecutive records of movement paths, including speed and directions 2)Efficient in traffic situation evaluation	1) Time intervals affect data quality 2) Sensor infrastructure may affect data collection 3) Difficulty to connect with other big data sources	Hu, Gao, et al., 2021; Zhang et al., 2017; Zheng et al., 2016.

Type	Definition	Advantages	Complexities	Studies
SCD	Data collected when users enter/exit transit stations	1) Massive data quantity 2) Spatial coverage is large	1) Geo-coded by station, cannot reflect real origin/destination 2) Transfers remain unknown 3) Some systems only report onboard data 4) Public transits could be a small fraction of all trips	Briand et al., 2017; Zhao et al., 2018; Zhou et al., 2021 .
Location-based social media	Data collected when users post or check-in on various platforms	1) Provides the geolocations, timestamps, and content of social media posts 2) Provides a large amount of data 3) Reports high-resolution geolocations	1) Coverage limited 2) Only locations posted by users can be detected 3) Positioning indoors is less precise	Candipan et al., 2021; Gkiotsalitis & Stathopoulos, 2016; Hasan & Ukkusuri, 2014.
Others	Data collected when users log-in a Wi-Fi network, Camera signaling or browse specific contents on Internet	1) Reflect certain groups' behaviors 2) Limited sample size 3) Limited sample	1) Limited sample size 2) Often collected in a short period 3) Lower degree of reproducibility	Su et al., 2019; Wang & Vermeulen, 2021.

2.3 Revisiting spatial and mobility elements of social dimensions

Moreover, the social dimensions of transport and mobility have emerged as an increasingly contentious area (Lucas, 2012). Travel activities, behaviors, and mode choices are utilized as innovative research lenses to reflect complex social, economic, and spatial interactions (Jensen, 2013; Sheller, 2014). Research streams adopting these lenses based on our review of the refereed articles collected are synthesized as follows.

Firstly, using mobility information to infer spatiotemporal and demographic features of social groups has gained increasing research interest over the decades (Cranshaw et al., 2010; Liben-Nowell et al., 2005; Toole et al., 2015). As Kaufmann noted (2014), mobility "reflects and creates differences in a society," and a vast number of literatures have used mobility-related data as a new toolkit to uncover such differences across social groups. For example, Luo et al. (2016) extracted users' space-time trajectories from geo-tagged Twitter contents and identified their home and activity locations. By combining user's demographic information (extracted from user's profiles) and their mobility characteristics, they found that race/ethnicity was the dominant attribute that affects urban human mobility patterns. Similarly, Wang et al. (2018) compared mobility behaviors of groups with different internet browsing habits. They found that travellers' internet usage preferences could significantly predict their destination choices, even though such datasets hardly contain any socioeconomic attributes. Alongside the empirical studies, some scholars contributed to the methodological advancements that capture different social groups' mobility behaviors. For example, Yoo (2019) proposed a robust approach to the temporal aspect of human mobility and suggested a benchmark of 11 to 14 days as the minimum observing period to capture people's daily mobility routines.

While the first category of studies emphasizes on social groups or/and individuals, the second category zooms into the linkages among them. The latter investigated how existing social network influenced people's travel patterns (Carrasco et al., 2008), and how linkages among people's online/ offline activities predicted their potential social interactions (Calabrese et al., 2011). Toole et al. (2015) coupled human mobility and social ties by linking people's mobility similarity and their predicated visitation patterns. Findings suggested that the strengths of social ties among people were highly correlated with their mobility similarity, and such a relationship was capable of conducting empirical measurements in different urban contexts. The massive and fine-grained bigdata also enables scholars to investigate some more nuanced social ties, such as the familiar stranger phenomenon in urban transit systems (Sun et al., 2013; Zhou et al., 2018a, 2018b). By analyzing the transit riders who share a portion of trajectories for multiple times, they validated and revised the familiar stranger phenomenon that was originally defined and discussed by sociologist Stanley Milgram in 1977. With the assistance of transit smartcard data, this small set of studies contributed to a better understanding of social life in transit systems and the probable attitudes and personalities of urbanites that shape social life and are shaped by social life (Geng & Yang, 2017; Wu et al., 2022).

The third category emphasizes on the inequality and segregation reflected in collective travel patterns. Zhang and Zhao reviewed a series of studies on urban transport equity in the context of China and re-conceptualized the concept of transport equity, which is characterized by the equitable accessibility distribution. This distribution is mediated by the local institutional architecture, which is supposed to achieve equality in social opportunities (2021). In the Western context, there existed more and more studies on the socio-spatial segregation or exclusion across urban communities and/or groups, with the assistance of mobility big data (Davidson & Ryerson, 2021; Trasberg & Cheshire, 2021; Xu et al., 2018). Age (Kandt & Leak, 2019), ethnicity (Jarv et al., 2015; Zhai et al., 2021), gender (Bocker et al., 2020; Ooi et al., 2021), disability (Grisé et al., 2019), and access to innovative transport modes (Lovelace et al., 2020; Zhang et al., 2021) are widely discussed research focuses. These studies reflected a wide range of research scales, from individuals to regions and even the world.. They produced practical insights into public health measures development, infectious diseases containment, and general social policymaking.

Studies assessing urban vitality and vibrancy using human mobility patterns as proxies lie closer to the practice end of the research spectrum and are the fourth stream. Largely inspired by Jan Gehl and Jane Jacob's concepts, numerous attempts have been made to identify, evaluate, and analyze human dynamics through the lens of diversity. The fine-grained big data concerning mobility trajectories across various transport means have opened the possibility of utilizing those trajectories as an efficient proxy to measure urban vitalities, which can be measured by the spatiotemporal diversity of human movement activities (Kang et al., 2021), temporal variations of mobility flows (Sulis et al., 2018), human-land interaction intensity and diversity (Liu et al., 2020), and home-cantered activity coverage (Xu et al., 2015). This stream of studies enlarged Jacob's concept that pedestrian flows on streets represent vitality in cities; moreover, massive mobility data revealed the diversity of people's movement from limited street observations to large-scale data computations. Corresponding findings assist in the evolution of New Urbanism and smart growth strategies in a modern urban context (Lan et al., 2020).

Lastly, Covid-19 has spurred researchers to apply the above concepts under a specific pandemic context. Subsequently, a series of works emerged, contributing to a better understanding of the pandemic, containment strategies, and corresponding mobility impacts (Hu, Xiong, et al., 2021; Kim and Kwan, 2021). However, Given the limited room of one paper, we do not separately consider these works as a fifth stream.

2.4 Summary

Since its advent in the early 2010s, big data has penetrated into various research domains and practices. Its cherished expectations and questioned deficiencies stirred up heated discussions (Aiden & Michel, 2013; Buhl et al., 2013; Walker, 2014). From a fashionable dictum to a leading topic across various research domains and practices, big data's applications in human mobility attracted extensive research interests as well. Figure 1 shows a simple uptick of published research articles themed on both "big data" and "human mobility" by year, according to the Scopus database. Alongside the general accelerating research outputs, we are also interested in how different research realms demonstrated (possibly) various focuses and whether interdisciplinary dialogues have been initiated by the tool advancement of big data. We aim to answer the above questions in the ensuing sections

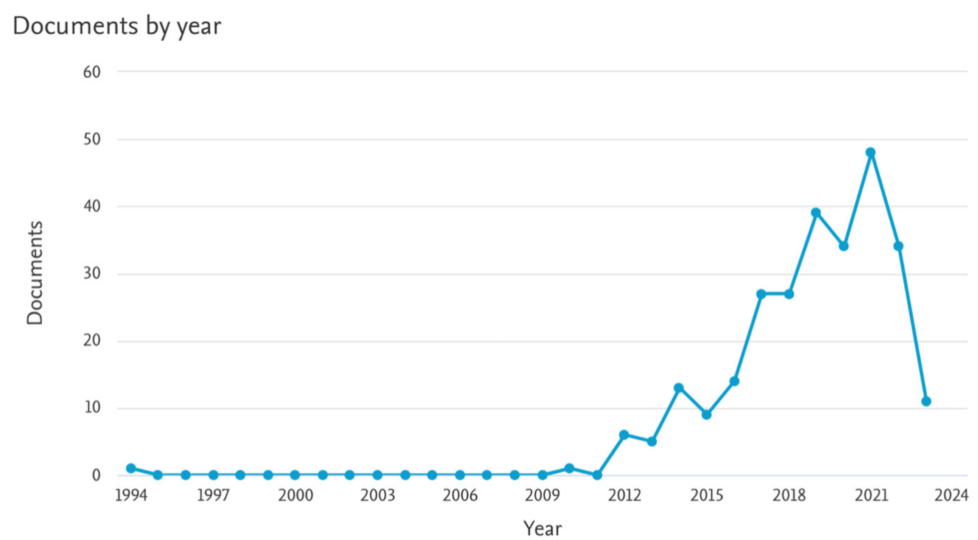

Figure 1. The number of published research articles themed in "big data" and "human mobility" by year in Scopus

3 Methods

In this section, a stepwise procedure of review methodology is presented, guided by the PRISMA guidelines (Page et al., 2021). Besides the methods, a detailed PRISMA checklist is available in Appendix.

3.1 Defining the keyword search strategy

To ensure an effective searching output, each research article is expected to cover the following three aspects of information, including at least one elements of social dimension attributes, at least one urban mobility attributes, and at least one application of big data sources. The Boolean logic operator AND is used between these three sets of search keywords, and logic operator OR is applied inside each keyword set. Three sets of search terms are identified as below.

- Keyword set A: those denote social dimension attributes, including: s*ocial ties, social network, social relationship, social interaction, human activity, social connection, social group, segregation, inclusion, exclusion, pr inequality*

- Keyword set B: those denote urban mobility attributes, including: *urban transport, urban transit, bus, public transport, metro, rail, bicycle, subway, automobile, walking, carpooling tram, ride-hailing or shared mobility.*
- Keyword set C: those denote big data related technologies, including: *big data, smartcard data, call detail record, location-based services, GPS data, geotagged data, mobile data, signal data or crowdsource*

3.2 Identifying corpus collections

After a comprehensive review of available online bibliographic databases, Web of Science (WoS) and Scopus are selected as the two major search databases, considering both comprehensive coverage and consistency in advance search features. The same searching strategy has been employed in these two databases to detect specific keywords in the publication title, abstract, and keywords. Only research articles published in English have been screened in this review.

As a result, WoS reported 119 papers covering all three sets of keywords, and Scopus reported 180 papers in total. To ensure that all papers in the corpus discuss the application of big data in revealing the social dimensions of human mobility activities, rigorous screening of titles and abstracts has been conducted to exclude papers indexed with the search keywords but unrelated to the intended topic. For instance, the definition of the keyword "vertical mobility" could be either one's social hierarchy movement or the vertical take-off and landing transport means. To avoid this confusion between keywords written in a similar way but with different meanings, we pre-processed the corpus by manually screening as the previous section presents, then input the manually fine-tuned corpus into the further analytical process. After this screening, 85 papers remained from WoS and 77 remains from Scopus. At last, 31 duplicate papers are found and removed, and 130 unique papers from two databases are prepared for machine learning powered topic modeling. Figure 2 presents the screening process.

