ISTANBUL UNIVERSITY PRESS
Journal of Technology in Architecture Design and Planning
JTADP 2023, 1 (1), 64–83
DOI: 10.26650/JTADP.01.004

RESEARCH ARTICLE

A Systematic Literature Review of Big Data in Urban Studies

Gülce KIRDAR[1]

[1]Istanbul Technical University, Graduate School, Department of Informatics, Architectural Design Computing, Istanbul, Türkiye

ABSTRACT

The paper aims to explore the use of big data in urban studies by analyzing selected state-of-the-art studies in urban informatics that utilize big data to support urban decision-making. The study conducts exploratory research to gain insight into the association patterns of big data-related concepts. The researchers use the VOSviewer tool to analyze 30 selected references based on keyword occurrences, abstracts, and titles. The study also focuses on how the references handle decision support and examines the relationship network of decision support with other terms. The qualitative and quantitative analysis results are presented to show the association and numeric distribution of the terms. The paper finds that decision support in the selected studies is mainly provided through data-driven computational methods, spatial statistical methods, and mapping of the spatiotemporal pattern of urban phenomena. The reference studies mainly support decisions related to urban activities and functioning, user activities and movement, visiting, and urban perception. The study contributes to presenting the trend in big data studies for urban planning and decision-making.

Keywords: Big data, decision support, urban informatics, data-driven, exploratory analysis

INTRODUCTION

Understanding spatiotemporal changes in user behavior has been a crucial challenge for urban studies. With developments in ubiquitous computing in the 1990s, information technology has permeated urban life through web technologies, information and communication technologies (ICT), sensors, and the Internet of Things (IoT). Pervasive computation hyperextends urban spaces, reinforces multiple interaction networks between users and computers, and consequently generates networked societies and net localities. Net locality is composed of a mixture of digital information and physical localities (Gordon and E Silva, 2011, p.56). In this networked ecosystem, large amounts of data trails are produced. Analyzing the vast volume and variety of data provides researchers with opportunities to understand urban sociability through citizens' digital trails. The information network manages 'the metabolism of urban lives' from environmental, social, and economic aspects, as Townsend (2013, p.6) states. Researchers have begun measuring data, enabling us to control the city's network. The information fallacy enriches the decision-making process.

In this data-rich environment, big data bring about more efficient and effective decision-making in urban planning. The use of big data in urban studies is a relatively new field, and therefore there is a need for more research and exploration in this area. Big data has wide use in urban studies in understanding urban functioning and supporting decisions. Categorizing the references is a crucial step in conducting an efficient literature review in the scope of big data studies. Categorizing references based on interrelationships and grouping them under specific topics in urban informatics allows researchers to better understand hot research topics and identify impactful studies. A systematic and consistent categorization approach ensures a comprehensive and accurate literature review, enhancing research quality. In this regard, this study aims to create a guideline for urban researchers by conducting a systematic literature review.

The research problem explores the potential of big data in understanding urban functioning and supporting urban planning decisions, particularly in the context of urban informatics. The study aims to answer the question of how big data can be utilized to support decision-making in urban studies and to identify the current trends and methods in the use of big data for this purpose. This paper aims to investigate the potential of big data in understanding urban functioning and supporting urban planning decisions. This study poses the question of how to use big data to support decision-making in the context of urban informatics. To that end,

Corresponding Author: Gülce Kırdar **E-mail:** gulcekirdar@gmail.com
Submitted: 26.02.2023 • **Revision Requested:** 30.03.2023 • **Last Revision Received:** 10.04.2023 • **Accepted:** 12.04.2023 • **Published Online:** 09.06.2023

this paper focuses on decision support in urban studies using big data. It provides insight into how big data is utilized in urban informatics studies to understand urban functioning and support decision-making based on the references.

The paper begins with an introduction to the research context, encompassing decision-support systems within data-driven decision-support methods, and the use of big data in urban informatics.

In the second section, the paper delves into the primary subject of this study—big data—by elucidating its definitions and contributions to urban studies, and the locative media data concept which is the scope of the study. The methodology section (Section 3) outlines the qualitative analysis process employed in the study. It is followed by a review of the selected case studies for exploratory analysis in the fourth section. The study presents both qualitative and quantitative results derived from the analysis, in the fifth section. The discussion part (Section 6) addresses ongoing debates concerning the use of big data in urban studies. Lastly, the paper concludes by offering a comprehensive overview of the study, discussing its limitations, and providing recommendations for future research.

LITERATURE REVIEW

Decision support in urban informatics

Advancements in computational systems and the ubiquity of data have given rise to urban informatics. Urban informatics combines theoretical developments in social physics, ecology, urban economics, and location theories with technological innovations in computer systems to understand and model urban functioning (Batty, 2013b). According to Kitchin (2016) and Townsend (2015), "Urban informatics is a field concerning the generation, management, processing, and analysis of urban data" (p. 3, cited in Kitchin, 2017). Big data, advances in ICT, social networks, data from social media platforms, citizen science, participatory practices, and an open data culture contribute to the growth of urban informatics (Batty, 2013b). Urban informatics has been widely used in the following research and application areas: strategies for urban development, planning and policy analysis, management and civic participation, theoretical insights, and knowledge discovery of urban patterns (Thakuriah, Tilahun, and Zellner, 2017). Data-intensive research in urban informatics shifts the focus from computational modeling and simulation of observed phenomena to data-driven modeling, hypothesis generation, and visual data description (Thakuriah, Tilahun & Zellner, 2017).

Data-driven decision support empowers urban planners to make more informed, evidence-based decisions, ultimately leading to better urban policies and outcomes. Decision Support Systems (DSS) are computer-based systems that support users by providing alternatives in the decision-making process for unstructured and multi-criteria problem solutions (Çağdaş et al., 2015). In DSS, the system formalizes the problem and computes the reasoning method to solve it, offering alternative solutions whose selection depends on expert judgment (Drudzel and Flynn, 2002). Decision-support systems have been improved with data-driven analysis methods and big data. Decision-making transforms into a data-driven process, enabled by the expanding scale and types of data available for analysis (Loo & Tang, 2019). Urban planners and designers utilize urban data as evidence in decision-making to support decisions with use cases (Tunçer and You, 2017). More types and amounts of data enable designers to address domain and scale simultaneously, and as a result, they gain more information or evidence to support urban design decisions (Tunçer and You, 2017). Knowledge discovery with data analysis methods generates evidence to be used in design or planning. It offers evidence-based decision support in urban planning decisions. According to Thakuriah, Tilahun, and Zellner (2017, p.33), "The knowledge discovery aspects of data-driven models are important to attract the attention of citizens and decision-makers to urban problems and to stimulate new hypotheses about urban phenomena, which could potentially be rigorously tested using inferential urban models."

Big data

Big urban data has been widely utilized in urban research paradigms to restructure existing computational approaches to urban modeling and simulations, and data analysis methods have been strengthened with machine learning and data mining (Thakuriah, Tilahun, and Zellner, 2017). Big data has three dimensions, which are volume, variety, and complexity (Laney, 2001). Kaisler et al. (2013) describe the big data dimensions as follows:

> "Data volume measures the amount of data, data velocity measures the speed of data creation, streaming, and aggregation; data variety measures the richness of data sources (text, image, video, etc.); data value measures the usefulness of data. In addition to three dimensions, data value measures the usefulness of data in making decisions, and complexity measures the degree of interconnectedness and interdependence in data structures (Kaisler et al., 2013, p. 997)."

Big data approaches support evidence-based decision support through data-driven analysis and decision-making. Urban studies take advantage of big urban data to collect evidence with the intent to support decisions. Batty and colleagues (2012) indicate that integrating big data analysis into urban design leads to informed decision-making about the city's functioning. Big data analytics implement data-driven approaches. Big data analytics support the decision-making process by testing decisions with large amounts and varieties of data. It contributes to the field of evidence-based urban design. Through big data analysis, designers can gain insight into the behavior patterns of the community, which can be valuable evidence in decision-making. The variability of data enables designers to address domain and scale simultaneously, and as a result, they gain more information or evidence to support urban design decisions (Tunçer and You, 2017). As the analyzed data types, scale, and time expand, more valuable information can be gathered and transformed into more useful evidence in the urban decision process. According to Loo and Zhang (2016), "Decision-making becomes a data-driven process" (p. 139). Big data analysis contributes to knowledge generation and decision-making in urban planning by improving the quality, trustworthiness, and legitimacy of decisions (Fredriksson, 2018).

In addition to decision support, the big data approach opens up new possibilities in data collection, analysis, evaluation, representation, and organization methods in urban informatics. Big data paves the way for 'sensing the urban environment' (Batty, 2001). The ubiquity of big data affects the role of urban planners, spatiotemporal resolution, and scaling of urban planning. The role of urban planners expands to include observing changes in urban dynamics ranging from land use and traffic to socio-economy and demography through multiple databases and sensors (Laurini, 2001). Space is disconnected from time with high spatiotemporal resolution of data, and scaling expands from a physical to a socio-economical perspective (Batty, 2001). Big data obtained through digital technology represents a rich and valuable alternative to conventional observation data with high space-time resolution, accuracy, and quality. The use of big data shifts the focus of urban planning to shorter time periods through real-time streaming and user data (Batty, 2013a). Real-time streaming enables urban planners to measure changes and understand user behavior on a shorter time scale (Batty, 2016). Big data presents a valuable source for data visualization and mapping purposes. Big data expands the scale and content of data mapping for information dissemination. The integration of spatiotemporal data on an urban data platform allows for monitoring city performance in real-time and making smarter decisions to enhance performance (Loo and Tang, 2019). Lastly, the big data approach changes urban organization and management methods by fostering the participatory urban design approach.

Big Locative media data

Wilken and Goggin (2017) define locative media as media intrinsically connected to the places and spaces where they are utilized and encountered. These media types are location-based or location-aware, incorporating GPS coordinates, geotagged images or videos, social media check-ins, or any other digital content associated with a particular physical location (Wilken & Goggin, 2017). Locative media data, as defined by Wilken and Goggin (2017), plays a crucial role in understanding the relationship between individuals, places, and digital technology. Locative big data has the potential to greatly impact various fields by providing valuable spatial and temporal information regarding collective user patterns. By integrating locative big data with other datasets, stakeholders can gain a comprehensive understanding of urban dynamics, leading to more effective decision-making processes and improved urban living conditions. Web 2.0 technology, GIS, geotagging, and GPS are the enablers of locative media data (Goodchild, 2007). Locative media data encompasses various types such as geolocated (Adelfio et al., 2020; García-Palomares et al., 2015; Girardin et al., 2008), locative social media (Martí et al., 2017), geospatial (Ensari & Kobaş, 2018; Schlieder & Matyas, 2009), and georeferenced data (Wood et al., 2013).

This study employs Location Based Social Network (LBSN) data as a primary source of locative media data, collected from location-based technologies such as websites, platforms, apps, and online services (W.-C. Lee & Ye, 2014). LBSN and locative media data, both relying on location-based information, provide tailored and interactive experiences for users. By leveraging advanced data analysis tools (Tasse & Hong, 2017), the study aims to achieve an affordable, scalable, and insightful understanding of urban functioning through locative media data from LBSN. This approach enables the exploration of intangible urban aspects from user experiences and perceptions (Martí et al., 2019). LBSN data analysis offers several benefits, such as large sampling, unobtrusive data collection, user bias avoidance, cost-effectiveness, and time efficiency (Martí et al., 2019).

Geolocated data is the data sharing of inhabitants about their activities, perceptions, and interactions when using ICT-based smart mobile applications. Geolocated data (Girardin et al., 2008; Garcia Palomares et al., 2015; Adelfio et al., 2020) has different definitions in the literature: locative social media data (Marti et al., 2017), geospatial data (Kobaş and Ensari, 2018, Schlieder and Matyas, 2009), georeferenced data (Wood et al., 2018; De Choudhury et al., 2010; Shlieder and Matyas, 2009; Manal et al., 2018; Jang et al., 2019; Mart et al., 2017) user-generated content (UGC) (Mora et al., 2018; Girardin et al., 2008), and Volunteered Geographic Information (VGI) (Goodchild, 2007; Jiang et al., 2015; Deng & Newsam, 2017). Geolocated data sources consist of (i) sensing data (GPS traces), (ii) crowdsourcing mapping data (mapping applications like Open Street Map (OPM)), and (iii) LBSN data or social media data (Instagram, Twitter, Flickr, Foursquare). Batty and other researchers (2012) emphasize that analytics of user data provides the possibility to capture the interaction, flows, and networks between the user and the city.