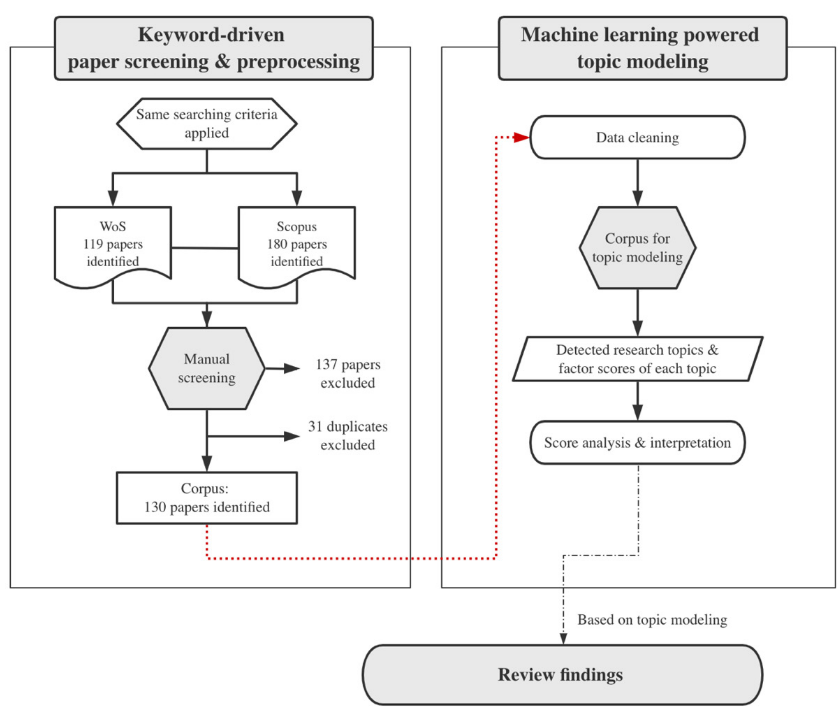

Figure 2. Screening process of the corpus

3.3 Topic modeling: Latent Dirichlet Allocation

The Latent Dirichlet Allocation (LDA), one of the most widely accepted unsupervised, probabilistic modeling methods in natural language processing, was selected as the research tool. First proposed by Pritchard et al. in 2000 and introduced by Blei et al. in 2003, the LDA analyses the given corpus to extract the latent topics by modeling the corpus as discrete distribution across topics (Blei, 2012). By using a matrix decomposition technique, LDA assumes each document is a mixture of topics, and is capable to visualize the topical clusters. Thus, it allows users to better understand the latent topics presented through different documents has gained popularity in the literature review (Asmussen & Muller, 2019), content recommendation (Younus et al., 2014), search engine optimization (Mavragani et al., 2018), consumer complaint analysis (Bastani et al., 2019), and a wide range of other applications. For the implementation of LDA in this study, coding and visualization were conducted on Python 3.

4 Results

4.1 Optimal number of detected topics

As one of the challenges in applying LDA analytics, determining appropriate number of detected topics is of great importance to ensure the quality and interpretation of the resulting topics. Although there is no single established method for determining the optimal number of topics (Gan & Qi, 2021), several commonly accepted evaluation tools are helpful in providing insights on topic number solutions, including perplexity measurements, coherence analysis, and visualization evaluations. In this review, a coherence analysis is conducted before topic modeling to identify the number of topics that results in the most cohesive theme and reports a highest coherence score of 0.333 with topic number as 2, and a second high coherence score of 0.322 with topic number as 6. The results are presented in Figure 3.

According to the coherence analysis results, the optimal number of detected topics for the corpus is identified as 6. This determination takes into careful consideration both interpretation and visualization aspects, signifying that a model encompassing six topics is anticipated to effectively capture the most coherent themes and representative patterns for this corpus.

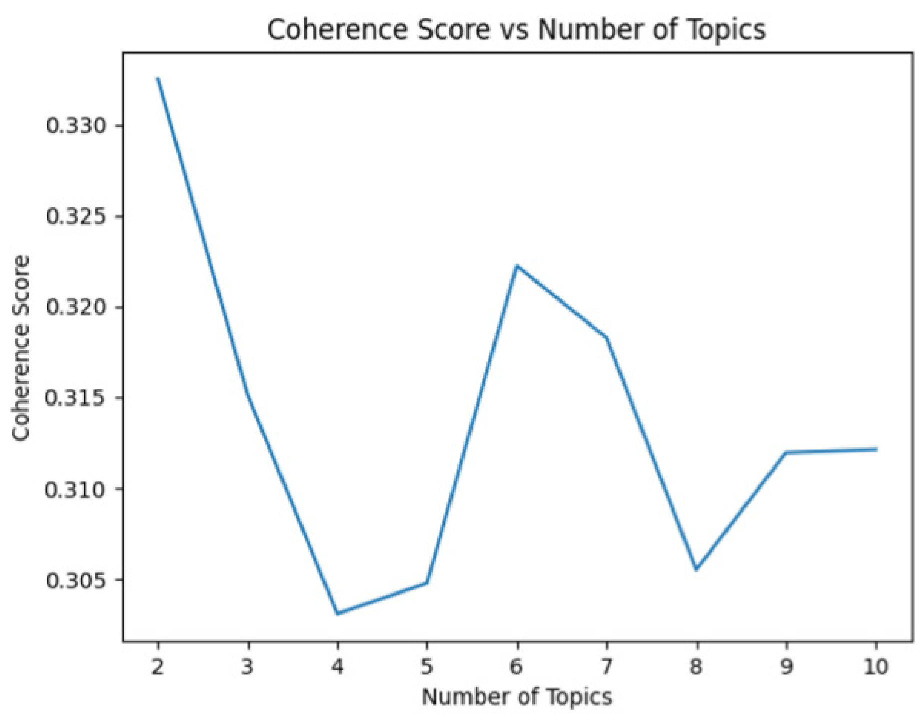

Figure 3. Coherence analysis to determine optimal number of detected topics

4.2 Topic detection, interpretation, and visualization

With the assistance of the LDA, 6 major topics were detected from the pre-processed corpus. Table 2 reports the key terms of each detected topic, topic label, count of related research papers to this topic, and top 5 high loading documents contributing to the detected topic. It is worth noting that labelling each bag of key terms from the modeling results can be a challenging task, as it requires researchers to subjectively grasp the main concepts represented by a set of key terms. To proceed the labelling process, both manual inspection and word frequency analytics are combined to formulate a set of conclusive topic labels.

Table 2. Topic modeling results

Topic	Key terms and weightings	Count	Topic label	Top 5 high loading documents
T1	0.011*"network" + 0.011*"data" + 0.008*"mobility" + 0.007*"study" + 0.007*"analysis" + 0.007*"gps" + 0.006*"urban" + 0.006*"transportation" + 0.006*"road" + 0.006*"time"	19	Networked mobilities	Liu & Engels, 2012; Wu et al., 2011; Eftelioglu et al., 2022; Li et al., 2022; Stipancic et al., 2018.
T2	0.015*"urban" + 0.013*"data" + 0.011*"activity" + 0.010*"city" + 0.010*"space" + 0.009*"behavior" + 0.008*"system" + 0.008*"public" + 0.007*"mobility" + 0.007*"human"	15	Human dynamics in spaces	Woo & Suh, 2020a; Hu et al., 2016; Xu, Santi, & Ratti, 2022; Szell, 2018; Hu et al., 2023.
T3	0.016*"data" + 0.012*"event" + 0.007*"model" + 0.006*"passenger" + 0.006*"travel" + 0.005*"functional" + 0.005*"bus" + 0.005*"traffic" + 0.005*"transport" + 0.005*"mode"	12	Event modeling	Cahill et al., 2022; Fan & Stewart, 2021; Cottrill, 2020; Aoki et al., 2021; Nguyen & Armoogum, 2020.
T4	0.023*"data" + 0.011*"spatial" + 0.011*"urban" + 0.009*"study" + 0.008*"social" + 0.008*"city" + 0.006*"mobility" + 0.006*"network" + 0.006*"big" + 0.006*"using"	43	Spatial underpinnings	Liu & Shi, 2016; Woo & Suh, 2020b; Ji et al., 2021; P. Wang et al., 2021; Nelson et al., 2021.
T5	0.012*"travel" + 0.010*"data" + 0.008*"activity" + 0.008*"time" + 0.007*"urban" + 0.007*"model" + 0.006*"traffic" + 0.006*"transportation" + 0.006*"user" + 0.006*"trip"	21	Travel behaviors and mobility patterns	Kang, 2019; Hu & Cheng, 2021; Herrera et al., 2016 Behrens & Newlands, 2022; Suciu et al, 2017.
T6	0.011*"data" + 0.009*"social" + 0.007*"city" + 0.007*"model" + 0.007*"activity" + 0.006*"people" + 0.006*"user" + 0.006*"study" + 0.005*"network" + 0.005*"transport"	28	Sociodemographic Heterogeneity	Sajeevan, 2019; Weth et al., 2017; Badji et al., 2021; Su et al., 2017; Zhao et al., 2019

Figure 4 (a-f) presents the visualized topics patterns and their most relevant terms. Each circle on the left represents one major topic detected, and their sizes reflect the proportions of the topics across the N total tokens in the given corpus. Accordingly, blue bars represent the overall frequency of each term in the corpus, while red bars represent the estimated number of times a given term was generated by a given topic. More technical details are available in Sievert and Shirley (2014).

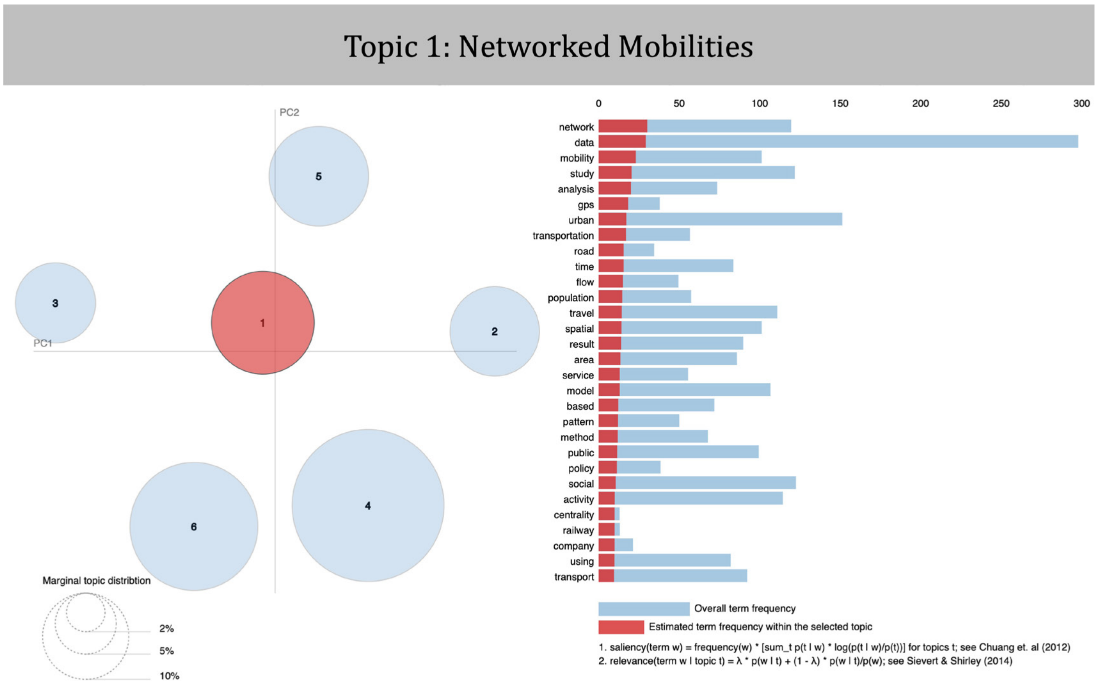

Figure 4a. Networked mobilities (Topic 1)