METHOD

This research conducts exploratory research to gain a better understanding of the use of big data in urban studies. In this article, the researcher applies qualitative analysis to a systematic literature review to identify relevant publications on big data in urban studies. The review was conducted on various leading databases of peer-reviewed literature with a focus on articles related to big data, data analysis, urban planning and design, and the contribution of big data analysis to supporting urban planning decisions. The researcher carries out a systematic literature review process using different filters to identify scientific publications as reference studies. A substantial volume of publications have been collected from leading database platforms of peer-reviewed literature, including ScienceDirect, Tandfonline, Arxiv, IEE Explore, Nature, ACM Library, Journal SagePub, and PlosOne. These platforms are renowned for hosting esteemed journals characterized by high impact factors and superior quartile rankings. A total of 30 articles were selected for this study, all of which originate from journals indexed in Q3 or higher quartile rankings. This criterion ensures the inclusion of reputable and impactful sources in the analysis.

In selecting articles for this study, the researcher establishes specific criteria to align with the main research topics: big data, data analysis, urban planning, design, and decision support. The selection process involves two main criteria. First, articles published in journals indexed in Q3 or higher quartile rankings are selected. Second, articles containing relevant keywords, ***big data, geolocated data types (e.g., LBSN, VGI, UGC, SMD), urban analysis, and urban form and functional attributes***, are chosen. The coherence of the articles in terms of purpose, methods used, and results are also crucial for selection. By employing these criteria, the researcher ensures that the selected sources are both pertinent and valuable for the research objectives while upholding academic standards. Based on these criteria, 30 relevant articles have been selected. This study conducts a bibliometric analysis of the selected studies with VOSviewer to describe the overview and evolution of big data use in urban studies, as qualitative analysis.

As explained by Van Eck and Waltman (2014), there are various visualization approaches for bibliometric networks, with three common ones being distance-based, graph-based, and timeline-based approaches. The distance-based approach positions nodes so that the distance between them represents their relatedness, utilizing techniques such as multidimensional scaling. The graph-based approach places nodes in a two-dimensional space and uses edges to indicate relatedness, employing algorithms like the Kamada-Kawai method. The timeline-based approach connects each node to a specific point in time, making it ideal for visualizing publication networks. In this method, one dimension represents time, while the other shows the relatedness of nodes. VOSviewer uses the distance-based visualization approach for bibliometric networks. It specifically employs the VOS (Visualization of Similarities) technique to determine the position of nodes in the visualization (Van Eck and Waltman, 2014).

VOSviewer is a software tool used for constructing and visualizing bibliometric mapping based on the visualization of similarities (VOS) (VOSviewer, 2023). VOSviewer has been utilized for bibliometric analysis of the publications. In VOSviewer, a map is created via the map wizard based on network data, bibliographic data, or text data. In network data, the VOSviewer map and network files are the data source. The bibliographic database and reference manager files within referencing APIs are data sources in bibliographic data and text data (Van Eck and Waltman, 2023). The map of bibliographic data provides connections between items according to co-authorship, co-occurrence of keywords, citation, co-citation, and bibliographic coupling. The map of text data provides relations according to the co-occurrence of terms in the abstract and title. There are two ways of counting occurrences to create links: binary and full counting. Binary counting counts how many documents the term occurs in at least once. Full counting counts how many occurrences of the term occur in all documents (Van Eck and Waltman, 2023).

In this study, four networks were created from (i) abstract, (ii) title, (iii) keywords, and (iv) citations. The software (VOSviewer) reads data from reference manager files. Before visualization in VOS software, the citations of the publications were conveyed into reference files (RIS) in Mendeley. The RIS files are the input of this network. The abstracts of the references were used to create a network. The terms were collected under specific groups, based on the widely used terms related to the research topic, scope, and method. A dictionary was created to collect term groups with similar meanings. To create a dictionary, first, all the keywords extracted by the VOSviewer Wizard are obtained and converted into a comma-separated file (CSV) format. In this way, the keywords are defined. During the scrapping process, there is an option to scrape the terms as word groups. With this option, terms like city and smart city, or data and big data, are not confused and are considered as separate terms. The dictionary is imported into the map wizard as a thesaurus file. A thesaurus file has two columns: a label column and a replace-by column. The first line includes the keywords, while the second line contains the corresponding terms, defined by the researcher, to merge synonyms and similar-meaning words into a single term. For instance, data-related terms are grouped under big data, activity places (venues, restaurants, etc.) are grouped under urban activity, and decision support and decision-making terms are grouped under decision support. Different city names are grouped under city-level analysis. The plaza, public space, square, recreational site and recreational area, leisure, urban park, and park are grouped under urban POS category. Urban POS stands for the urban public open space. Geolocated data, geotagged data, georeferenced data, location information, geo-tag, geospatial, location-based social network (LBSN), and volunteered geographic information (VGI) data are grouped under geolocated data. Visualization-related terms and methods are grouped under data visualization. Similarly, different statistical methods are grouped under spatial statistical techniques, and various social media data-related terms are grouped under social media network. Accordingly, 725 terms were grouped into 84 items extracted by grouping. In this way, the occurrence of each term is increased. The dictionary database is shown in **Appendix B.**

The group names were extracted by the researcher from the keywords that indicate the scope and topic of the studies, as shown in **Table A.1**, which displays the reference studies. The network was created by extracting the text data from abstracts. The full counting method was employed to increase the occurrences of the terms and complicate the network. The minimum number of occurrences was set at 8, and 51 items met the threshold. Based on the relevance score, 90% of the items were selected. For 45 items, 3 clusters of 15 items were created each. In this network, there were 45 items, 874 links, and a total link strength of 16,362. For the title network, the same dictionary was used to compile closely related terms under the same group. The words in the title of the references were imported from RIS files as text data and counted using the full counting method. Eighty-six items were obtained, and twenty of them occurred more than 2 times. As a result, 3 clusters were obtained with 42 links and a total link strength of 58. **Table A.1** displays the exported results of the terms in the abstracts of the references. In a similar vein, another map was created based on bibliographic data according to keyword co-occurrence. The input information about the publications was taken from the reference database in RIF format. The same dictionary was used to group the terms in the keyword. In the keyword network, there were 26 items with 3 clusters, 112 links, and 166 total strengths. The keywords are listed with their occurrences and relevance scores in **Table A.2**. Another network map was created for citations of the authors using the bibliographic database. The documents were imported via the Digital Object Identifier (DOI) of the publication. Citation is the unit of analysis for the documents. The threshold for the citation of each document was set at a minimum of 10. Among 27 documents, 19 items met the threshold. For these 19 items, 4 clusters were created with 33 links and a total strength of 65. Citation data are listed with their occurrences and relevance scores in **Table A.3.**

Data visualization includes network visualization, overlay visualization, and density visualization. All of the networks were analyzed based on the publication years in overlay visualization and occurrence frequency in density visualization. In ***network visualization,*** items are represented by labels and circles (Van Eck and Waltman, 2023). The size of the label and circle depends on the weight of the items. The items are connected with links, which display the relatedness of the two publications in terms of co-citation, keyword, abstract, title, etc. ***Overlay visualization*** utilizes the same network, but the colorization changes based on quantitative information such as the publication year, citation score, or impact factor. ***Density visualization*** indicates the density of the clusters based on the number of neighboring items and their weights, represented by colors ranging from blue to green to yellow. Accordingly, as the number of items increases in the neighborhood of a point within their weight, the color changes from blue to yellow (Van Eck and Waltman, 2023). Overall, this methodology provides a clear and coherent approach to identifying and analyzing relevant articles on the use of big data in urban studies and provides a detailed explanation of the processes and tools used in the analysis.

REFERENCE STUDIES

Researchers have shown an increased interest in the use of geolocated data coming from LBSN and social media networks in urban analysis to generate useful knowledge about cities' functioning and livability within the scope of the big data approach. As the literature review demonstrates, the locative information obtained from geolocated data is useful to reveal the ephemeral layer of cities: where people are concentrated (Garcia et al., 2015; Cranshaw et al., 2011; Kisilevich et al., 2013), what the common travel patterns are (Batty et al., 2012; Sun et al., 2016), and visit patterns (Girardin et al., 2008), as well as the urban form factors driving the use of a place (Lie et al., 2019; Zhang & Zhou, 2018).

In terms of urban activities, Wu and colleagues (2016) employ POI data from Foursquare check-ins and online real estate data for land price analysis in order to uncover the relation between POIs and land price. Jiang et al. (2015) use the Yahoo platform to estimate land use based on POI data and achieve a higher level of accuracy compared with traditional methods. Zhang and Zhou (2018) utilize Foursquare check-in data with other spatial attributes through regression analysis to examine how the attributes, locations, and accessibility of a park affect the check-in number in the park. Kobas and Ensari (2018) enrich social network data analytics by overlapping multiple datasets - including real estate data, geolocation from social media (Instagram) within its context, accessibility data from public transportation, and commercial resources - with the intent to examine demographic and economic trends through data visualization.

The location data (check-ins) from Foursquare is employed to identify city venues and POIs (Sun et al., 2016). Cranshaw and colleagues (2012) use Foursquare check-ins within Twitter tweets to understand the collective movement patterns of users and identify lively urban areas. Abbasi and colleagues (2015) use a text mining approach to reveal the trip purposes of tourists from tweets using Twitter. Marti and other researchers (2017) identify successful public areas through Foursquare and analyze their relation with historic centers and the main axes of the city to understand the correlation between location and the vibrant characteristics of the space. Text data from Twitter has been used to identify public sentiment about social life (Tuncer and You,

2018) and urban happiness (Guo et al., 2021). Image data from geolocated services (such as Instagram and Flickr) is useful for defining the place identity of the public. Huang et al. (2021) and Jang et al. (2019) utilize images and texts to generate cognitive mapping referring to Lynchian elements.

User-generated and sports data apps (fitness and other sports applications) have been studied to understand the sports habits of citizens and detect popular exercise areas (Mora et al., 2018; Delhoyo et al., 2018; Balaban and Tuncer, 2016). Balaban and Tuncer (2016) measure the runnability of a city and factors related to sports activities of citizens using fitness data from mobile applications, crime data, weather data, and spatial data to understand the sports habits of residents, and accordingly, predict recreational spaces in the city. Mora and colleagues (2018) overlap GPS tracking data with mobile applications used for sports activities to create knowledge aiming to support decision-making for planning recreational areas. Furthermore, travel patterns in public transportation have been evaluated using travel card data (London Oyster card) to understand daily travel flow and the volumes in hubs on the rail network (Batty et al., 2012). Photo-sharing platforms (Panoramio and Flickr) have been utilized for spatiotemporal analysis to identify attractive places (Kisilevich et al., 2013; Garcia-Palomares et al., 2015) and estimate the number of visitors at tourist attractions (Wood et al., 2013; Girardin et al., 2008). Manal and colleagues (2018) analyze viewpoints and view scenes/tags of Flickr photos to elicit preferred heritage attributes and their significance. This research examines related works under the subjects of (i) urban activities and density, (ii) user movement and visits, and (iii) urban perception. ***Table 1*** exhibits related studies using LBSN data in urban research.

RESULTS

Figure 1 and **Figure 2** present quantitative facts about the terms as clustered pie charts and pie chart diagrams, respectively. The first diagram **(Figure 1)** displays the occurrence of each term group. Notably, the most occurred term groups are city (73), city scale analysis (65), research (60), place concept (55), big data (50), and social media network (40). The second most occurred term groups include place identity (34), urban points of interest (32), spatiotemporal analysis (30), urban functioning (30), user (29), urban activity (27), neighborhood (24), visit (24), analysis (22), points of interest (22), check-in (21), location (21), sentiment analysis (20), photo-sharing services (20), quantitative (20), urban decision support (20), and user activity (20).

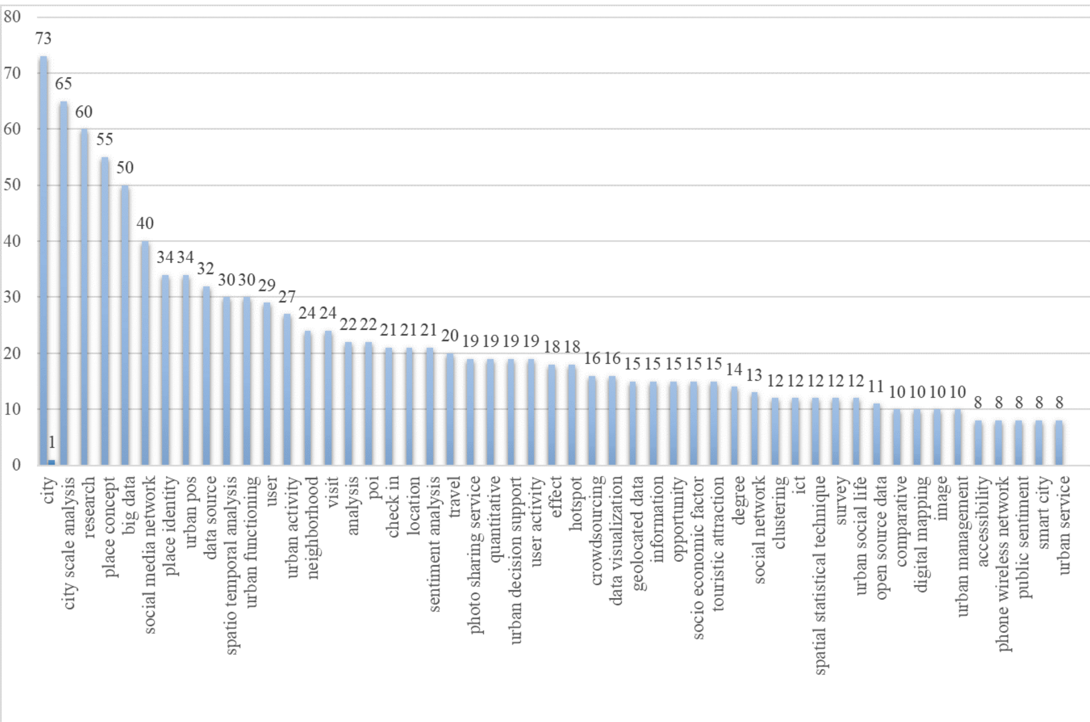


Figure 1. Clustered bar chart graphic showing the occurrence of the terms.

In the reference studies, 54% of the overall terms are related to research topics, which primarily focus on urban phenomena, place, context, and the scale of the studies. **Figure 2** displays the distribution of terms based on specific categorization. Among the most cited terms, city scale analysis (city names), place concept, neighborhood, urban points of interest (parks, green areas, plazas), touristic attractions (historic city centers, cultural sites), points of interest (POIs), and hotspots (city centers) define the location.

Table 1. Reference studies to reveal the urban activities, user movement, density, perceptions, and preferences.

Subject: Urban activities (POI) and user density **aim:** *to understand land use, POIs, social life, land value, socio-economic trends*						
Study	**Reference**	**Topic/ Scope**	**Keywords**	**Data types and tools used**	**Methodology**	**Decisions supported**
Spatial and Social Media Data Analytics of Housing Prices in Shenzhen, China	Wu, C., Ye X., Ren, F., Wan, Y., Ning P., Du, Q. (2016).	**LAND VALUE** the factors affecting house prices / the impact of POI on land value	Housing prices, urban China, big data, spatial patterns, geographically weighted regression.	Check-in data and online real estate data from websites	Kernel density estimation (KDE) for POI density and the hedonic price method (HPM) to estimate pricing and the geographically weighted regression (GWR) method for spatial analysis	To support decisions related to land value
Mining Point of Interest Data from Social Networks for Urban Land Use Classification and Disaggregation	Jiang, S., Alves A., Rodrigues, F., Ferreira, J., Pereira, F.C. (2015).	**LAND USE -POIs** the relationship between land use attraction points through social media platforms	Information extraction Points of interest Volunteered geographic information	Yahoo platform (for POIs) and land use data	Machine learning for POI classification (destiny-based clustering) and estimation of the land use based on POIs	To support decisions related to land value
Crowdsourcing functions of the living city from Twitter and Foursquare data	Zhou & Zhang (2016)	**URBAN FUNCTIONS (LAND USE)** the relation between urban functions and social sensing	Social media data Geographic Information Systems City characteristics Urban planning	geo-referenced tweets from Twitter and Foursquare check-ins of venues (restaurant/store/ cinema)	Foursquare API to classify venues and Support Vector Machine (SVM) to classify tweets/ PostGIS for mapping and R for analyzing	Support decisions related to urban functions
Quantitative Comparison of Open-Source Data for Fine-Grain Mapping of Land Use	Deng, X., Newsam, S. (2017)	**URBAN FUNCTIONS (LAND USE)** the comparison of land uses in administrative data and activities in VGI	Land use, points of interest, volunteered geographic information (VGI)	Google Places listings Yellow Pages and Bing Map Open Street Map	Web scraping and quantitative evaluation with statistics test for comparing pairwise comparison of administrative and VGI data	Support decisions related to urban functions and land use planning
Understanding User Activity Patterns of the Swarm App: A Data Driven Study	Lin S., Xie R., Xie Q., Zhao H., Chen Y. (2017)	**URBAN ACTIVITIES** the comparison of urban activities in different cities	Location-Based Services; Check-In, Swarm App, Spatial-Temporal Analysis, City Computing, Human-centered computing	Foursquare check-ins	Web scraping of Foursquare data and spatio-temporal analysis within the correlation tests	Support decisions related to urban activities and user activity pattern
Measuring urban activities using Foursquare data and network analysis: a case study of Murcia	Agryzkov T., Martí P. , Tortosa L. & Vicent, J.F. (2016)	**URBAN ACTIVITIES** urban activities and social networks	Urban analysis Social networks analysis Street networks PageRank algorithms Data visualization	Foursquare check-ins and user comments	Web scraping of Foursquare data of venues Analysis through network centrality algorithm Compare field study with social media analysis in terms of venue categorization	To support decisions related to user preferences on urban activities
Social Activity in Gothenburg's Intermediate City: Mapping Third Places through Social Media Data	Adelfio, M., Serrano-Estrada, L., Ciriquian P.M., Kain, J., Stenberg, J. (2020)	URBAN ACTIVITIES - Urban sprawl Understanding urban activities in the intermediate cities	Intermediate city third places Social media data (SMD)	Foursquare Twitter Google Places	Web scraping Descriptive analysis	To support decisions in urban planning related to urban sprawl
Revisiting the Spatial Definition of Neighborhood Boundaries: Functional Clusters versus Administrative Neighborhoods	Marti, P., Serrano-Estrada, L., Nolasco-Cirugeda, A., Baeza, J.L.(2021)	**NEIGHBORHOOD BORDERS** define the borders of neighborhoods based on functions and compare with	Neighborhood boundaries, functional clusters, urban economic activities, Google Places social networks	Google Places listings	Web scraping Functional clustering	To support decisions related to urban economic activities
Social Activity in Gothenburg's Intermediate City: Mapping Third Places through Social Media Data	Adelfio, M., Serrano-Estrada, L., Ciriquian P.M., Kain, J., Stenberg, J. (2020)	**URBAN ACTIVITIES -** Urban sprawl understanding urban activities in the intermediate cities	Intermediate city third places Social media data (SMD)	Foursquare Twitter Google Places	Web scraping Descriptive analysis	To support decisions in urban planning related to urban sprawl

The context of the studies makes up 13% of the most occurred terms, including city, big data, geolocated data, information, urban decision support, and urban management. Urban phenomena account for 28% of the occurrences, encompassing urban functioning, place identity, user activities, urban activity (user movement, sports), travel, visit, urban social life, and socio-economic factors. Data sources represent 19% of the overall terms. Widely used data sources for analysis include geolocated data sources, social media networks, Twitter, check-ins (Foursquare, Swarm, Weibo), photo-sharing services (Flickr, Panoramio), open data sources (real estate listings, Airbnb listings, websites), and surveys. Analysis methods constitute 21% of the overall terms, featuring spatiotemporal analysis, sentiment analysis, data-driven analysis, computational methods, digital data mapping, data visualization, spatial statistical techniques (statistical methods), clustering, comparison, and classification. The results indicate that big data studies have been employed to understand urban dynamics from various perspectives, ranging from social and economic factors to user activity, movement, visit, and perception. Numerous data-driven analysis methods have been utilized to comprehend the spatial distribution of urban phenomena, extending from statistics to visualization and mapping. The results also showcase the diversity of geolocated data sources, grouped under geolocated data, used to measure the pulse of places.

The network results of the text data analysis of the abstract terms (Url-1) are shown in **Figure 3**. This network is normalized with association strength. In this network, **Cluster 1** (represented in red color) includes terms related to urban activities, urban

Table 2. Reference studies to reveal the urban activities, user movement, density, perceptions, and preferences.

Impact of Airbnb on the Gentrification Process: The Case of Rasimpaşa Neighborhood	Uzgören, G. & Türkün, A. (2018)	**GENTRIFICATION THROUGH ACCOMMODATIONS**	Airbnb, gentrification, displacement, globalization.	Airbnb accommodation properties and rental price (real estate web site)	Mixed research methods Quantitative data with web scraping, Survey field study with people, Comparing the two analysis methods	To support decisions related to land use planning regarding gentrification
Recreational visits to urban parks and factors affecting park visits: Evidence from geotagged social media data	Zhang & Zhou (2018)	**PARK USE** understanding the effective factors of urban park use	Park use, geotagged check-in data, social media, park attributes, park location	Social media platform (Weibo) check-ins and shares and spatial data	Statistical methods (regression) to investigate the correlation Weibo API (4year datasets) to scrape recreations ArcGIS for spatial analysis	To support spatial decisions related to the use of public open space
Deciphering the recreational use of urban parks: Experiments using multi-source big data for all Chinese cities	Li, F., Li, F., Li, S. Long, Y. (2019)	**PARK USE** The relationship of urban park use with other social reinforcement in different cities	Weibo check-ins Park attributes Regression models Park usage China	Social media platform (Weibo) check-ins of the park, population GDP per capita	Weibo API check-ins for park visits Multiple linear regression models to analyze the affecting factors ArcGIS to data mapping	To support spatial decisions and socio-economic trends related to the use of public open space
Web Scraping and Mapping Urban Data to Support Urban Design Decisions	Kobas & Ensari (2018)	**URBAN ACTIVITIES AND USE** understanding how people use the urban area and how urban use creates spatio-temporal change	Geospatial data, Urban data, Urban design decision support, Web scraping	Instagram shares, data from online food order/delivery, online real estate, land use data, public transportation	Web scraping methods with Python apis Data visualization for decision support with Meerkat plugin of Grasshopper	To support urban design decisions regarding socio-economic trends
Subject: User movement, travel and urban sports **aim:** *to detect travel routes and identify recreational/sport areas*						
Study	**Reference**	**Topic/ Scope**	**Keywords**	**Data types used**	**Methodology**	**Decisions supported**
Smart Cities of the Future.	Batty M., Axhausen K., Fosca G., …. Portugali Y. (2012).	**DAILY TRAVEL MOVEMENT** monitor daily flow of community and volume of public transportation	Geotagged photographs, Photo-sharing services, Spatial distribution, Tourist attractions	London Oyster travel data	Construct multimodal trips (origins and destinations) Calculate movement flows and passenger volume to estimate optimum passenger capacity	Support decisions related to mobility and travel
Visualizing urban sports movement	Balaban and tunçer (2016)	**Urban sports** the runnability of the city reveals related factors of sport	Sports activity, big data, urban visualization, fitness applications	Data from mobile applications fitness related / crime data/ weather data/ spatial data/ fitness data	Integrate endomondo fitness application data with gps data as relational databases in mysql data visualization of different datasets in gis	Support decisions related to urban sport activities
Analysis of Social Networking Service Data for Smart Urban Planning	Mora, H., Perez-Delhoyo, R., Parades-Perez, J.F., Molla-Sirvent, R.A. (2018).	**URBAN SPORTS** identify the most preferred fitness route	Social networks, ambient behavioral analysis, urban planning, decision making, sustainability, accessibility	The social network platforms used for sport activities/data from mobile applications (GPS tracking data)	Scrape geolocation data Establish route map and overlap street view to understand the frequently-visited places and the reason for selection	Support decisions related to recreational areas
Subject: User density, perception and preferences **aim:** *to identify attractions and points of interests (POIs)/understand the purpose and reasons for visit/offer proposal for visit*						
Study	**Reference**	**Topic/ Scope**	**Keywords**	**Data types used**	**Methodology**	**Decisions supported**
Using social media to quantify nature-based tourism and recreation	Wood, S.A., Guerry, A.D., Silver J.M., Lacayo, M. (2013).	**NATURAL PARKS** define visitor density/rate in year and week	Sports Activity, Big Data, Urban Visualization, Fitness Applications	Flickr georeferenced data	Scrape visitation data through Flickr, Calculate the visitation rate, Estimate user-days per year and each week Compare visitation rate with real visits through statistical analysis ANCOVA	To support decisions related to recreational visits
Automatic construction of travel itineraries using social breadcrumbs	De Choudhury, M., Feldman, M., Amer-Yahia, S., Golbandi, N., Lempel, R., & Yu, C. (2010).	**TRAVEL ROUTES** create travel itinerary to monitor visitors' route	Flickr, geo-tags, mechanical turk, orienteering problem, social media, travel itinerary	Flickr georeferenced data	Scrape photos from Flickr API Take POIs from Yahoo and combine with photos' geolocation photo-POI mapping and construct time paths to sequence based on visiting time Create itineraries from time paths, justify with surveys	To support decisions related to visitor routes