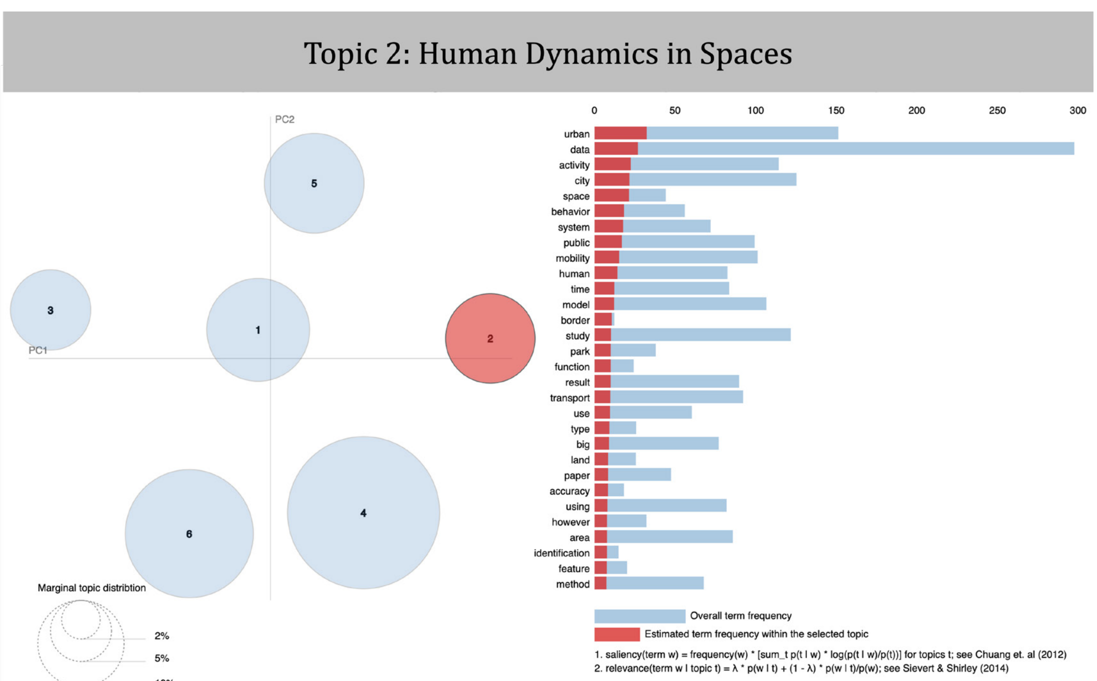

Figure 4b. Human dynamics in spaces (Topic 2)

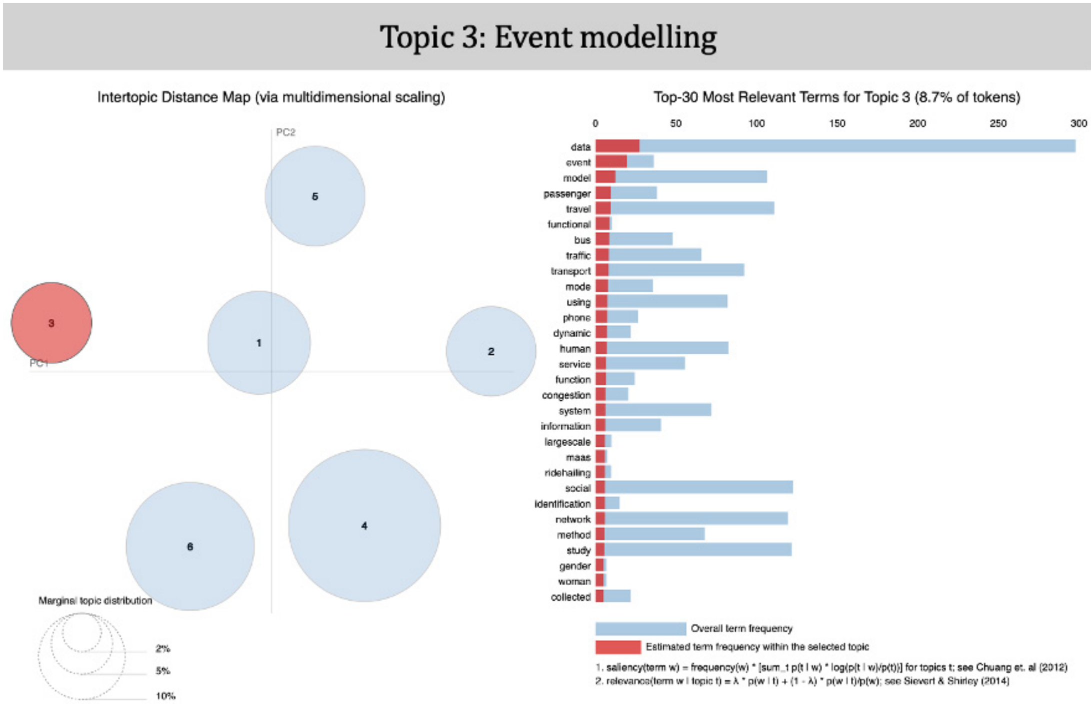

Figure 4c. Event modeling (Topic 3)

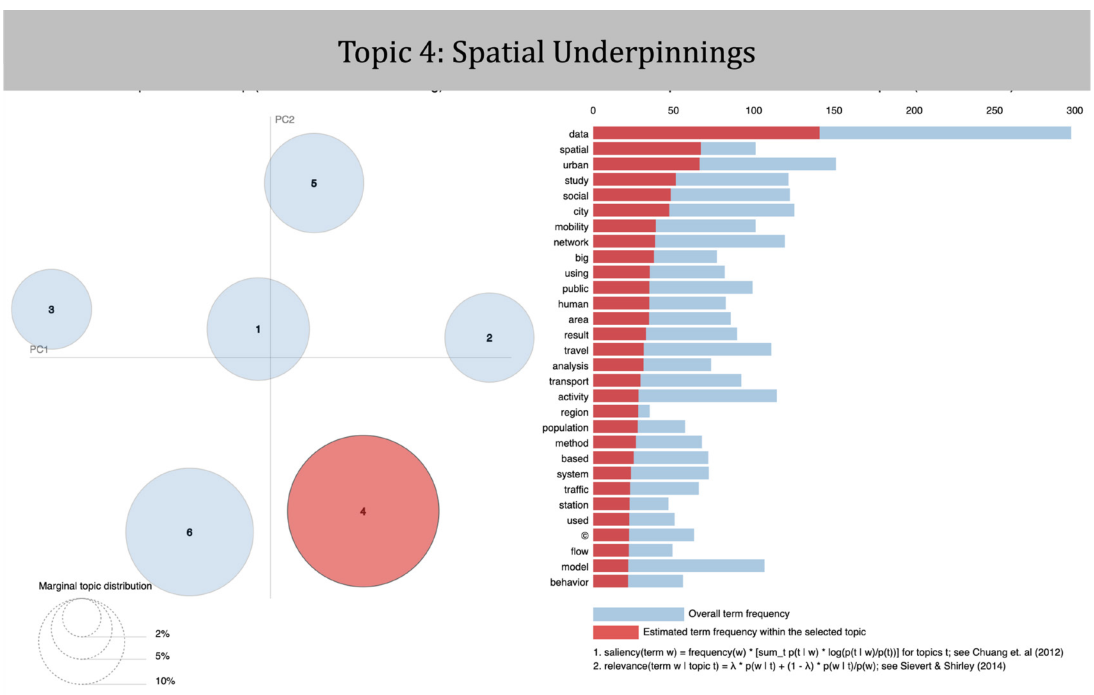

Figure 4d. Spatial underpinnings (Topic 4)

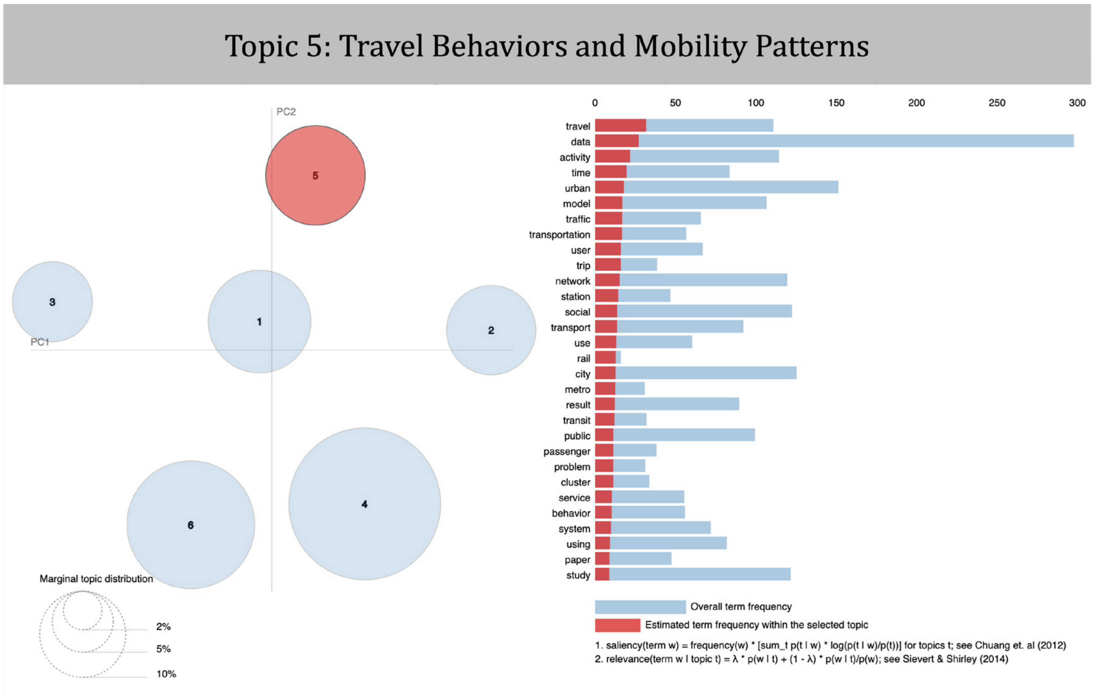

Figure 4e. Travel behaviors and mobility patterns (Topic 5)

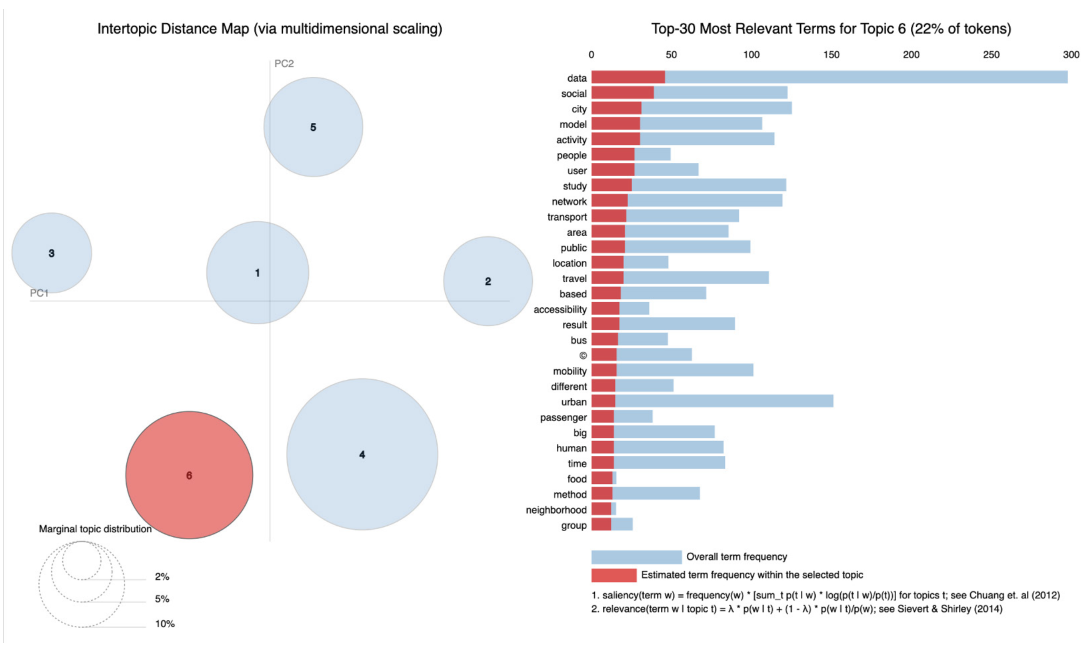

Figure 4f. Sociodemographic heterogeneity (Topic 6)

Among the six detected topics, *Networked mobilities* (T1) reflects a wide range of research interests in utilizing emerging big data sources to recognize and analyze the complex transport networks, thus, it is labelled *Networked mobilities*, with 14.2% of tokens captured under this topic. Representative terms include "network," "flow," "centrality," "railway," etc. Various transport modes and their derived networks have been discussed, including public transit (Liu & Engels, 2012), railway (Li et al., 2022), and innovative shared mobilities (Zhang et al., 2021). Research papers in Topic 2, which we labelled *Human dynamics* in spaces, mainly discuss the details and features of human travel activities, including travel mode, daily ridership, temporal dynamics, and transit usage. This stream of studies are conducted in various case cities across the globe, such as Beijing, China (D. Xu et al., 2022), Seoul, Korea (Park & Chang, 2020), and Rio de Janeiro (Costa, 2022), etc. *Event modeling* (T3) is rather a niched research area that represents thematic studies on the early detection of major events from mobility activities (Aoki et al., 2021) or human's collective movement patterns during large-scale events (Fan & Stewart, 2021). Papers categorized in Topic 4 are labelled as *Spatial underpinnings*, which deal with the geographical infrastructure's impact on urban mobility and associated spatial equality. Representative terms include "spatial," "network," "area," "region," "station," etc. Topic 4 is also the largest circle visualized in Figure 4, as 31% of tokens are captured under this topic, reflecting a long-lasting research interests on this topic. Topic 5 is labelled as *Travel behaviors and mobility patterns*, and papers contributing to this topic mainly analyze the spatial and temporal patterns of travel activities and the derived collective movement patterns are discussed; representative terms include "time," "trip," "user," "passenger," "cluster," "service," etc. This stream of studies highlight the common usage of big data sources in assessing public transit usage (Deng & Zhao, 2022; Hu & Cheng, 2021; Sohn, 2016). At last, Topic 6 represents thematic studies on Sociodemographic Heterogeneity, and it covers a wide range of research topics exploring socio-demographic differences arising from human mobility. As an illustration, a notable study examined the correlation between public transit accessibility and healthcare utilization among individuals aged 18-60, both with and without disabilities, in Australia (Badji et al., 2021). The findings of this study emphasized that enhancing public transport availability could potentially mitigate certain obstacles to healthcare access experienced by individuals with disabilities.