functioning (urban function and urban form attributes), and neighborhoods. According to the keywords in this cluster, the studies focus on urban activity and functioning attributes within the socio-economic and socio-cultural attributes of neighborhoods. Socio-cultural factors are associated with the diversity, ethnicity, age, and gender of the citizens within the cultural attributes of the city, while socio-economic factors are associated with economic trends, housing, and real estate data. Clustering analysis of check-ins is useful within the spatial statistical technique to measure density in urban activities and specify points of interest (POIs). Socio-economic data aids in understanding the social and economic value of urban activities. This cluster comprises 15 terms, with the most frequent ones being decision support, digital mapping, urban activity, urban functioning, hotspot, check-in, big data, and user **(Figure 3).** This cluster is classified as ***urban activities and density.*** The studies in this cluster aim to support decisions in urban activities and functioning. The results of these studies would be useful for decision-making in neighborhood density and activities **(Figure 3).** Based on the occurrence of the terms in the abstract, the most associated terms are, respectively, contextual concepts (city, place concepts), then urban phenomena (urban POS (public open space), functioning, social life, and identity), and lastly, these urban phenomena are connected with analysis methods (data-driven analysis, mapping, and visualization).

Table 3. Reference studies to reveal the urban activities, user movement, density, perceptions, and preferences.

Utilising Location Based Social Media In Travel Survey Methods: Bringing Twitter Data Into The Play	Abbasi A., Rashidi T.H., Maghrebi M., Waller S.T. (2015)	**TRAVEL ROUTES** Twitter data for travel demand analysis (as topic modeling)	Social media data, Crowdsourcing, Travel attributes, Twitter, Tourists, Travel pattern	Twitter data	Detect activity location from geolocation through Twitter Search API - Data mining of personal tweets' location to query the activity location Text mining techniques to understand trip purpose from tweets	To support decisions related to travel pattern
Photographing a City: An Analysis of Place Concepts Based on Spatial Choices	Schlieder and Matyas, (2009)	**TOURISTIC ATTRACTIONS AND VISITING ROUTE** analyze point of views (POV) of photographers	Geospatial semantics Place concepts Geographic recommender systems Social tagging	Flickr georeferenced data	Measure the popularity of POV (point of visit) with Flickr API Cluster the POVs with clustering algorithm (Heatmapper) To understand the visit route of users	To support decisions related to visitor density at touristic pois
Digital Footprinting: Uncovering Tourists with User-Generated Content	Girardin, F., Calabrese, F., Fiore, F. D., Ratti, C., & Blat, J. (2008).	**TOURISTIC ATTRACTIONS AND VISITING ROUTE** define user-related spatio-temporal data	Pervasive user-generated content Reality mining Information visualization	Flickr georeferenced data and cell phone network data	Scrape visit data through Flickr Compare with statistical analysis (standard deviation analysis) use KML* for visualization of geolocated data process data into Google Earth	To support decisions related to visitor density and movement at touristic pois
Towards acquisition of semantics of places and events by multi-perspective analysis of geotagged photo collections	Kisilevich, S., Keim, D., Andrienko, N., & Andrienko, G. (2013).	**ATTRACTIVE CITY CENTERS** analyze events and places using geotagged photo collections	Spatio-temporal clustering Semantic enrichment Geotagged photo collections	Panoramio and Flickr photo shares Multisource data types (geolocation and text data)	Analysis of movement data Define spatio-temporal clusters of visit using DBSCAN* algorithm/ Apply semantic analysis (titles/tag/content/GeoNames) with text mining and grouping	To support decisions related to visitor density and movement at touristic pois
Identification of tourist hot spots based on social networks: A comparative analysis on European metropolises using photo-sharing services and GIS	Garcia-Palomares, J.C., Gutierrez, J., Minguez, C. (2015).	**TOURISTIC ATTRACTIONS** analyze tourist dynamics in tourism geography	Geotagged photographs, Photo-sharing services, Spatial distribution patterns, Tourist attractions, European cities. Spatial statistics	Panoramio photo data and GIS spatial data	Scrape Panaramio API and Flickr API Apply spatio-temporal analysis in ArcGIS to identify interest points and determine tourist dynamics Identify clusters with statistical technics in GIS (Moran I)	To support decisions related to visitor density touristic pois
Identifying the City Center Using Human Travel Flows Generated from Location Based Social Networking Data	Sun, Y., Fan H., Li, M., Zipf, A. (2016).	**ATTRACTION CENTERS** define activity centers and venues in the city	Center identification, Human mobility, Local Getis-Ord, DBSCAN, Grivan-Newman	Foursquare check-ins and mobility data (travel data from public transportation and taxi) GPS tracking data	Conduct spatial statistical analysis (DBSCAN/ Grivan-Newman* and LGOG* methods), Detect hotspots and clustering of hotspots (attraction centers), Identify the borders of city center with different clustering methods	To support decisions related to user density and movement at pois
The Livehoods Project: Utilizing Social Media to Understand the Dynamics of a City	Cranshaw, J., Hong, J. I., & Sadeh, N. (2011)	**ATTRACTION CENTERS** livehoods: dynamic areas in the city urban vibrancy/ vitality	smart cities, location-based social networks, Foursquare, check-ins	Foursquare check-ins and Twitter data	Clustering model for mapping based on collective behavior of users, Combine quantitative methods with qualitative methods through interviews for identifying livehoods	To support decisions related to the attractiveness of hotspots in the city center
Using Locative Social Media And Urban Cartographies To Identify And Locate Successful Urban Plazas	Marti, P., Serrano-Estrada, L., Nolasco-Cirugeda, A. (2017)	**URBAN PLAZAS** identify driving factors of successful plazas (relation with urban axes)	Public space Plaza Square Social networks Livable spaces Social spaces	Foursquare georeferenced data (check-ins)/ category/ total users/ pictures and shares with spatial data from cartography	Data collection through Foursquare API classification of venues/ Detect most popular plazas and compare with spatial data (from cartography)	To support decisions related to the attractiveness of urban squares
Mapping historic urban landscape values through social media	Manal, G., Rodersb, A. P., Teller, J. (2018)	**HISTORICAL URBAN HERITAGE SCAPE** the analysis of the view- point from Flickr photos to elicit preferred heritage attribute	Historic urban landscape Social media Flickr Cultural heritage Everyday landscape	Flickr georeferenced photos/tags	Map the spatial distribution of Flickr photos to identify preferred heritagescape. Create classification model to elucidate the heritage attributes Reveal tag analysis to determine heritage significance via semantic map	To support decisions related to visitations, visitor perception and preferences
Understanding Happiness in Cities using Twitter: Jobs, Children, and Transport	Guo, W., Gupta, N., Pogrebna, G., Jarvis, S.A. (2021)	**PUBLIC SENTIMENT**	Happiness, Social media data, City demographics Socioeconomic parameters, Positive sentiments	Twitter data	web scraping of tweets from Twitter apply natural language processing (NLP) techniques sentiment analysis	To support decisions related to citizen satisfaction from socioeconomic conditions for urban happiness

Cluster 2 (represented in green color) includes terms concerned with social life, place identity, and public sentiment. According to the keywords in this cluster, the studies deal with identifying public sentiment, place identity, and urban social life through sentiment analysis of images and Twitter. The place identity group encompasses the image of the city constructed through Lynchian elements and cognitive mapping, while the social life group encompasses the everyday life habits of citizens. Sentiment analysis of Twitter is useful for revealing public sentiment and social habits. Sentiment analysis of images from photo-sharing services is useful for revealing place identity elements. Data visualization is the method used to represent place identity and sentiments. This cluster involves 15 items, with the most frequent ones being crowdsourcing, data visualization, image, place identity, public sentiment, social life, and sentiment analysis. This cluster is classified as ***urban perception.*** Accordingly, the studies in this cluster aim to support decisions in place relationships in terms of place identity and sentiment. The results of these studies would be useful for decision-making in user perception of urban places. The analysis unit of the studies in clusters 2 and 3 is the city scale **(Figure 3).**

Table 4. Reference studies to reveal the urban activities, user movement, density, perceptions, and preferences.

Exploring public sentiments for livable places basedon a crowd-calibrated sentiment analysis mechanism	Tuncer, B. You (2018)	**PUBLIC SENTIMENT** the effects of public sentiments of social networks in the domain of place design	Sentiment analysis Twitter, Machine learning algorithms Pattern classification Livable places Crowd-calibrated sentiment analysis mechanism	Twitter and Instagram hashtags and content	Web scraping of tweets from Twitter sentiment analysis Develop a crowd calibrated geo-sentiment analysis mechanism Classify the positive-negative-neutral sentiments	To support decisions related to urban design considering public sentiments
The image of the City on social media: A comparative study using "Big Data" and "Small Data" methods in the Tri-City Region in Poland	Huang, j., obracht-prondzynska, h., kamrowska-zaluska, d., sun, y., & li, l. (2021).	**Public image**	Kevin lynch city image social media analytics tri-city poland	Instagram posts and hashtags	Web scraping of hashtags social media data analytics lynch methods mapping-survey, gis analysis of administrative data clustering analysis (dbscan- kde) comparison of three datasets	To support decisions related to urban identity, urban image
Crowd-sourced cognitive mapping: A new way of displaying people's cognitive perception of urban space	Jang, k. M., & kim, y. (2019).	**Public image**	Cognitive mapping, georeferenced text data, crowd-sourced cognitive map, urban identity, visualization	Textual data of instagram hashtags	Scrape instagram photos and photo data Create crowdsourced cognitive map using instagram hashtags	To support decisions related to urban perception through cognitive map, evaluation of urban identity

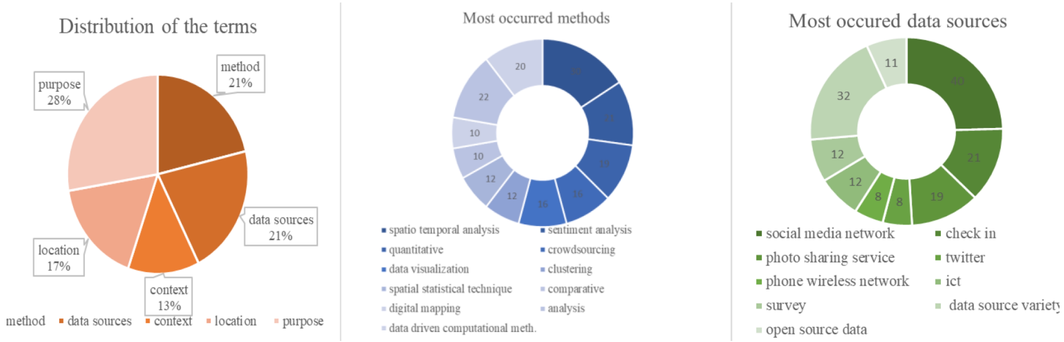


Figure 2. Pie chart graphics displaying distribution of the terms.

Cluster 3 (represented in blue) includes terms related to tourism, visit, urban POS, travel, and accessibility. Based on the keywords, the studies deal with visits to urban POS, touristic visit sites, and visitation rates. Photo-sharing services are the geolocated data source used to measure density in visited places. Phone and wireless networks are other data sources to track visitors' movement and density. The spatiotemporal patterns of visitor and travel data have been revealed in the studies. This cluster comprises 14 items, with the most frequent ones being travel, accessibility, tourism, urban POS, visit, spatiotemporal pattern, photo-sharing service, and phone wireless network. This cluster is classified ***as user movement and visit.*** Accordingly, the studies in this cluster aim to support decisions in visits and travel. The results of these studies would be useful for decision-making by visitors and mobility or accessibility **(Figure 3).**

Table 5. The terms in these three clusters.

Cluster	Cluster Color	Terms in this cluster
Cluster 1	Red	urban activities, urban function, urban form attributes, neighborhood, socioeconomic attributes, sociocultural attributes, urban density
Cluster 2	Green	public sentiment, place identity, urban social life, sentiment analysis, images, Twitter, urban perception
Cluster 3	Blue	tourism, visits, public open spaces, touristic visit sites, visitation rates, user movement, accessibility, visit and travel, urban POS (public open spaces)

The overlay visualization map **(Figure 4)** illustrates the chronological development of the reference studies. The years in the legend represent the average publication year of the references. As seen, the publication years of the sources range from 2015 to 2018, indicating that they are up-to-date and closely related. There is a notable increase in published references over the five-year span. Studies related to the use of geolocated data began around 2008 and accelerated until 2018, showing that big data is a popular research topic in urban studies. The initial studies using big data were associated with urban movement (visit and travel) and points of interest (POIs) through the use of photo-sharing services (Flickr, Panoramio) and tracking from phone and wireless

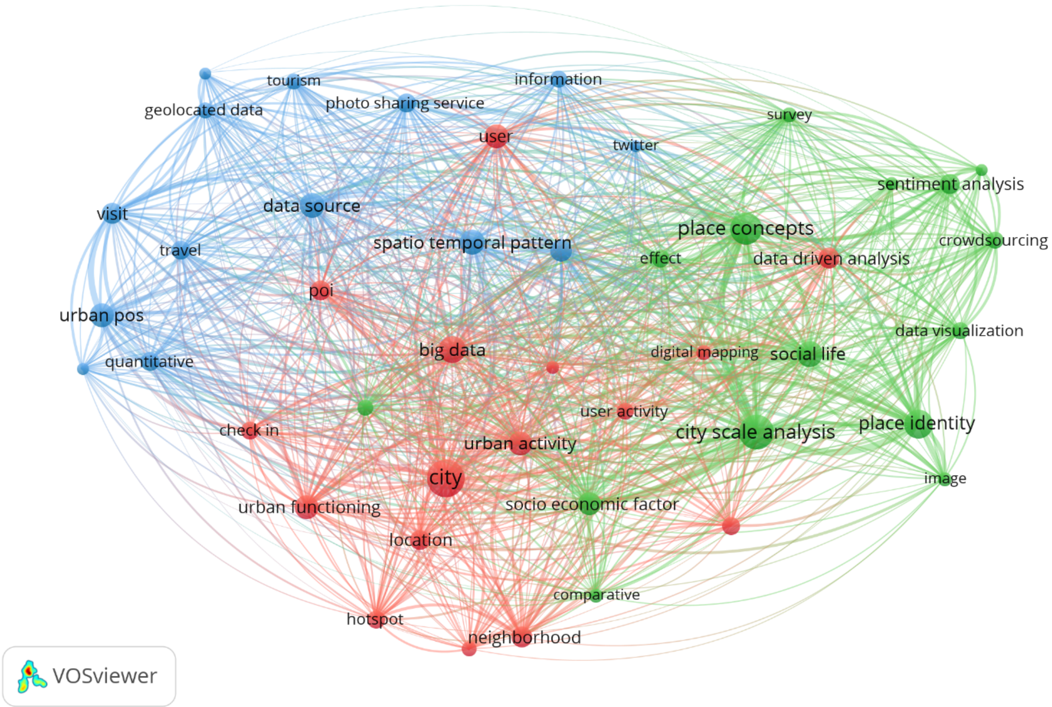


Figure 3. Network visualization results.

networks around 2010. With the ubiquity of social media and check-in data (Twitter, Instagram, Foursquare, etc.) and various analysis methods, research topics expanded to include urban activities and perceptions around 2015 and later. The density map **(Figure 5a)** displays that city, city scale analysis, big data, place concepts, place identity, social media network, spatiotemporal pattern, and social life are the most densely clustered items. Cluster density results **(Figure 5b)** show that cluster 1 (urban function) is the most densely clustered group, cluster 2 (urban perception) is moderately dense, and cluster 3 (urban movement) is the least dense cluster.

Figure 6 presents the network visualization of the titles, while **Figure 7** displays the network visualization of the keywords based on their co-occurrence. Both networks provide information about the research topic, applied methods, and geolocated data types. The results of the title and keyword networks support the classification derived from the abstract text. Accordingly, the ***urban activities and density*** cluster **(Cluster 1)** includes check-in, user activity, urban activity, digital mapping, and data-driven analysis. Based on the keyword network results, urban activities, functioning, and points of interest are clustered together. The urban perception cluster **(Cluster 2)** comprises public sentiment, sentiment analysis, place identity, crowdsourcing, place concepts, and spatiotemporal patterns. Keyword network results confirm the clustering of public sentiments, place identity, urban activities, functioning, and points of interest.

The **urban movement and visit cluster (Cluster 3)** encompasses digital mapping, tourism, travel, urban points of interest (POS), user activity, tourism, urban functioning, and social media networks. As justified by the keyword network results, urban POS, user activity, and touristic attractions are clustered together. The keyword network highlights data-driven computational analysis methods, urban decision support, and social media networks as two significant nodes connecting the terms. Based on the results, the urban perception cluster has clearer borders with its terms. However, the boundaries between urban activities-density and urban movement-visit are somewhat blurred due to the hotspot and urban functioning terms.

The term *"urban decision support"* has the strongest relationship with city, neighborhood, and place concepts (urban site, area, place, environment), as well as city scale analysis (cities as case studies). The relationship strength between these terms and urban decision support is higher than 30, which indicates a strong relationship. Following this, decision support has strong relationships with methods such as digital mapping, data-driven analysis, sentiment analysis, crowdsourcing, data visualization, spatiotemporal patterns, and clustering. It also has strong relationships with urban research topics, including place identity, data sources, urban functioning, social life, and socioeconomic factors, with a minimum link strength of 15. As the minimum link strength decreases, the relationships expand to encompass the majority of research topics and methods used in the reference studies.

The term "urban decision support" appears 18 times and connects to 40 items with a total link strength of 482. **Figure 8** displays the position of this term in the relationship network. As a result, the term "urban decision support" is primarily strongly related to

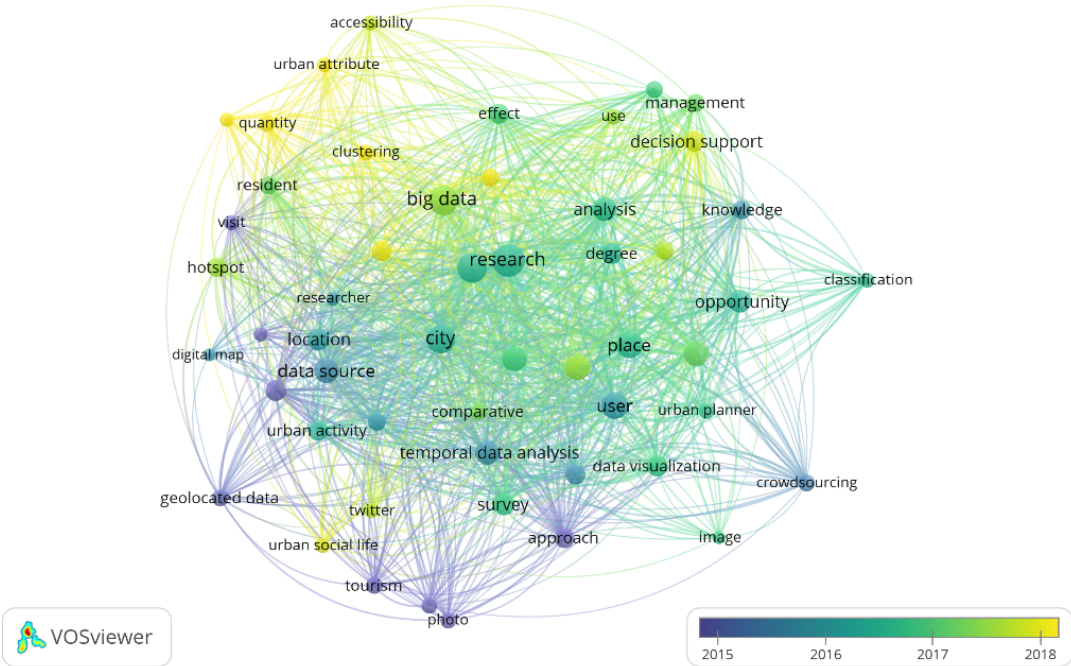
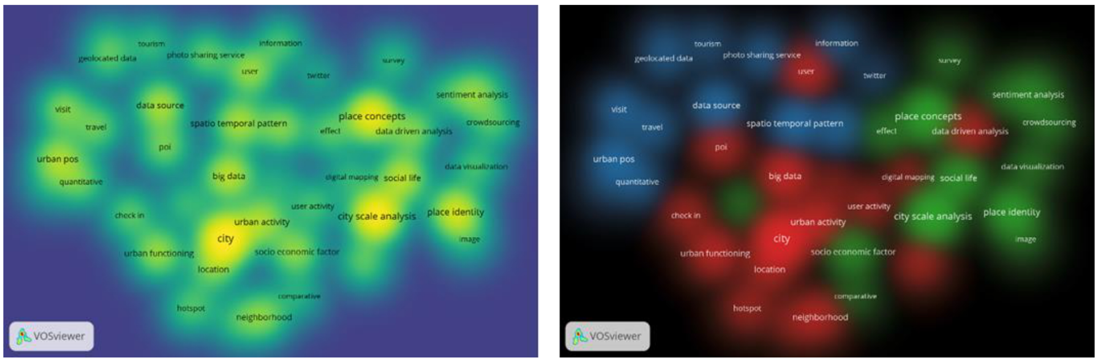


Figure 4. Overlay visualization results of the network.



a b

Figure 5. Density visualization results based on items **(5.a)** and cluster density **(5.b)**.

contextual concepts (city, neighborhood, and place concepts), followed by applied methods (such as data visualization, mapping, data-driven analysis, and spatial statistical techniques) and urban research topics (such as urban activities, urban functioning, user movement, and place identity). In terms of keywords, decision support has strong associations with social media networks, urban social life, spatial statistical techniques, Geographic Information Systems (GIS), urban management, smart city, geolocated data, data-driven computational analysis, urban functioning, and urban activities **(Figure 9).** These associations also indicate a significant relationship between decision support and analysis methods, and their contribution to understanding urban dynamics (urban activity, sentiment, social life, management).