In sum, the visualized LDA results showcased an efficient way to classify major research streams with a given corpus, and provide a powerful means to comprehend and interpret the outcomes of the analysis by offering an intuitive representation of the main topics, and highlighting connections between topics. Obviously, some detected topics are proximate to one another and even share some top relevant terms, indicating a stronger logical linkage and shared research interests between these two separate studies identified by the LDA.

5 Discussion

5.1 Research question 1: Big data's role in urban mobility studies' evolution

The first proposed research question aims to reveal whether and to what degree has big data helped us better reveal the social dimensions of urban mobility across different disciplines. More specifically, what kind of related studies are more significantly accelerated by the usage of big data across the review period. It is evident that big data's influence on human mobility studies is evolving as it develops, and impact different research domains from time to time. Table 3 summarizes count of papers of each detected topics from 2012 to 2023. The darker the red, the greater number of papers contributing to that topic has been identified in the specific year.

Table 3. Topic evolutions from 2010 to 2023

	Topic1	Topic2	Topic3	Topic4	Topic5	Topic6
2010	0	0	0	0	0	0
2011	1	0	0	0	0	0
2012	1	0	0	0	0	0
2013	0	0	0	1	0	1
2014	0	1	0	1	1	0
2015	1	0	0	1	1	2
2016	0	2	0	3	2	0
2017	0	2	0	2	2	3
2018	3	1	2	2	2	2
2019	1	2	4	7	4	6
2020	4	3	0	7	0	0
2021	2	2	1	10	5	2
2022	6	2	2	3	3	4
2023	1	1	0	3	2	0

The presence of big data application in the scholarship on human mobility started in the early-2010s, with a focus on uncovering mobility network attributes (Topic 1) and spatial underpinnings of human movement (Topic 4) at its early stage. For example, Wu et al. (2011) utilized global positioning system (GPS) to collect time-location data of participants' exposure to air pollutions in various scenarios; Wang et al. (2013) analyzed the complex behavior of 10,000 frequent users of Location Based Social Networks (LBSNs) and examined the recommendation system based on user's temporal regularities. At this stage, major big data sources employed in these studies were either GPS or mobile phone signal datasets, trying to extract mobility patterns by analyzing the spatial and temporal patterns of different human activities, namely, GPS trajectories of invited volunteers or Twitter footprints of specific social groups. Later on, *Networked mobilities* (T1) shows a more consistent presence in the research domain, with a peak in 2022, and *Spatial Underpinnings* (Topic 4) presents a trending research direction since 2019 . A wider range of emerging big data sources have been employed, and the identification of precise human activities or assessment of derived mobility networks are the dominant topics. For example, Eftelioglu et al. (2022) proposed a novel deep neural network architecture is to infer modalities of GPS points in a trajectory with frequent activity changes. Zhang et al. (2021) examined the network of bike sharing in Shenzhen, China, and found the specific networks are assortative and positive autocorrelated with evident communities.

Human dynamics in spaces (Topic 2), *Travel behavior and mobility patterns* (Topic 5), and *Sociodemographic Heterogeneity* (Topic 6) initially show a rather modest presence in early 2010s, and then increasingly attract research interest since 2017. These three topics share an overlapped interests in the direction to understand mobility as an important metric for evaluating spatial inequity, as well as policy tool for mitigating inequity in future development. For *Human dynamics in spaces* (Topic 2), some scholars examined accessibility to care facilities for ageing people (Yoon & Park, 2022), and some utilized the crowdsourced data to validate the prevalent imbalance between different transport modal share and urban surface space allocation in 23 cities across the globe (Szell, 2018); another interesting research interests lie in the specific context of Coronavirus of 2019, rising research interest has been seen to uncover the social disparities that are (sometimes implicitly) reflected in mobility behaviors, especially under an uncertain and risky context rendered by the pandemic (Hasselwander et al., 2021). *Travel behavior and* *mobility patterns* (Topic 5) presents a view to utilize human mobility indicators as proxies to understand the spatial inequity of accessibility for urbanites with different levels of resources. Qiao et al. (2023) used mobile phone trajectories of gig drivers working for the ride-hailing platform and mapped their residential locations. Findings suggest that ride-hailing drivers demonstrate a higher propensity to originate from neighborhoods characterized by lower income levels and limited access to stable employment opportunities. Similarly, Chen, Yan, et al. (2021) conducted an investigation on the "last mile problem" utilizing the Google Map API as a tool. Their findings revealed a clustering of "last mile problem" areas in economically disadvantaged regions of Chicago. Moreover, the study identified positive associations between income levels, housing sale prices, and last mile performance scores. *Sociodemographic Heterogeneity* (Topic 6) reflects a new understanding of sociodemographic heterogeneity as innovative mobility options emerged, and goes further to capture the unequal access to various resources and opportunities across social groups, for example, food (Su et al., 2017), health care facilities (Badji et al., 2021), public transit (Kandt & Leak, 2019), etc. Notably, the big-data-driven urban mobility studies in these research domains reflected a deeper linkage between empirical applications and classic geography theories.

Among the diverse range of topics under scrutiny, *Event modeling* (Topic 3) emerges as a distinctive and specialized realm of inquiry. It involves the meticulous examination of human movement patterns during specific events, including sports events, religious ceremonies, festivals, even natural disasters. With the assistance of five-month long bus data collected from Beijing, China, Aoki et al. (2018) proposed a framework to detect mega events in advance without incurring any privacy invasion. Compared to traditional small datasets, big datasets collected from long-span continuous vehicles are typically recorded in real-time or near real-time, enabling researchers to analyze human movement patterns as events unfold. Such important feature facilitates timely decision-making, rapid response planning, and the ability to adapt strategies based on evolving situations. Nevertheless, the current body of literature in this stream of studies remains scarce, albeit imbued with optimistic prospects in practical transport planning and management.

To sum up, we have witnessed an evolution of research topics and an enhancement of analytical tools employed over time. Inevitably, the corpus analyzed in our paper could not reflect the full scope of human mobility studies in the past decade due to the limited sample size; for example, some important topics were not detected by conducting the LDA, including but not limited to gender inequalities in mobility (Gauvin et al., 2020), health services (Fu et al., 2021), education (Moreno-Monroy et al., 2018), and different public facilities (Tahmasbi et al., 2019), and mobility as a service (MaaS) and related policy studies (Butler et al., 2021). Through this section, it is hoped that studies from various disciplines could benefit from our work and further enlarge the research scope by taking advantage of the assistance of advanced analytical tools and multi-sourced big datasets. Together, they contribute more to a more holistic and up to date understanding of efficient, inclusive, green, and sustainable urban mobility.

5.2 Research question 2: Opportunities and limitations

To date, a series of review papers have examined the big data applications in the existing scholarship on human mobility, for example, Tao et al. (2021) classified big data utilization into five categories: 1) data enrichment and travel behavior mining, 2) travel behavior variability, 3) transport system assessment and planning, 4) evaluating urban structure and function, and 5) environmental and social implications; Wang et al. (2022) attempted to understand cities by reconceptualizing human mobility as spatial movements, a social phenomenon, a policy tool, and an indicator for economic activities; and more reviews are conducted by focusing on a specific topical area, for instance, public transportation (Welch and Widita, 2019), urban transport equity (Zhang & Zhao, 2021), smart city development (A. Wang et al., 2021), data mining techniques (Zhao et al., 2016), and Covid-19 related analytics (Hu, Xiong, et al., 2021). Few papers zoomed into the social dimensions of human mobility, although mobility itself is an important slice of urban daily life, and creates differences everywhere and every time (Kaufmann, 2014). Based on the above review, this section aims to respond to the previously proposed research questions that what are the opportunities and limitations faced by the identified papers, and which are a sample of the increasingly enlarging body of literature on social dimensions of urban mobility facilitated by big data?

5.2.1 Big data as efficiency advancement

As discussed earlier, the massive, fine-granularized, time-stamped, and geocoded big data not only help scholars to validate classic concepts investigated by traditional surveys and observations before, but also enlarge the research scope by supporting some research questions previously considered not possible, such as the revelation of global multi-layer network of human mobility (Belyi et al., 2017) or the detailed substructures and spatiotemporal flows characterizing human mobility (Bagrow & Lin, 2012). From the information richness theory's viewpoint, big data enriches human mobility research by providing a more comprehensive and granular understanding of individuals' mobility patterns, however, it also emphasizes the importance of matching the richness of the communication channel with the complexity and ambiguity of the information.

More specifically in urban mobility research context, some research areas are naturally more data-sensitive and could accumulate more advantages as the data sources improve. For example, network studies in the small-data era heavily rely on exquisite experimental design, such as the small-world experiment conducted by sociologist Stanley Milgram to examine the average path length and social networks in the United States (Travers & Milgram, 1967). Compared to the original research procedures, which was tracking and analyzing the postcards mailing between Nebraska and Kansa, the big data-driven replication of it involved over 24,000 emails chains and extended the geographic scope to the entire earth (Dodds et al., 2003). Zoomed into more recent studies, the network attributes of human mobility patterns attracted great research interest and contribute greatly to accelerating interdisciplinary dialogues (Alessandretti et al., 2020; Sun et al. 2013;). On the contrary, some research areas focus on the mobility attributes that could hardly be captured by big data, for example, the subjective perception in a transit environment, the individual life event changes, or the socio-economic features of individuals. Therefore, the significance of small data in the era of big data remains irreplaceable, and more big-and-small data integrations are expected in human mobility studies (Chen et al., 2016).

Furthermore, some innovative transport modes are intrinsically data-driven. The boosting of shared mobility since 2017 has attracted scholars' interest and promoted a vast number of relevant studies. Due to the nature of modern shared mobility, every action of the user is recorded digitally, and the companies exploit data analytics as a key tool to enhance their services. On the one hand, a series of works on shared mobility behaviors emerged (Raux et al., 2017; Saberi et al., 2018; Zhou, 2015); on the other hand, privacy concerns over shared mobility data usage are flagged. Section 5.2.3 will further discuss such limitations and pitfalls of big data applications.

5.2.2 Big data as equity lens

Alongside technological advancement, big data provide new avenues for our understanding of societal fabric of cities through the lens of equity. In Section 4, three equity-concerned research topics were detected from the corpus; they are Human dynamics in spaces (T2), Travel behaviors and mobility patterns (T5), and Sociodemographic Heterogeneity (T6). This reflects increasing research efforts by authors to reveal the inequality in various forms of urban mobility and potential mitigations. The major contributions of big data in equity-oriented mobility research are twofold.