The citation results, as shown in **Figure 10,** indicate that the selected references span the years between 2008 and 2022. Approximately 33% of the references (9) were published up until 2016, while about 66% of the references (18) were published up until 2022. These results demonstrate the growing popularity of big data in urban studies in recent years. The journals in which the reference articles were published have an average of 25 citations, with some outliers. Among them, The European Physical Journal Special Topics, Scientific Reports, IEEE Pervasive Computing, Applied Geography, Computers, Environment and Urban Systems, and Landscape and Urban Planning are the most cited journals, with citations above 100 (as seen in **Table A.3**). According to the documented citation results, the citations of the documents range from 10 to 100 and above. As the citation results show, the

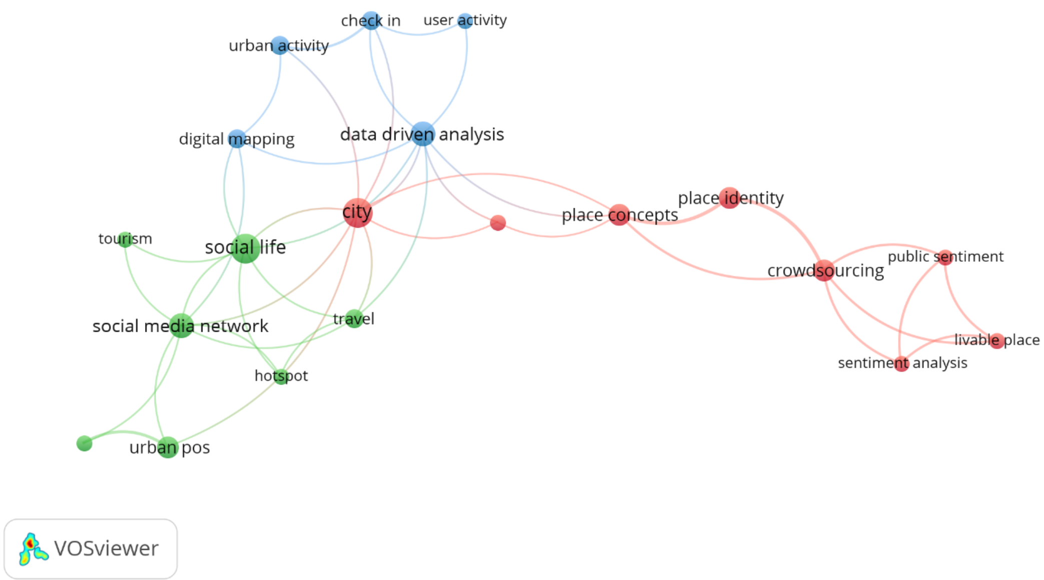


Figure 6. Network visualization of the titles.

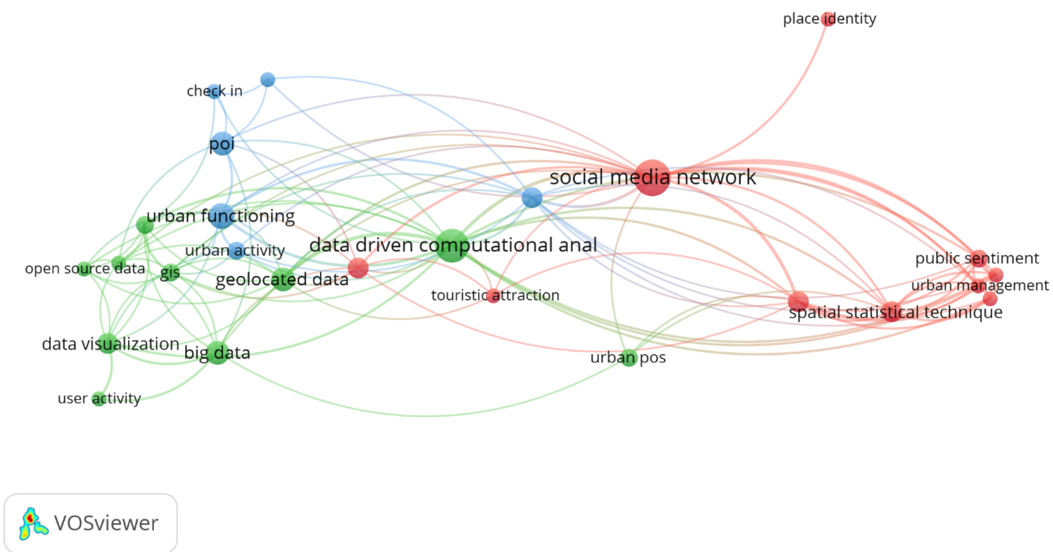


Figure 7. Network visualization of the keywords.

selected articles represent the state of the art research in the big data field. The network results also reveal that the majority of the references are connected to each other in terms of the research topic and applied method.

DISCUSSION

There is a growing trend in using big data to understand urban phenomena and support decision-making. However, urban researchers need to consider the discussions surrounding the drawbacks of big data and data-driven analysis in this data-rich environment. Data-driven analysis of urban systems has faced criticism for oversimplifying the intricate relationship patterns inherent in city systems, drawing parallels to Alexander's "city as tree" concept from the modernist movement (Mattern, 2021). According to Mattern (2021), much like the reduction of complex urban systems to tree-like structures in modernist approaches, data-driven analysis risks simplifying cities into similar structures through data networks, potentially neglecting the inherent complexity and interconnectivity of urban environments. Moreover, other criticisms of data-driven urban analysis are derived from the lack of methodological and theoretical background (Berry, 2011), the absence of sociological, economic, and psychological aspects (Thakuriah et al., 2017), and the missing epistemological links between data-driven modeling and critical urban theory (Thakuriah et al., 2017). In response to these counterarguments, Kitchin (2017) states that data-driven methods change the way

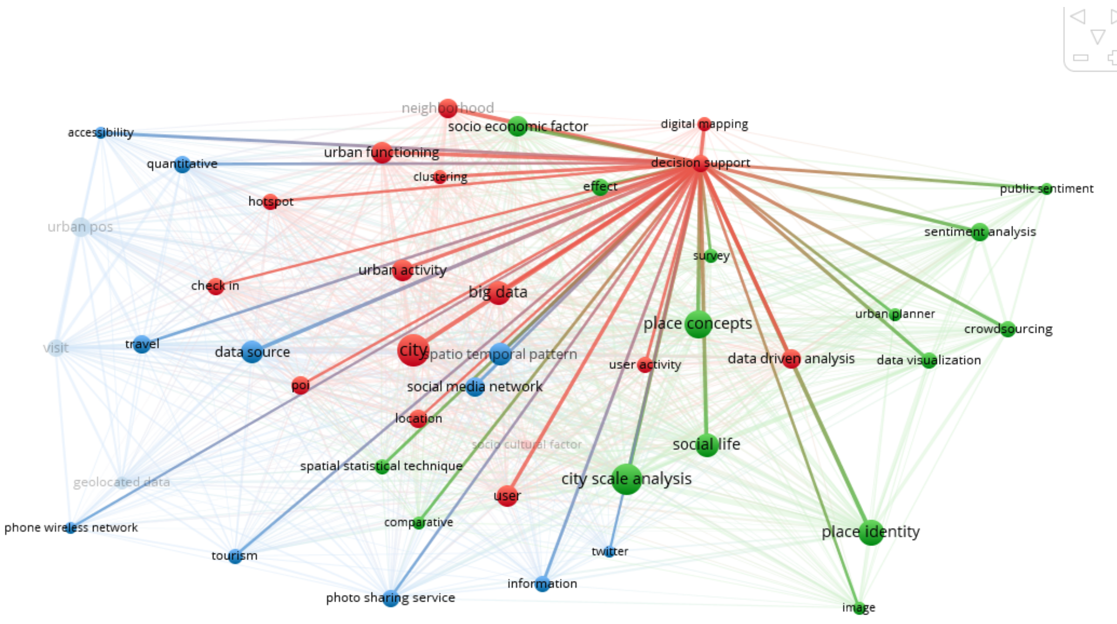


Figure 8. The linkage of the urban decision support with other terms in the abstract.

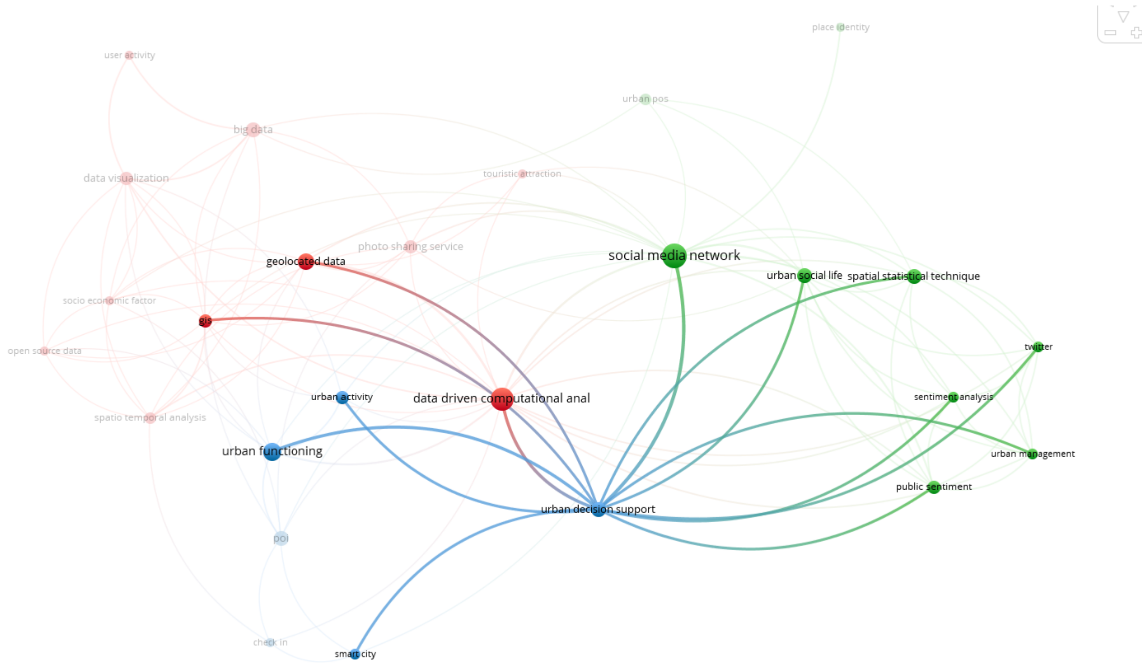
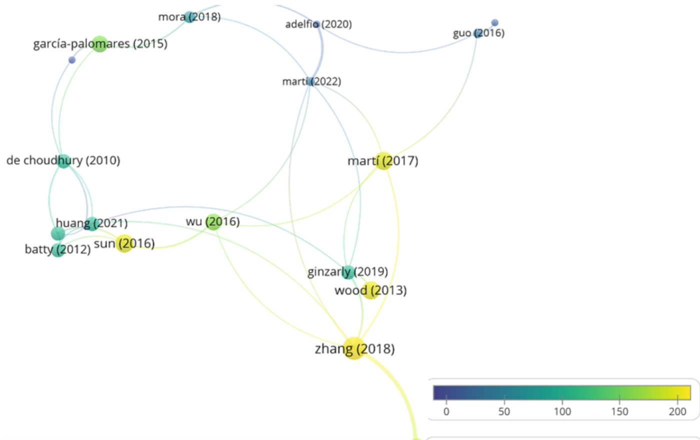


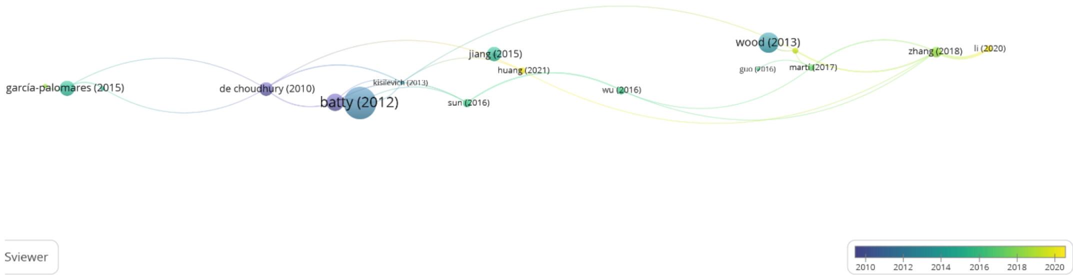
Figure 9. The linkage of the urban decision support with other terms in keywords.

hypotheses are created, shifting from theory to data. Data is processed through algorithms, rather than theories, to discover urban phenomena (Kitchin, 2017).

The drawbacks of big data stem from issues related to representation, ethics, technique, and methodology. In terms of representativeness, ongoing discussions include the adequacy of LBSN sampling in representing collective user patterns, unclear and vague attributes of the sampling, and comparisons with surveys in terms of target coverage. Additionally, there is skepticism about the context, content, ownership of information (Chan, 2015), and blurred lines between commercial and academic research fields (Boyd & Crawford, 2012). Technical drawbacks can be observed in data generation, capturing, processing, and management stages due to the fragmented structure of API raw data (Thakuriah et al., 2017). Different LBSN platforms also have specific technical



a



b

Figure 10. Network visualization of the references based on the citations **(Figure 10.a)** and years **(Figure 10.b)**.

drawbacks, such as complexity in requesting and retrieving data, differences in categories and data types, distinct methods used for validation, selection, and interpretation of data, inconsistencies in the retrieval data scale, and user-generated content errors (such as duplications or inaccurate place information, and locative descriptors) (Marti et al., 2019).

The bibliometric mapping network based on the visualization of similarities has the potential to represent the exploration of state-of-the-art research. However, the bibliometric mapping network has challenges in the scope of the study. These challenges are derived from the oversensitivity of the relationship network to minor changes, the existence of multiple correspondence words to describe the same term, and limitations of the tool. This network is very sensitive to minor changes. The network structure changes completely even with minor changes, making it difficult to read the associations and clusters in the network structure each time. The fragile structure creates a limitation. In addition, there are many different definitions to describe the same term. For instance, more than 15 different terms have been used to describe geolocated data. Having so many definitions describing the same term complicates the filtering process. Another difficulty arises from the use of the tool. The fact that the tool only analyzes the text in the abstract prevents the effective analysis of the terms in the articles. In this study, the researcher generates a dictionary to filter the close-meaning terms and creates an appropriate group for them. The filtering process through the dictionary is conducted manually by the researcher. This process is labor and time-intensive and prone to human error. The labor intensiveness of filtering is another limitation of this study.

To overcome these limitations, the filtering process can be furthered with machine learning to provide efficiency. The use of machine learning algorithms in semantics can be employed to construct an automated dictionary in further studies. Further studies could involve reviewing a wider range of literature and exploring different perspectives and approaches to conduct comprehensive research. Additionally, this study can be supported by the collaboration of experts in the field.

CONCLUSION

This paper examines state-of-the-art studies in urban informatics that exploit big data to support urban decisions. In this paper, a qualitative analysis has been conducted to classify the references based on the interrelationships of the terms. The qualitative map results give insight into the association of keywords and the state-of-the-art research in big data, while the quantitative diagram results show the numeric distribution of the keywords in this field. The use of VOSviewer software for the bibliometric analysis of selected studies and the creation of a dictionary to filter close-meaning terms are unique aspects of the study. This study is useful for revealing hot research topics and data analysis methods in the scope of big data in the urban informatics domain. The identification of hot research topics and data analysis methods can help inform the development of decision support systems that use data-driven computational methods, spatial statistical methods, and mapping spatiotemporal patterns of urban phenomena.

In this study, the VOSviewer tool has been used for the literature analysis, which constructs bibliometric networks based on the visualization of similarities (VOS). In this study, VOSviewer contributes to the management, representation, and customization of the relationships and versatility of the analysis. This tool provides the ability to handle large datasets in an efficient way. It represents the complex data network in an easy-to-understand manner. It allows users to customize the visualization of specific relationships or nodes to emphasize the networks. In this study, the researcher customizes the decision support relations with other terms. Moreover, VOSviewer offers to analyze and visualize various types of bibliometric relationships. This study conducts different bibliometric analyses based on the keyword co-occurrence and term co-occurrence in the abstract and titles and citations. These different analysis methods add value to the versatility of this tool.

In this study, the selected reference studies were analyzed with a systematic literature review through the VOSviewer software based on the terms occurring in the title, keywords, and abstracts. In the literature, a systematic literature review was applied by searching the keywords with induction. Unlike the reference studies, the process of the literature study review is based on deduction: from the references to the terms. This literature study aims to map the general trends in big data and urban informatics regarding their intersection with decision support in urban planning and design. It provides an opportunity for further research on the integration of decision support systems in big data studies. In the selected studies, decision support is mostly provided through data-driven computational methods (machine learning and data analytics), spatial statistical methods for clustering and classifications, and mapping the spatiotemporal pattern of urban phenomena. The reference studies support decisions mainly related to urban activities and functioning, user activities (movement), visiting, and urban perception.

Acknowledgements: I would like to acknowledge to Prof. Dr. Gülen Çağdaş for her contributions and support throughout the process.
Ethics Committee Approval: The author declared that this study does not require ethics committee approval.
Peer Review: Externally peer-reviewed.
Grant Support: The author declared that this study have received no financial support.
Note: The assistant guest editor was not involved in the evaluation, peer-review and decision processes of the article. These processes were carried out by the Editor-in-Chief and the member editors of the editorial management board.

REFERENCES

Abbasi, A., Rashidi, T. H., Maghrebi, M., & Waller, S. T. (2015). Utilising Location Based Social Media in Travel Survey Methods: Bringing Twitter Data into the Play. *Proceedings of the 8th ACM SIGSPATIAL International Workshop on Location-Based Social Networks*.https://doi.org/10.1145/2830657.2830660

Adelfio, M., Serrano-Estrada, L., Martí-Ciriquián, P., Kain, J.-H., & Stenberg, J. (2020). Social Activity in Gothenburg's Intermediate City: Mapping Third Places through Social Media Data. *Applied Spatial Analysis and Policy, 13*(4), 985–1017. https://doi.org/10.1007/s12061-020-09338-3

Balaban, O., & Tuncer, B. (2016). Visualizing Urban Sports Movement. *Proceedings of the International Conference on Education and Research in Computer Aided Architectural Design in Europe, 2*, 89–94. https://doi.org/10.52842/conf.ecaade.2016.2.089

Batty (2001). Foreword. In R. Laurini (ed.), *Information systems for urban planning: a hypermedia cooperavite approach.* Taylor and Francis

Batty M. (2016). Big Data and The City. *Built Environment 42 (3)*, 321-337.https://doi.org/10.2148/benv.42.3.321

Batty M., Axhausen K., Fosca G., Pozdnoukov A., Bazzani A., Wachowicz M., Ouzonunis G., Portugali Y. (2012). Smart Cities of the Future. *Eur. Phys. J. Special Topics 214*, pp. 481-518. https://doi.org/10.1140/epjst/e2012-01703-3

Batty, M. (2013a). Urban Informatics and Big Data. London: CASA, University College London. Retrieved date: 03/05/2019. Retrieved from:http://www.spatialcomplexity.info/files/2015/07/Urban-Informatics-and-Big-Data

Batty, M. (2013b). The New Science of Cities. Cambridge, MA: MIT Press.

Berry, D. (2011). The computational turn: thinking about the digital humanities. *Culture Machine*, 12. https://sro.sussex.ac.uk/id/eprint/49813/

Boyd, D., Crawford, K. (2012). Critical Questions about big data: Provocations for a cultural,technological, and scholarly phenomenon. *Information Communication and Society, 15* (5), 662-679. https://doi.org/10.1080/1369118X.2012.678878

Çağdaş G., Bacınoğlu Z. S., Çavuşoğlu Ö., (2015). Mimarlıkta Hesaplamalı Yaklaşımlar. Mimaride Sayısal Fırsatlar, 34-42.

Chan, A. (2015). Big data interfaces and the problem of inclusion. Media, Culture & Society, 37(7), 1078-1083. https://doi.org/10.1177/0163443715594106

Cranshaw, J., Schwartz, R., Hong, J., & Sadeh, N. (2021). The Livehoods Project: Utilizing Social Media to Understand the Dynamics of a City. *Proceedings of the International AAAI Conference on Web and Social Media, 6*(1 SE-Full Papers), 58–65. https://doi.org/10.1609/icwsm.v6i1.14278

De Choudhury, M., Feldman, M., Amer-Yahia, S., Golbandi, N., Lempel, R., & Yu, C. (2010). Automatic Construction of

Travel Itineraries Using Social Breadcrumbs. *Proceedings of the 21st ACM Conference on Hypertext and Hypermedia*, 35–44. https://doi.org/10.1145/1810617.1810626

Deng, X., & Newsam, S. D. (2017). Quantitative Comparison of Open-Source Data for Fine-Grain Mapping of Land Use. *CoRR, abs/1711.0*. http://arxiv.org/abs/1711.03641

Drudzel, M.J. & Fylnn R.R. (2002). Decision support systems. In A. Kent (ed.), *Encyclopedia of library and information science* (2 nd ed., pp. 3-12). Marcel Dekker Inc. https://sites.pitt.edu/ druzdzel/psfiles/dss.pdf

Ensari, E., & Kobaş, B. (2018). Web scraping and mapping urban data to support urban design decisions. *A/Z ITU Journal of the Faculty of Architecture, 15*(1), 5–21. https://doi.org/10.5505/itujfa.2018.40360

Fredriksson, C. (2018). Big data creating new knowledge as support in decision-making: practical examples of big data use and consequences of using big data as decision support. Journal of Decision System, 27(3), 1-18. http://dx.doi.org/10.1080/12460125.2018.1459068

García-Palomares, J. C., Gutiérrez, J., & Mínguez, C. (2015). Identification of tourist hot spots based on social networks: A comparative analysis of European metropolises using photo-sharing services and GIS. *Applied Geography, 63*, 408–417. https://doi.org/https://doi.org/10.1016/j.apgeog.2015.08.002

Ginzarly, M., Pereira Roders, A., & Teller, J. (2019). Mapping historic urban landscape values through social media. *Journal of Cultural Heritage, 36*. https://doi.org/10.1016/j.culher.2018.10.002

Girardin, F., Calabrese, F., Fiore, F. D., Ratti, C., & Blat, J. (2008). Digital Footprinting: Uncovering Tourists with User-Generated Content. *IEEE Pervasive Computing*, 7(4), 36–43. https://doi.org/10.1109/MPRV.2008.71

Goodchild, M. F. (2007). Citizens as Sensors: The World of Volunteered Geography. *GeoJournal (69)*, 211-221. DOI: 10.1007/s10708-007-9111-y.

Gordon, E., & E Silva, A. D. S. (2011). *Net locality: Why location matters in a networked world*. John Wiley & Sons.

Guo, W., Gupta, N., Pogrebna, G., & Jarvis, S. (2016). Understanding happiness in cities using Twitter:Jobs, children, and transport. 2016 *IEEE International Smart Cities Conference (ISC2)*, 1–7. https://doi.org/10.1109/ISC2.2016.7580790

Huang, J., Obracht-Prondzynska, H., Kamrowska-Zaluska, D., Sun, Y., & Li, L. (2021). The image of the City on social media: A comparative study using “Big Data” and “Small Data” methods in the Tri-City Region in Poland. *Landscape and Urban Planning, 206*, 103977. https://doi.org/https://doi.org/10.1016/j.landurbplan.2020.103977

Jang, K. M., & Kim, Y. (2019). Crowd-sourced cognitive mapping: A new way of displaying people’s cognitive perception of urban space. *PLOS ONE, 14*(6), e0218590. https://doi.org/10.1371/journal.pone.0218590

Jiang, S., Alves, A., Rodrigues, F., Ferreira, J., & Pereira, F. C. (2015). Mining point-of-interest data from social networks for urban land use classification and disaggregation. Computers, Environment and *Urban Systems, 53*, 36–46. https://doi.org/https://doi.org/10.1016/j.compenvurbsys.2014.12.001

Kaisler, S., Armour, F., Espinosa, J.A., Money, W. (2013). Big data redux: Issues and challenges moving forward. 46 th Hawaii International Conference on System Services. http://dx.doi.org/10.24251/HICSS.2019.131

Kisilevich, S., Keim, D., Andrienko, N., & Andrienko, G. (2013). *Towards Acquisition of Semantics of Places and Events by Multi-perspective Analysis of Geotagged Photo Collections BT - Geospatial Visualisation* (A. Moore & I. Drecki (eds.); pp. 211–233). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642122897_10

Kitchin, R. (2017, February 1). Urban Science: A Short Primer. https://doi.org/10.31235/osf.io/sdsu2

Laurini, R. (2001). *Information systems for urban planning: a hypermedia cooperavite approach*. Taylor and Francis.