Firstly, big data enable scholars to unravel those "invisible groups" that could hardly be investigated in "small" data from traditional sources. For example, extreme commenters, who have particularly burdensome commutes, can imply the existence of social and geographic inequalities. These specific groups might account for a minor share of the entire commuting population, which means they could possibly be neglected in traditional travel surveys. However, understanding their behaviors is of great importance in revealing linkages between mobility and well-being, and mental health. With the assistance of a traditional travel survey and census data, researchers identified a series of explanatory factors of extreme commuters, such as sex, age, occupation type, mode of transportation, migration, employment status, etc. (Maoh & Tang, 2012). In a later study using one-week smartcard data of Beijing transit system, Long et al. (2016) identified 188.9 thousands of riders who have at least one pre-defined extreme travel behaviors. On the one hand, their findings are consistent with previous studies using traditional travel surveys; namely, the ratio of extreme commuters to total commuters (Rapino & Fiedls, 2013) and the explanatory socioeconomic attributes for those behaviors (Maoh & Tang, 2012). On the other hand, the application of big data enabled Long et al. (2016) to depict the detailed spatiotemporal patterns of the extreme commuters in a one-week timeframe and detected the job and housing places of them. Besides extreme commuters, big data also contributed to the study on other disadvantaged groups, such as the seniors (Kandt & Leak, 2019; Liu et al., 2021), the disabled (Ferrari et al., 2014), and the general vulnerable population (Zhang et al., 2021). In sum, big data enable researchers to identify, detect, and understand more subgroups, which can deepen our understanding of transport justice.

Secondly, the trajectories of individuals reflect their "real use of the city" (Ahas et al. 2010), and empower a series of transport performance assessments, transportation demand management (TDM), real-time simulation, and policy construction in the short or long term. g et al. (2022) summarized the "mobility-as-a-policy-tool" studies in four categories: monitoring and mitigating traffic anomaly, understanding and predicting urban crimes, understanding and relieving disaster impacts, and predicting and mitigating the spread of pandemics. This series of studies examined mobility as a source of improving potential urban practices, and a majority of them have discussed equity issues in their research contexts. For instance, human mobility data serve as the key input in understanding evacuation movements after natural disasters (Solmaz & Turgut, 2017), post-disaster recovery in cities (Yabe et al., 2020), and vulnerability across different social groups (Martin et al., 2020). Similarly, a series of studies have been conducted in the context of Covid-19 to uncover the heterogeneous adaptions and responses of various social atoms and groups (Loo et al., 2021; Pan & He, 2022; Zhou et al., 2021), and contribute to more efficient disease containment strategies and more inclusive policies.

5.2.3 Limitations and opportunities

Given the above-mentioned progresses, multi-source we could tell that big data have greatly unleashed the research potential to better understand various social dimensions in human mobility in the past decade or so. Although the contributions it brings to different scholars and practitioners are heterogeneous, big data have generally promoted interdisciplinary dialogues and given rise to a better understanding of urban social fabric through the lens of urban mobility.

As noted by Chen et al. (2016), the validation and representativeness of big data remain unresolved in the research domain, and we further argue that these problems have not been adequately addressed to date. For the majority of big-data-driven human mobility studies we identified and reviewed, the big data utilized were aggregated at different levels, which therefore poses the danger of making an ecological mistake and causes problems in integration with other big or small datasets (Chen et al., 2016). Furthermore, even though the size and amount of such mobility datasets could be massive, they could only represent the groups using a particular transport mode (e.g., subways or shared bikes) or living in a specific region (e.g., urban center or suburban areas); therefore, biased sampling remains challenging to be addressed in future studies. Besides, it's worth noting that the boom of big data-facilitated research could possibly accelerate digital inequality in academia, especially for regions with weak digital infrastructure, e.g., the global south. As the importance of digital infrastructure in daily lives and urban practices increases, it generates profound implications for related human mobility studies; therefore, such awareness of the inequality and divide in digital infrastructure across the globe could be beneficial to create a more inclusive research agenda. Finally, privacy has been constantly discussed in related studies and scholars argued the application and employment of big data should be used for "social good" (Poom et al., 2020).

6 Conclusions

Big-data-driven human mobility has attracted considerable research interest in the past decade across various academic fields. This paper has systematically reviewed the social dimensions discussed in this arena by conducting a keyword-driven literature selection and the LDA method on the corpus collected from two widely used academic databases (Web of Science and Scopus). The results suggest six major topics and assist in re-conceptualizing the social demission reflected in human movement patterns, including Networked mobilities, Human dynamics in spaces, Event modeling, Spatial underpinnings, Travel behaviors and mobility patterns, and Sociodemographic Heterogeneity.

Through the review, an evolution of focused research topics from time to time has been observed. As the analytical techniques developed and the understanding of the societal fabric of human mobility deepened, the nexus of the studies evolved accordingly. Drawing upon the evolution of research topics, this paper identifies Networked Mobilities as a resilient and enduring area of research, providing a solid foundation for future studies on human mobility. In addition, Spatial underpinnings and Travel behaviors and mobility patterns provides valuable insights into the intersection of social justice and human mobility. These topics have emerged as prominent areas of research, reflecting the current trends in scholarly inquiry. Among the various topics explored, Event modeling stands out as a specialized area of investigation. This particular field delves into the intricate analysis of human movement patterns, leveraging the power of big data. The implications derived from such analyses hold significant practical relevance.

The review also identifies a few prospective directions that call for additional work. First, active transport (e.g., walking or bicycling) and the linkages between active human mobility and health, well-being, and a sustainable environment deserve more focus. Due to the data collection difficulty, few studies have been seen to explore such relationships with the assistance of big data. Recently, several works have attempted to understand pedestrian's satisfaction with traffic flow by using Internet of Things connected sensors (Carter et al., 2020), or analyzing pedestrian's route preferences with the assistance of walking trajectories collected from an activity-based smartphone application (Sevtsuk et al., 2021). These explorations serve as a good starting point for more work in this field. Furthermore, more integrative analysis of multimodal transport modes is expected. To date, most empirical studies rely on a single source of mobility datasets and data fusion across different transport modes or analytical scales is still scarce. Lastly, understanding attitudinal factors and individual perceptions of travel behaviors serve as a promising field for better transport service operation and policy making (van Acker et al., 2010). Text analysis on travel blogs (Haris & Gan, 2021; Wang et al., 2019), image recognition (Desjardins et al., 2021), and internet content related to transport services analysis (Sarram & Ivey, 2022) are considered as some beneficial attempts at this stage.

When looking back, we can see a multidisciplinary discussion attempting to elucidate the many socioeconomic components of human movement patterns. It is hoped that this article will serve as a starting point for academics, planners, policymakers, and professionals who are interested in collective efforts on a more supportive, inclusive, and sustainable urban living environment.

Acknowledgments

The research underlying this study is financially supported by the Platform Technology Funding (URC012530226) and Guangdong-Hong Kong-Macau Joint Laboratory Program (Project #: 2020B1212030009). We also thank three anonymous reviewers for their insightful comments and suggestions, which have enhanced the paper in multiple dimensions.

Appendix

Appendix available as a supplemental file at https://jtlu.org/index.php/jtlu/article/view/2281.

References

Ahas, R., Silm, S., Järv, O., Saluveer, E., & Tiru, M. (2010). Using mobile positioning data to model locations meaningful to users of mobile phones. *Journal of Urban Technology.* https://doi.org/10.1080/10630731003597306

Aiden, E., & Michel, J.-B. (2013). *Uncharted: Big data as a lens on human culture*. New York: Penguin, Riverhead.

Alessandretti, L., Aslak, U., & Lehmann, S. (2020). The scales of human mobility. *Nature.* https://doi.org/10.1038/s41586-020-2909-1

Aoki, S., Sezaki, K., Yuan, N. J., & Xie, X. (2018). Busbeat: Early event detection with real-time bus GPS trajectories. *IEEE Transactions on Big Data, 7*(2), 371–382.

Asmussen, C. B., & Møller, C. (2019). Smart literature review: A practical topic modelling approach to exploratory literature review. *Journal of Big Data.* https://doi.org/10.1186/s40537-019-0255-7

Badji, S., Badland, H., Rachele, J. N., & Petrie, D. (2021). Public transport availability and healthcare use for Australian adults aged 18–60 years, with and without disabilities. *Journal of Transport & Health, 20,* 101001.

Bagrow, J. P., & Lin, Y. R. (2012). Mesoscopic structure and social aspects of human mobility. *PLoS ONE.* https://doi.org/10.1371/journal.pone.0037676

Bastani, K., Namavari, H., & Shaffer, J. (2019). Latent Dirichlet allocation (LDA) for topic modeling of the CFPB consumer complaints. *Expert Systems with Applications.* https://doi.org/10.1016/j.eswa.2019.03.001

Behrens, R., & Newlands, A. (2022). Revealed and future travel impacts of COVID-19 in sub-Saharan Africa: Results of big data analysis and a Delphi panel survey. *Journal of Transport and Supply Chain Management, 16,* 758.

Belyi, A., Bojic, I., Sobolevsky, S., Sitko, I., Hawelka, B., Rudikova, L., Kurbatski, A., & Ratti, C. (2017). Global multi-layer network of human mobility. *International Journal of Geographical Information Science.* https://doi.org/10.1080/13658816.2017.1301455

Birenboim, A., Reinau, K. H., Shoval, N., & Harder, H. (2015). High-resolution measurement and analysis of visitor experiences in time and space: The case of Aalborg Zoo in Denmark. *Professional Geographer.* https://doi.org/10.1080/00330124.2015.1032874

Blau, P. M. (1977). A Macrosociological theory of social structure. *American Journal of Sociology.* https://doi.org/10.1086/226505

Blei, D. M. (2012). Introduction to probabilistic topic models. *Communications of the ACM, 55*(4), 77–84.

Böcker, L., Anderson, E., Uteng, T. P., & Throndsen, T. (2020). Bike sharing use in conjunction to public transport: Exploring spatiotemporal, age and gender dimensions in Oslo, Norway. *Transportation Research Part A: Policy and Practice.* https://doi.org/10.1016/j.tra.2020.06.009

Briand, A. S., Côme, E., Trépanier, M., & Oukhellou, L. (2017). Analyzing year-to-year changes in public transport passenger behavior using smart card data. *Transportation Research Part C: Emerging Technologies.* https://doi.org/10.1016/j.trc.2017.03.021

Brockmann, D., Hufnagel, L., & Geisel, T. (2006). The scaling laws of human travel. *Nature.* https://doi.org/10.1038/nature04292

Buhl, H. U., Röglinger, M., Moser, F., & Heidemann, J. (2013). Big data: A fashionable topic with(out) sustainable relevance for research and practice? *Business and Information Systems Engineering*. https://doi.org/10.1007/s12599-013-0249-5

Burgess, E. W. (1926). The growth of the city. In R. Park., & E. W Burgess (Eds), *The city.* Chicago: The University of Chicago Press.

Butler, L., Yigitcanlar, T., & Paz, A. (2021). Barriers and risks of mobility-as-a-service (MaaS) adoption in cities: A systematic review of the literature. *Cities.* https://doi.org/10.1016/j.cities.2020.103036

Cahill, R. A., Mac Aonghusa, P., & Mortensen, N. (2022). The age of surgical operative video big data: My bicycle or our park? *The Surgeon, 20*(3), e7–e12.

Calabrese, F, Smoreda, Z., Blondel, V. D., & Ratti, C. (2011). Interplay between telecommunications and face-to-face interactions: A study using mobile phone data. *PLoS ONE.* https://doi.org/10.1371/journal.pone.0020814

Candipan, J., Phillips, N. E., Sampson, R. J., & Small, M. (2021). From residence to movement: The nature of racial segregation in everyday urban mobility. *Urban Studies.* https://doi.org/10.1177/0042098020978965

Carrasco, J. A., Hogan, B., Wellman, B., & Miller, E. J. (2008). Collecting social network data to study social activity-travel behavior: An egocentric approach. *Environment and Planning B: Planning and Design.* https://doi.org/10.1068/b3317t

Carter, E., Adam, P., Tsakis, D., Shaw, S., Watson, R., & Ryan, P. (2020). Enhancing pedestrian mobility in smart cities using big data. *Journal of Management Analytics.* https://doi.org/10.1080/23270012.2020.1741039

Chen, C., Ma, J., Susilo, Y., Liu, Y., & Wang, M. (2016). The promises of big data and small data for travel behavior (aka human mobility) analysis. *Transportation Research Part C: Emerging Technologies.* https://doi.org/10.1016/j.trc.2016.04.005

Chen, S., Chen, Y., Lei, Z., & Tan-Soo, J.-S. (2021). Chasing clean air: Pollution-induced travels in China. *Journal of the Association of Environmental and Resource Economists.* https://doi.org/10.1086/711476

Chen, S., Yan, X., Pan, H., & Deal, B. (2021). Using big data for last mile performance evaluation: An accessibility-based approach. *Travel Behavior and Society, 25,* 153–163.

Cheng, Z., Trépanier, M., & Sun, L. (2021). Probabilistic model for destination inference and travel pattern mining from smart card data. *Transportation.* https://doi.org/10.1007/s11116-020-10120-0

Cheng, Z., Caverlee, J., Lee, K., & Sui, D. Z. (2011). Exploring millions of footprints in location sharing services. *Icwsm.* https://doi.org/papers3://publication/uuid/0C46BD5D-4908-4A8A-BD06-5BCB2F1DE282

Cho, E., Myers, S. A., & Leskovec, J. (2011). Friendship and mobility: User movement in location-based social networks. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.* https://doi.org/10.1145/2020408.2020579

Chudyk, A. M., Winters, M., Moniruzzaman, M., Ashe, M. C., Gould, J. S., & McKay, H. (2015). Destinations matter: The association between where older adults live and their travel behavior. *Journal of Transport and Health.* https://doi.org/10.1016/j.jth.2014.09.008

Costa, A. (2022). *Intelligent urban mobility systems in Rio de Janeiro: A critical assessment.* CUADERNOS DE VIVIENDA Y URBANISMO. DOI10.11144/Javeriana.cvu15.simu

Cottrill, C. D. (2020). MaaS surveillance: Privacy considerations in mobility as a service. *Transportation Research Part A: Policy and Practice, 131,* 50–57.

Cranshaw, J., Toch, E., Hong, J., Kittur, A., & Sadeh, N. (2010). Bridging the gap between physical location and online social networks. *UbiComp'10 - Proceedings of the 2010 ACM Conference on Ubiquitous Computing.* https://doi.org/10.1145/1864349.1864380

Crompton, R. (2008) 40 years of sociology: Some comments. *Sociology 42*(6), 1218–1227.

Cwerner, S., Kesselring, S., & Urry, J. (2008). *Aeromobilities.* https://doi.org/10.4324/9780203930564

Davidson, J. H., & Ryerson, M. S. (2021). Modeling regional disparity and the reverse commute. *Transportation Research Part A: Policy and Practice.* https://doi.org/10.1016/j.tra.2021.06.005

Deng, Y., & Zhao, P. (2022). The impact of new metro on travel behavior: Panel analysis using mobile phone data. *Transportation Research Part A: Policy and Practice, 162,* 46–57.

Desjardins, E., Higgins, C. D., Scott, D. M., Apatu, E., & Páez, A. (2021). Using environmental audits and photo-journeys to compare objective attributes and bicyclists' perceptions of bicycle routes. *Journal of Transport and Health.* https://doi.org/10.1016/j.jth.2021.101092

Dodds, P. S., Muhamad, R., & Watts, D. J. (2003). An experimental study of search in global social networks. *Science.* https://doi.org/10.1126/science.1081058

Eftelioglu, E., Wolff, G., Nimmagadda, S. K. T., Kumar, V., & Chowdhury, A. R. (2022, November). Deep classification of frequently changing activities from GPS trajectories. *Proceedings of the 15th ACM SIGSPATIAL International Workshop on Computational Transportation Science,* 1-10.

Elliott, A., & Urry, J. (2010). *Mobile Lives.* https://doi.org/10.4324/9780203887042

Etter, V., Kafsi, M., Kazemi, E., Grossglauser, M., & Thiran, P. (2013). Where to go from here? Mobility prediction from instantaneous information. *Pervasive and Mobile Computing.* https://doi.org/10.1016/j.pmcj.2013.07.006

Fan, J., & Stewart, K. (2021). Understanding collective human movement dynamics during large-scale events using big geosocial data analytics. *Computers, Environment and Urban Systems, 87,* 101605.

Ferrari, L., Berlingerio, M., Calabrese, F., & Reades, J. (2014). Improving the accessibility of urban transportation networks for people with disabilities. *Transportation Research Part C: Emerging Technologies.* https://doi.org/10.1016/j.trc.2013.10.005

Fu, L., Xu, K., Liu, F., Liang, L., & Wang, Z. (2021). Regional disparity and patients mobility: Benefits and spillover effects of the spatial network structure of the health services in China. *International Journal of Environmental Research and Public Health.* https://doi.org/10.3390/ijerph18031096

Gan, J., & Qi, Y. (2021). Selection of the optimal number of topics for LDA topic model—taking patent policy analysis as an example. *Entropy, 23*(10), 1301.

Gauvin, L., Tizzoni, M., Piaggesi, S., Young, A., Adler, N., Verhulst, S., Ferres, L., & Cattuto, C. (2020). Gender gaps in urban mobility. *Humanities and Social Sciences Communications.* https://doi.org/10.1057/s41599-020-0500-x

Geng, W., & Yang, G. (2017). Partial correlation between spatial and temporal regularities of human mobility. *Scientific Reports.* https://doi.org/10.1038/s41598-017-06508-1

Gkiotsalitis, K., & Stathopoulos, A. (2016). Joint leisure travel optimization with user-generated data via perceived utility maximization. *Transportation Research Part C: Emerging Technologies.* https://doi.org/10.1016/j.trc.2016.05.009

Goldthorpe, J. H. (2015). Sociology as a population science. *Sociology as a Population Science.* https://doi.org/10.1017/cbo9781316412565

González, M. C., Hidalgo, C. A., & Barabási, A. L. (2008). Understanding individual human mobility patterns. *Nature.* https://doi.org/10.1038/nature06958

Grisé, E., Boisjoly, G., Maguire, M., & El-Geneidy, A. (2019). Elevating access: Comparing accessibility to jobs by public transport for individuals with and without a physical disability. *Transportation Research Part A: Policy and Practice.* https://doi.org/10.1016/j.tra.2018.02.017

Halford, S., & Savage, M. (2017). Speaking sociologically with big data: Symphonic social science and the future for big data research. *Sociology.* https://doi.org/10.1177/0038038517698639

Haraguchi, M., Nishino, A., Kodaka, A., Allaire, M., Lall, U., Kuei-Hsien, L., ... & Kohtake, N. (2022). Human mobility data and analysis for urban resilience: A systematic review. *Environment and Planning B: Urban Analytics and City Science.* https://doi.org.10.1177/23998083221075634

Haris, E., & Gan, K. H. (2021). Extraction and visualization of tourist attraction semantics from travel blogs. *ISPRS International Journal of Geo-Information*. https://doi.org/10.3390/ijgi10100710

Hasan, S., Schneider, C. M., Ukkusuri, S. V., & González, M. C. (2013). Spatiotemporal patterns of urban human mobility. *Journal of Statistical Physics.* https://doi.org/10.1007/s10955-012-0645-0

Hasan, S., & Ukkusuri, S. V. (2014). Urban activity pattern classification using topic models from online geo-location data. *Transportation Research Part C: Emerging Technologies.* https://doi.org/10.1016/j.trc.2014.04.003

Hasan, S., Zhan, X., & Ukkusuri, S. V. (2013). Understanding urban human activity and mobility patterns using large-scale location-based data from online social media. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.* https://doi.org/10.1145/2505821.2505823

Hasselwander, M., Tamagusko, T., Bigotte, J. F., Ferreira, A., Mejia, A., & Ferranti, E. J. (2021). Building back better: The COVID-19 pandemic and transport policy implications for a developing megacity. *Sustainable Cities and Society, 69,* 102864.

Herrera, N. H., Santamaría, H. S., Macías, M. M., & Gómez, E. (2016, March). Analysis of the factors generating vehicular traffic in the city of Quito and its relation to the application of sensorial and social data with big data as a basis for decision making. *2016 Third International Conference on eDemocracy and eGovernment,* 133–137.

Howe, E. G., & Elenberg, F. (2020). Ethical challenges posed by big data. *Innovations in Clinical Neuroscience, 17*(10–12), 24–30.

Hu, J., Gao, Y., Wang, X., & Liu, Y. (2023). Recognizing mixed urban functions from human activities using representation learning methods. International Journal of Digital Earth, 16(1), 289–307.

Hu, M., Li, W., Li, L., Houston, D., & Wu, J. (2016). Refining time-activity classification of human subjects using the global positioning system. *PloS one, 11*(2), e0148875.

Hu, S., Gao, S., Wu, L., Xu, Y., Zhang, Z., Cui, H., & Gong, X. (2021). Urban function classification at road segment level using taxi trajectory data: A graph convolutional neural network approach. *Computers, Environment and Urban Systems.* https://doi.org/10.1016/j.compenvurbsys.2021.101619

Hu, S., Xiong, C., Yang, M., Younes, H., Luo, W., & Zhang, L. (2021). A big-data driven approach to analyzing and modeling human mobility trend under non-pharmaceutical interventions during COVID-19 pandemic. *Transportation Research Part C: Emerging Technologies.* https://doi.org/10.1016/j.trc.2020.102955

Hu, W., & Cheng, F. (2021). Application research of urban subway traffic mode based on behavior entropy in the background of big data. *Journal of High-Speed Networks, 27*(3), 291–304.

Huang, W., Li, S., Liu, X., & Ban, Y. (2015). Predicting human mobility with activity changes. *International Journal of Geographical Information Science.* https://doi.org/10.1080/13658816.2015.1033421

Huang, Z., Ling, X., Wang, P., Zhang, F., Mao, Y., Lin, T., & Wang, F. Y. (2018). Modeling real-time human mobility based on mobile phone and transportation data fusion. *Transportation Research Part C: Emerging Technologies.* https://doi.org/10.1016/j.trc.2018.09.016

Jacobs, J. (1961). *The death and life of great American cities.* New York: Random House.

Järv, O., Müürisepp, K., Ahas, R., Derudder, B., & Witlox, F. (2015). Ethnic differences in activity spaces as a characteristic of segregation: A study based on mobile phone usage in Tallinn, Estonia. *Urban Studies.* https://doi.org/10.1177/0042098014550459

Jensen, O. B. (2013). Staging mobilities. *Staging Mobilities.* https://doi.org/10.4324/9780203070062

Ji, H., Ci, Z., Jiang, L., Li, G., & Li, B. (2021). Analysis of spatial inequality in taxi ride and its relationship with populations structure.

Jiang, S., Ferreira, J., & Gonzalez, M. C. (2016). Activity-based human mobility patterns inferred from mobile phone data: A case study of Singapore. *IEEE Transactions on Big Data.* https://doi.org/10.1109/tbdata.2016.2631141

Jung, J., & Sohn, K. (2017). Deep-learning architecture to forecast destinations of bus passengers from entry-only smart-card data. *IET Intelligent Transport Systems.* https://doi.org/10.1049/iet-its.2016.0276

Kandt, J., & Leak, A. (2019). Examining inclusive mobility through smartcard data: What shall we make of senior citizens' declining bus patronage in the West Midlands? *Journal of Transport Geography.* https://doi.org/10.1016/j.jtrangeo.2019.102474

Kang, C., Fan, D., & Jiao, H. (2021). Validating activity, time, and space diversity as essential components of urban vitality. *Environment and Planning B: Urban Analytics and City Science.* https://doi.org/10.1177/2399808320919771

Kaufmann, V. (2014). Mobility as a tool for sociology. *Sociologica.* https://doi.org/10.2383/77046

Kim, J., & Kwan, M. P. (2021). The impact of the COVID-19 pandemic on people's mobility: A longitudinal study of the U.S. from March to September of 2020. *Journal of Transport Geography.* https://doi.org/10.1016/j.jtrangeo.2021.103039

Lan, F., Gong, X., Da, H., & Wen, H. (2020). How do population inflow and social infrastructure affect urban vitality? Evidence from 35 large- and medium-sized cities in China. *Cities.* https://doi.org/10.1016/j.cities.2019.102454

Larsen, J., Axhausen, K. W., & Urry, J. (2006). Geographies of social networks: Meetings, travel and communications. *Mobilities.* https://doi.org/10.1080/17450100600726654

Li, Z., Wang, S., Cheng, L., Shi, X., Guan, H., & Shu, C. (2022). Differences and relationship between population flow and transportation networks in Northeast China. *Prog. Geogr, 41*(6), 985998.

Liang, D., Li, X., & Zhang, Y. Q. (2016). Identifying familiar strangers in human encounter networks. *EPL.* https://doi.org/10.1209/0295-5075/116/18006

Liben-Nowell, D., Novak, J., Kumar, R., Raghavan, P., & Tomkins, A. (2005). Geographic routing in social networks. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.0503018102

Liu, G. J., & Engels, B. (2012). Accessibility to essential services and facilities by a spatially dispersed aging population in suburban Melbourne, Australia. *Advances in location-based services: 8th International Symposium on Location-Based Services, Vienna, 2011,* 327–348.

Liu, Q., Liu, Y., Zhang, C., An, Z., & Zhao, P. (2021). Elderly mobility during the COVID-19 pandemic: A qualitative exploration in Kunming, China. *Journal of Transport Geography.* https://doi.org/10.1016/j.jtrangeo.2021.103176

Liu, S., Zhang, L., Long, Y., Long, Y., & Xu, M. (2020). A new urban vitality analysis and evaluation framework based on human activity modeling using multi-source big data. *ISPRS International Journal of Geo-Information.* https://doi.org/10.3390/ijgi9110617

Liu, W., & Shi, E. (2016). Spatial pattern of population daily flow among cities based on ICT: A case study of "Baidu migration." *Acta Geographica Sinica, 71*(10), 1667–1679.

Long, Y., Liu, X., Zhou, J., & Chai, Y. (2016). Early birds, night owls, and tireless/recurring itinerants: An exploratory analysis of extreme transit behaviors in Beijing, China. *Habitat International.* https://doi.org/10.1016/j.habitatint.2016.08.004

Loo, B. P. Y., Tsoi, K. H., Wong, P. P. Y., & Lai, P. C. (2021). Identification of superspreading environment under COVID-19 through human mobility data. *Scientific Reports.* https://doi.org/10.1038/s41598-021-84089-w

Lovelace, R., Beecham, R., Heinen, E., Vidal Tortosa, E., Yang, Y., Slade, C., & Roberts, A. (2020). Is the London cycle hire scheme becoming more inclusive? An evaluation of the shifting spatial distribution of uptake based on 70 million trips. *Transportation Research Part A: Policy and Practice*. https://doi.org/10.1016/j.tra.2020.07.017

Lucas, K. (2012). Transport and social exclusion: Where are we now? *Transport Policy.* https://doi.org/10.1016/j.tranpol.2012.01.013

Luo, F., Cao, G., Mulligan, K., & Li, X. (2016). Explore spatiotemporal and demographic characteristics of human mobility via Twitter: A case study of Chicago. *Applied Geography.* https://doi.org/10.1016/j.apgeog.2016.03.001

Manovich, L. (2015). Trending: The promises and the challenges of big social data. In *Debates in the digital humanities.* https://doi.org/10.5749/minnesota/9780816677948.003.0047

Maoh, H., & Tang, Z. (2012). Determinants of normal and extreme commute distance in a sprawled midsize Canadian city: Evidence from Windsor, Canada. *Journal of Transport Geography.* https://doi.org/10.1016/j.jtrangeo.2012.07.003

Martín, Y., Cutter, S. L., & Li, Z. (2020). Bridging Twitter and survey data for evacuation assessment of Hurricane Matthew and Hurricane Irma. *Natural Hazards Review.* https://doi.org/10.1061/(asce)nh.1527-6996.0000354

Marx, K. (1867). *Capital: A critique of political economy. Volume 1, Part 1: The process of capitalist production.* New York, NY: Cosimo.

Mavragani, A., Ochoa, G., & Tsagarakis, K. P. (2018). Assessing the methods, tools, and statistical approaches in Google trends research: Systematic review. *Journal of Medical Internet Research.* https://doi.org/10.2196/jmir.9366

Milgram, S. (1977). The familiar stranger: An aspect of urban anonymity. *Chi Letters, 6*(1), 223–230.

Moreno-Monroy, A. I., Lovelace, R., & Ramos, F. R. (2018). Public transport and school location impacts on educational inequalities: Insights from São Paulo. *Journal of Transport Geography.* https://doi.org/10.1016/j.jtrangeo.2017.08.012

Nelson, T., Roy, A., Ferster, C., Fischer, J., Brum-Bastos, V., Laberee, K., ... & Winters, M. (2021). Generalized model for mapping bicycle ridership with crowdsourced data. *Transportation Research Part C: Emerging Technologies, 125,* 102981.

Nguyen, M. H., & Armoogum, J. (2020). Hierarchical process of travel mode imputation from GPS data in a motorcycle-dependent area. *Travel Behavior and Society, 21,* 109–120.

Noulas, A., Scellato, S., Lambiotte, R., Pontil, M., & Mascolo, C. (2012). A tale of many cities: Universal patterns in human urban mobility. *PLoS ONE.* https://doi.org/10.1371/journal.pone.0037027

Ooi, K. B., Foo, F. E., Tan, G. W. H., Hew, J. J., & Leong, L. Y. (2021). Taxi within a grab? A gender-invariant model of mobile taxi adoption. *Industrial Management and Data Systems.* https://doi.org/10.1108/IMDS-04-2020-0239

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., … & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ.* Retrieved from https://www.bmj.com/content/372/bmj.n71

Pan, Y., & He, S. Y. (2022). Analyzing COVID-19's impact on the travel mobility of various social groups in China's Greater Bay Area via mobile phone big data. *Transportation Research Part A: Policy and Practice, 159,* 263–281.

Park, C., & Chang, J. S. (2020). Spatial equity of excess commuting by transit in Seoul. *Transportation planning and technology, 43*(1), 101–112.

Poom, A., Järv, O., Zook, M., & Toivonen, T. (2020). COVID-19 is spatial: Ensuring that mobile big data is used for social good. *Big Data and Society.* https://doi.org/10.1177/2053951720952088

Poorthuis, A., Shelton, T., & Zook, M. (2021). Changing neighborhoods, shifting connections: Mapping relational geographies of gentrification using social media data. Urban Geography. https://doi.org/10.1080/02723638.2021.1888016

Puura, A., Silm, S., & Ahas, R. (2018). The Relationship between social networks and spatial mobility: A mobile-phone-based study in Estonia. *Journal of Urban Technology.* https://doi.org/10.1080/10630732.2017.1406253

Qiao, S., Huang, G., & Yeh, A. G. O. (2023). Who are the gig workers? Evidence from mapping the residential locations of ride-hailing drivers by a big data approach. *Cities, 132*, 104112.

Rapino, M. A., & Fields, A. K. (2013). *Mega commuters in the US: Time and distance in defining the long commute using the American Community Survey* (No. Working Paper 2013-03).

Rashidi, T. H., Abbasi, A., Maghrebi, M., Hasan, S., & Waller, T. S. (2017). Exploring the capacity of social media data for modelling travel behavior: Opportunities and challenges. *Transportation Research Part C: Emerging Technologies.* https://doi.org/10.1016/j.trc.2016.12.008

Raux, C., Zoubir, A., & Geyik, M. (2017). Who are bike sharing schemes members and do they travel differently? The case of Lyon's "Velo'v" scheme. *Transportation Research Part A: Policy and Practice*. https://doi.org/10.1016/j.tra.2017.10.010

Saberi, M., Ghamami, M., Gu, Y., Shojaei, M. H., & Fishman, E. (2018). Understanding the impacts of a public transit disruption on bicycle sharing mobility patterns: A case of tube strike in London. *Journal of Transport Geography.* https://doi.org/10.1016/j.jtrangeo.2017.11.018

Sajeevan, A., Jyolsna, P. (2019). A survey on Location Recommendation Based on PSI Model and Trajectory Mining. National Conference on Advances in Smart Computing and Data Science(NCASCD-'19), Mar Baselios Christian College of Engineering & Technology.

Sarram, G., & Ivey, S. S. (2022). Evaluating the potential of online review data for augmenting traditional transportation planning performance management. *Journal of Urban Management.* https://doi.org/10.1016/j.jum.2022.01.001

Sevtsuk, A., Basu, R., Li, X., & Kalvo, R. (2021). A big data approach to understanding pedestrian route choice preferences: Evidence from San Francisco. *Travel Behavior and Society.* https://doi.org/10.1016/j.tbs.2021.05.010

Shalit, N., Fire, M., & Ben-Elia, E. (2020). A supervised machine learning model for imputing missing boarding stops in smart card data. https://arxiv.org/abs/2003.05285

Sheller, M. (2014). The new mobilities paradigm for a live sociology. *Current Sociology.* https://doi.org/10.1177/0011392114533211

Shi, Z., Liu, X., Lai, J., Tong, C., Zhang, A., & Shi, W. (2021). A data-driven framework for analyzing spatial distribution of the elderly cardholders by using smart card data. *ISPRS International Journal of Geo-Information.* https://doi.org/10.3390/ijgi10110728

Shoval, N., McKercher, B., Birenboim, A., & Ng, E. (2015). The application of a sequence alignment method to the creation of typologies of tourist activity in time and space. *Environment and Planning B: Planning and Design.* https://doi.org/10.1068/b38065

Sievert, C., & Shirley, K. (2014, June). LDAvis: A method for visualizing and interpreting topics. In *Proceedings of the workshop on interactive language learning, visualization, and interfaces* (pp. 63–70).

Simmel, G. (1995). L. R. Kramme, & A. Rammstedt (Eds.) *Aufsatze ybd Abhandlungen 1901-1908, vol.1.* Frankfurt: Suhrkamp.

Singh, V. K., Bozkaya, B., & Pentland, A. (2015). Money walks: Implicit mobility behavior and financial well-being. *PLoS ONE.* https://doi.org/10.1371/journal.pone.0136628

Sobolevsky, S., Sitko, I., Des Combes, R. T., Hawelka, B., Arias, J. M., & Ratti, C. (2014). Money on the move: Big data of bank card transactions as the new proxy for human mobility patterns and regional delineation. The case of residents and foreign visitors in Spain. *2014 IEEE International Congress on Big Data, Anchorage, AK,* 136–143. https://doi.org/10.1109/BigData.Congress.2014.28

Sohn, J. (2016). Development of dynamic passenger-trip assignment model of urban railway using Seoul-Incheon-Gyeonggi's transportation card. *KSCE Journal of Civil and Environmental Engineering Research, 36*(1), 105–114.

Solmaz, G., & Turgut, D. (2017). Modeling pedestrian mobility in disaster areas. *Pervasive and Mobile Computing.* https://doi.org/10.1016/j.pmcj.2017.05.005

Song, C., Qu, Z., Blumm, N., & Barabási, A. L. (2010). Limits of predictability in human mobility. *Science.* https://doi.org/10.1126/science.1177170

Stipancic, J., Miranda-Moreno, L., & Saunier, N. (2018). Vehicle maneuvers as surrogate safety measures: Extracting data from the GPS-enabled smartphones of regular drivers. *Accident Analysis & Prevention, 115,* 160–169.

Stopher, P. R. (2009). The travel survey toolkit: Where to from here? In *Transport survey methods: Keeping up with a changing world.* Bingley, UK: Emerald Group Publishing Limited.

Stopher, P., & Stecher, C. (2006). *Travel Survey methods: Quality and future directions.* Bingley, UK: Emerald Group Publishing.

Su, S., Li, Z., Xu, M., Cai, Z., & Weng, M. (2017). A geo-big data approach to intra-urban food deserts: Transit-varying accessibility, social inequalities, and implications for urban planning. *Habitat International, 64,* 22–40.

Su, S., Zhou, H., Xu, M., Ru, H., Wang, W., & Weng, M. (2019). Auditing street walkability and associated social inequalities for planning implications. *Journal of Transport Geography*. https://doi.org/10.1016/j.jtrangeo.2018.11.003

Suciu, G., Butca, C., Dobre, C., & Popescu, C. (2017, May). Smart city mobility simulation and monitoring platform. In *Proceedings of the 2017 21st International Conference on Control Systems and Computer Science (CSCS),* 685–689.

Sulis, P., Manley, E., Zhong, C., & Batty, M. (2018). Using mobility data as proxy for measuring urban vitality. *Journal of Spatial Information Science.* https://doi.org/10.5311/JOSIS.2018.16.384

Sun, L., Axhausen, K. W., Lee, D. H., & Huang, X. (2013). Understanding metropolitan patterns of daily encounters. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.1306440110

Szell, M. (2018). Crowdsourced quantification and visualization of urban mobility space inequality. *Urban Planning, 3*(1), 1–20.

Tahmasbi, B., Mansourianfar, M. H., Haghshenas, H., & Kim, I. (2019). Multimodal accessibility-based equity assessment of urban public facilities distribution. *Sustainable Cities and Society.* https://doi.org/10.1016/j.scs.2019.101633

Tao, S., Zhang, M., & Wu, J. (2021). Big data applications in urban transport research in Chinese cities: an overview. *Big Data Applications in Geography and Planning.* https://doi.org/10.4337/9781789909791.00020

Toole, J. L., Herrera-Yaqüe, C., Schneider, C. M., & González, M. C. (2015). Coupling human mobility and social ties. *Journal of the Royal Society Interface.* https://doi.org/10.1098/rsif.2014.1128

Trasberg, T., & Cheshire, J. (2021). Spatial and social disparities in the decline of activities during the COVID-19 lockdown in Greater London. *Urban Studies.* https://doi.org/10.1177/00420980211040409

Traunmueller, M. W., Johnson, N., Malik, A., & Kontokosta, C. E. (2018). Digital footprints: Using WiFi probe and locational data to analyze human mobility trajectories in cities. *Computers, Environment and Urban Systems.* https://doi.org/10.1016/j.compenvurbsys.2018.07.006

Travers, J., & Milgram, S. (1967). *An experimental study of the small world problem* (179–197). Cambridge, MA: Academic Press.

van Acker, V., van Wee, B., & Witlox, F. (2010). When transport geography meets social psychology: Toward a conceptual model of travel behavior. *Transport Reviews.* https://doi.org/10.1080/01441640902943453

Walker, J. S. (2014). Big data: A revolution that will transform how we live, work, and think. *International Journal of Advertising.* https://doi.org/10.2501/ija-33-1-181-183

Wang, A., Zhang, A., Chan, E. H. W., Shi, W., Zhou, X., & Liu, Z. (2021). A review of human mobility research based on big data and its implication for smart city development. *ISPRS International Journal of Geo-Information.* https://doi.org/10.3390/ijgi10010013

Wang, H., & Kilmartin, L. (2014). Comparing rural and urban social and economic behavior in Uganda: Insights from mobile voice service usage. *Journal of Urban Technology.* https://doi.org/10.1080/10630732.2014.888296

Wang, L., Gopal, R., Sankaranarayanan, R., & Pancras, J. (2013). *Checking in to check it out: An empirical analysis of customers' engagement on location based social media.* Paper presented at the 23rd Workshop on Information Technology and Systems: Leveraging Big Data Analytics for Societal Benefits, WITS Milan, Italy, December 14–15.

Wang, M., & Vermeulen, F. (2021). Life between buildings from a street view image: What do big data analytics reveal about neighborhood organizational vitality? *Urban Studies.* https://doi.org/10.1177/0042098020957198

Wang, P., Liu, K., Wang, D., & Fu, Y. (2021). Measuring urban vibrancy of residential communities using big crowdsourced geotagged data. *Frontiers in Big Data, 4,* 690970.

Wang, Q., Phillips, N. E., Small, M. L., & Sampson, R. J. (2018). Urban mobility and neighborhood isolation in America's 50 largest cities. *Proceedings of the National Academy of Sciences of the United States of America.* https://doi.org/10.1073/pnas.1802537115

Wang, R., Hao, J. X., Law, R., & Wang, J. (2019). Examining destination images from travel blogs: A big data analytical approach using latent Dirichlet allocation. *Asia Pacific Journal of Tourism Research.* https://doi.org/10.1080/10941665.2019.1665558

Wang, R., Zhang, X., & Li, N. (2022). Zooming into mobility to understand cities: A review of mobility-driven urban studies. *Cities, 130,* 103939. https://doi.org/https://doi.org/10.1016/j.cities.2022.103939

Wang, Y., Correia, G. H. de A., van Arem, B., & Timmermans, H. J. P. (Harry. (2018). Understanding travelers' preferences for different types of trip destination based on mobile internet usage data. *Transportation Research Part C: Emerging Technologies.* https://doi.org/10.1016/j.trc.2018.03.009

Wei, M. (2022). Investigating the influence of weather on public transit passenger's travel behavior: Empirical findings from Brisbane, Australia. *Transportation Research Part A: Policy and Practice*. https://doi.org/10.1016/j.tra.2021.12.001

Welch, T. F., & Widita, A. (2019). Big data in public transportation: A review of sources and methods. *Transport Reviews, 39*(6), 795-818.

Weth, C. V. D., Abdul, A. M., & Kankanhalli, M. (2017). *Cyber-physical social networks. ACM Transactions on Internet Technology (TOIT), 17*(2), 1–25.

Whyte, H. W. (1988). City: Rediscovering the centre. New York: Doubleday Dell Publishing Group.

Woo, K. S., & Suh, J. H. (2020a). A time series analysis of urban park behavior using big data. *Journal of the Korean Institute of Landscape Architecture, 48*(1), 35–45.

Woo, K. S., & Suh, J. H. (2020b). Analysis of behavior of Seoullo 7017 visitors with a focus on text mining and social network analysis. *Journal of the Korean Institute of Landscape Architecture, 48*(6), 16–24.

Wu, J., Jiang, C., Houston, D., Baker, D., & Delfino, R. (2011). Automated time activity classification based on global positioning system (GPS) tracking data. *Environmental Health, 10*(1), 1–13.

Wu, J., Zhou, J., & Ma, H. (2022). Revisiting the valuable locales in our cities? Visualizing social interaction potential around metro station areas in Wuhan, China. *Environment and Planning A*. https://doi.org/10.1177/0308518X211062227

Wu, W., Wang, J., & Dai, T. (2016). The geography of cultural ties and human mobility: Big data in urban contexts. *Annals of the American Association of Geographers.* https://doi.org/10.1080/00045608.2015.1121804

Xu, D., Zhang, X., Zhang, X., & Yu, Y. (2022). Type identification of land use in metro station area based on spatial–temporal features extraction of human activities. *Sustainability, 14*(20), 13122.

Xu, Y., Belyi, A., Bojic, I., & Ratti, C. (2018). *Human mobility and socioeconomic status: Analysis of Singapore and Boston. Computers, Environment and Urban Systems.* https://doi.org/10.1016/j.compenvurbsys.2018.04.001

Xu, Y., Liu, Y., Chang, X., & Huang, W. (2021). How does air pollution affect travel behavior? A big data field study. *Transportation Research Part D: Transport and Environment, 99,* 103007.

Xu, Y., Santi, P., & Ratti, C. (2022). Beyond distance decay: Discover homophily in spatially embedded social networks. *Annals of the American Association of Geographers.* https://doi.org/10.1080/24694452.2021.1935208

Xu, Y., Shaw, S. L., Zhao, Z., Yin, L., Fang, Z., & Li, Q. (2015). Understanding aggregate human mobility patterns using passive mobile phone location data: A home-based approach. *Transportation.* https://doi.org/10.1007/s11116-015-9597-y

Yabe, T., Tsubouchi, K., Fujiwara, N., Sekimoto, Y., & Ukkusuri, S. V. (2020). Understanding post-disaster population recovery patterns. *Journal of the Royal Society Interface.* https://doi.org/10.1098/rsif.2019.0532

Yoo, E.-H. (2019). How short Is long enough? Modeling temporal aspects of human mobility behavior using mobile phone data. *Annals of the American Association of Geographers.* https://doi.org/10.1080/24694452.2019.1586516

Yoon, Y., & Park, J. (2022). Equitable city in an aging society: Public transportation-based primary care accessibility in Seoul, Korea. *Sustainability, 14*(16), 9902.

Younus, A., Qureshi, M. A., Manchanda, P., O'riordan, C., & Pasi, G. (2014). Utilizing microblog data in a topic modelling framework for scientific articles' recommendation. *Lecture Notes in Computer Science.* https://doi.org/10.1007/978-3-319-13734-6_28

Yuan, N. J., Zheng, Y., Zhang, L., & Xie, X. (2013). T-finder: A recommender system for finding passengers and vacant taxis. *IEEE Transactions on Knowledge and Data Engineering.* https://doi.org/10.1109/TKDE.2012.153

Zhai, W., Fu, X., Liu, M., & Peng, Z. R. (2021). The impact of ethnic segregation on neighborhood-level social distancing in the United States amid the early outbreak of COVID-19. *Urban Studies*. https://doi.org/10.1177/00420980211050183

Zhang, H., Zhuge, C., Jia, J., Shi, B., & Wang, W. (2021). Green travel mobility of dockless bike-sharing based on trip data in big cities: A spatial network analysis. *Journal of Cleaner Production, 313,* 127930.

Zhang, M., & Zhao, P. (2021). Literature review on urban transport equity in transitional China: From empirical studies to universal knowledge. *Journal of Transport Geography.* https://doi.org/10.1016/j.jtrangeo.2021.103177

Zhang, S., Tang, J., Wang, H., Wang, Y., & An, S. (2017). Revealing intra-urban travel patterns and service ranges from taxi trajectories. *Journal of Transport Geography.* https://doi.org/10.1016/j.jtrangeo.2017.04.009

Zhao, J., Wang, J., Xing, Z., Luan, X., & Jiang, Y. (2018). Weather and cycling: Mining big data to have an in-depth understanding of the association of weather variability with cycling on an off-road trail and an on-road bike lane. *Transportation Research Part A: Policy and Practice.* https://doi.org/10.1016/j.tra.2018.03.001

Zhao, K., Tarkoma, S., Liu, S., & Vo, H. (2016). Urban human mobility data mining: An overview. *Proceedings - 2016 IEEE International Conference on Big Data, Big Data 2016.* https://doi.org/10.1109/BigData.2016.7840811

Zhao, P. S., Lee, J. H., & Lee, M. H. (2019). Characterizing the structure of China's passenger railway network based on the social network analysis (SNA) approaches: Focused on the 2008, 2013, and 2018 railway service data, respectively. *The Journal of the Korea Contents Association, 19*(10), 685-697.

Zhao, Z., Koutsopoulos, H. N., & Zhao, J. (2018). Individual mobility prediction using transit smart card data. *Transportation Research Part C: Emerging Technologies.* https://doi.org/10.1016/j.trc.2018.01.022

Zheng, Z., Rasouli, S., & Timmermans, H. (2016). Two-regime pattern in human mobility: Evidence from GPS taxi trajectory data. *Geographical Analysis.* https://doi.org/10.1111/gean.12087

Zhou, J., Wu, J., & Ma, H. (2021). Abrupt changes, institutional reactions, and adaptive behaviors: An exploratory study of COVID-19 and related events' impacts on Hong Kong's metro riders. *Applie*d *Geography*. https://doi.org/10.1016/j.apgeog.2021.102504

Zhou, J., Yang, Y., Ma, H., & Li, Y. (2020). "Familiar strangers" in the big data era: An exploratory study of Beijing metro encounters. *Cities.* https://doi.org/10.1016/j.cities.2019.102495

Zhou, X. (2015). Understanding spatiotemporal patterns of biking behavior by analyzing massive bike sharing data in Chicago. *PLoS ONE.* https://doi.org/10.1371/journal.pone.0137922

Zhou, J., Li, Y. and Yang, Y. (2018a). Familiar strangers: Visualizing potential metro encounters in Beijing. *Environment and Planning A 50*(2), 262–265.

Zhou, J., Yang, Y., Li, Y., & Maurer, V. (2018b). Someone like you: Visualizing co-presences of metro riders in Beijing. *Environmental and Planning A 50*(4), 752–755.