Li, F., Li, F., Li, S., & Long, Y. (2020). Deciphering the recreational use of urban parks: Experiments using multi-source big data for all Chinese cities. *Science of The Total Environment, 701*, 134896. https://doi.org/https://doi.org/10.1016/j.scitotenv.2019.134896

Lin, S., Xie, R., Xie, Q., Zhao, H., & Chen, Y. (2017). Understanding User Activity Patterns of the Swarm App: A Data-Driven Study. *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*, 125–128. https://doi.org/10.1145/3123024.3123086

Loo B. P. Y. & Tang W. S. M. (2019). “Mapping” Smart Cities. *Journal of Urban Technology, 26*(2), 129-146. https://doi.org/10.1080/10630732.2019.1576467

Marti, P., Serrano-Estrada L., Nolasco-Criugeda A. (2019). Social media data: Challenges, opportunities and limitations in urban studies. Computers, Environment and Urban Systems, 74 (2019), 161- 174. https://doi.org/10.1016/j.compenvurbsys.2018.11.001

Martí, P., Serrano-Estrada, L., & Nolasco-Cirugeda, A. (2017). Using locative social media and urban cartographies to identify and locate successful urban plazas. *Cities, 64*, 66–78. https://doi.org/https://doi.org/10.1016/j.cities.2017.02.007

Mattern, S. (2021). *A city is not a computer: Other urban intelligence*. Princeton University Press.

Mora, H., Pérez-delHoyo, R., Paredes-Pérez, J. F., & Mollá-Sirvent, R. A. (2018). Analysis of Social Networking Service Data for Smart Urban Planning. In *Sustainability* (Vol. 10, Issue 12). https://doi.org/10.3390/su10124732

Schlieder, C., & Matyas, C. (2009). Photographing a City: An Analysis of Place Concepts Based on Spatial Choices. *Spatial Cognition & Computation, 9*(3), 212–228. https://doi.org/10.1080/13875860903121848

Sprague, R.H., Carlson, E.D. (1982). *Building effective decision support systems*, Prentice-Hall.

Sun, Y., Fan, H., Li, M., & Zipf, A. (2015). Identifying the city center using human travel flows generated from location-based social networking data. *Environment and Planning B: Planning and Design, 43*, 480–498. https://doi.org/10.1177/0265813515617642

Thakuriah P., Tilahun N.Y., Zellner, M. (2017a). Big data and urban informatics: Innovations and challenges to urban planning and knowledge discovery. In P. Thakuriah, N. Tilahun, M. Zellner (eds.), *Seeing Cities Through Big Data* (pp. 11-48). Springer. https://doi.org/10.1007/978-3-319-40902-3

Townsend, M.A. (2013). *Smart Cities Big data, civic hackers, and the quest for a new utopia*. London: W.W. Norton & Company.

Tunçer, B. and You, L. (2017). *Informed Design Platform*. Proceedings of the 35th International Conference on Education and Research in Computer Aided Architecture and Design in Europe (eCAADe) (pp. 545-552). Rome, Italy.

Uzgören, G. (2018). Airbnb impact on gentrification process: the case of Rasimpaşa Neighborhood Kadıköy. *Journal of Planning, 28*(2), 154–170. https://doi.org/10.14744/planlama.2018.29491

Van Dicyk A.G.M. (2006). The digital divide achievements and shortcomings. *Poetic 34*, 221-235. https://doi.org/10.1016/j.poetic.2006.05.004

Van Eck, N.J., & Waltman, L. (2014). Visualizing bibliometric networks. In Y. Ding, R. Rousseau, & D. Wolfram (Eds.), Measuring scholarly impact: Methods and practice (pp. 285–320). Springer. http://dx.doi.org/10.1007/978-3-319-10377-8_13

Van Eck, N.J., Waltman, L. (2013). VOSviewer Manual. https://www.vosviewer.com/documentation/Manual_VOSviewer_1.6.15.pdf

VOSviewer, (2023). Welcome to VOSviewer. https://www.vosviewer.com/

Wilken, R., & Goggin, G. (2017). *Locative media*. Routledge.

Wood, S. A., Guerry, A. D., Silver, J. M., & Lacayo, M. (2013). Using social media to quantify naturebased tourism and recreation. *Scientific Reports, 3*. https://doi.org/10.1038/srep02976

Wu C, Ye X, Ren F, Wan Y, Ning P, Du Q (2016) Spatial and Social Media Data Analytics of Housing Prices in Shenzhen, China. PLoS ONE 11(10): e0164553. https://doi.org/10.1371/journal.pone.0164553

Wu, C., Ye, X., Ren, F., Wan, Y., Ning, P., & Du, Q. (2016). Spatial and Social Media Data Analytics of Housing Prices in Shenzhen, China. *PLOS ONE, 11*(10), e0164553. https://doi.org/10.1371/journal.pone.0164553

You, L., & Tunçer, B. (2016). Exploring public sentiments for livable places based on a crowd-calibrated sentiment analysis mechanism. *2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, 693–700. https://doi.org/10.1109/ASONAM.2016.7752312

Zhang, S., & Zhou, W. (2018). Recreational visits to urban parks and factors affecting park visits: Evidence from geotagged social media data. *Landscape and Urban Planning, 180*, 27–35. https://doi.org/https://doi.org/10.1016/j.landurbplan.2018.08.004

Zhou, X., & Zhang, L. (2016). Crowdsourcing functions of the living city from Twitter and Foursquare data. *Cartography and Geographic Information Science, 43*(5), 393–404. https://doi.org/10.1080/15230406.2015.1128852

Url 1: https://tinyurl.com/2pfbrhq9

How cite this article

Kirdar, G. (2023). A systematic literature review of big data in urban studies. *Journal of Technology in Architecture Design and Planning*. Advanced online publication. https://doi.org/10.26650/JTADP.01.004

Appendix A: Export results of the analysis in VOSviewer.

Table 6. Terms and occurrences from the abstracts.

Id	Term	Occurrences	Relevance Score
5	city	73	0.2196
6	city scale analysis	65	0.1696
31	research	60	0.0868
26	place concept	55	0.1679
34	social media network	40	0.1393
3	big data	35	0.037
27	place identity	34	0.923
46	urban pos	34	23.111
10	data source	32	0.4046
38	Spatio-temporal analysis	30	0.1099
44	urban functioning	30	0.7392
49	user	29	0.0752
42	urban activity	27	0.1801
21	neighborhood	24	0.2908
51	visit	24	2.582
2	analysis	22	0.4394
28	poi	22	0.2724

Id	Term	Occurrences	Relevance Score
4	Check-in	21	0.4094
20	location	21	0.2213
32	sentiment analysis	21	27.737
41	travel	20	12.811
25	photo sharing service	19	0.2165
30	quantitative	19	17.818
43	urban decision support	19	0.0867
50	user activity	19	0.1875
14	effect	18	0.1261
16	hotspot	18	0.2795
9	crowdsourcing	16	1.796
11	data visualization	16	0.8846
15	geolocated data	15	12.933
19	information	15	0.123
23	opportunity	15	0.4826
36	Socio-economic factor	15	0.5254
40	touristic attraction	15	0.4892

Table 7. The terms and occurrences of the keywords.

Id	Keyword	Occurrences	Total link strength
4	big data	5	8
14	ata driven computational analysis method	10	28
16	data visualization	4	9
19	geolocated data	5	13
20	GIS	3	11
36	photo sharing service	4	9
39	POI	5	6
64	urban social life	4	13

Id	Keyword	Occurrences	Total link strength
41	public sentiment	3	10
45	social media network	12	26
50	patial statistical technique	4	10
51	spatio temporal analysis	3	8
58	urban activity	3	7
60	urban decision support	4	13
61	urban functioning	6	13
63	urban pos	3	4

Table 8. The citation numbers of the references and journals.

Id	Document	Citations	Links
3	Batty (2012)	1038	0
8	Wood (2013)	410	2
27	Girardin (2008)	303	3
9	García-Palomares (2015)	229	2
22	Jiang (2015)	201	0
21	De Choudhury (2010)	182	2
7	Zhang (2018)	124	2
14	Sun (2016)	73	1
1	Wu (2016)	60	1

Id	Document	Citations	Links
6	Martí (2017)	50	1
13	Li (2020)	49	2
25	Huang (2021)	46	0
12	Ginzarly (2019)	43	1
11	Zhou (2016)	37	0
10	Abbasi (2015)	27	2
5	Schlieder (2009)	26	0
17	Mora (2018)	25	1
15	Agryzkov (2017)	21	0

Table 9. The journals citation of the reference articles.

id	source	citations	total link strength		id	source	citations	total link strength
26	The European Physical Journal Special Topics	1042	0		20	Proceedings of The 8th Acm Sigspatial International Workshop On LBSN	27	2
23	Scientific Reports	411	2		24	Spatial Cognition &Amp; Computation	26	0
11	Ieee Pervasive Computing	303	3		25	Sustainability	25	1
4	Applied Geography	229	4		12	International Journal of Geographical Information Science	21	1
8	Computers, Environment and Urban Systems	201	0		10	Geospatial Visualisation	17	1
19	Proceedings of the 21st ACM Conference on Hypertext and Hypermedia	182	2		1	2016 IEEE International Smart Cities Conference (ISC2)	15	0
16	Landscape and Urban Planning	171	2		21	Proceedings of The International AAAI Conference on Web and Social Media	14	0
9	Environment and Planning B: Planning and Design	73	1		18	The 2017 ACM International Symposium on Wearable Computers	9	0
17	Plos One	60	1		2	2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (Asonam)	8	0
7	Cities	50	3		5	Applied Spatial Analysis and Policy	6	2
22	Science of the Total Environment	50	2		3	A/Z : ITU Journal of Faculty of Architecture	3	0
13	Journal of Cultural Heritage	44	1		15	Journal of Urban Technology	3	2
6	Cartography and Geographic Information Science	37	0		14	Journal of Planning	0	0

Appendix B: The dictionary to group the related terms.

Label	Replace by		Label	Replace by		Label	Replace by
access	accessibility		Bundang	city scale analysis		amt	data driven computational analysis method
accessible green space	accessibility		China	city scale analysis		computational method	data driven computational analysis method
park accessibility	accessibility		Country	city scale analysis		machine learning	data driven computational analysis method
classification accuracy	accuracy		Dongtan	city scale analysis		calibration	data driven computational analysis method
best accuracy	accuracy		Ilsan	city scale analysis		reality mining	data driven computational analysis method
good agreement	agreement		Korea	city scale analysis		mechanical turk	data driven computational analysis method
new analysis capability	analysis		Rome	city scale analysis		orienteering problem	data driven computational analysis method
detailed analysis	analysis		San Francisco	city scale analysis		page rank algorithms	data driven computational analysis method
social app	app		Seoul	city scale analysis		city computing	data driven computational analysis method
new approach	approach		Songdo	city scale analysis		geographic recommender systems	data driven computational analysis method
interdisciplinary approach	approach		Alicante	city scale analysis		information extraction	data driven computational analysis method
twofold approach	approach		Amsterdam	city scale analysis		web scraping	data driven computational analysis method
available data	available		Athens	city scale analysis		topic modelling	data driven computational analysis method
data	big data		Bamberg	city scale analysis		semantics enrichment	data driven computational analysis method
check	check in		Barcelona	city scale analysis		machine learning algorithms	data driven computational analysis method
park check	check in		Beijing	city scale analysis		new source	data source
Foursquare data	check in		Benchmark	city scale analysis		new type	data source
Weibo	check in		Berlin	city scale analysis		additional source	data source
social media network Foursquare	check in		Cardiff	city scale analysis		additional type	data source
social network Foursquare	check in		Case area	city scale analysis		commercial resource	data source
swarm	check in		Dublin	city scale analysis		latent source	data source
check-ins	check in		Gothenburg	city scale analysis		utilizing resource	data source
check-in	check in		Gothenburg's intermediate city	city scale analysis		variety	data source
Foursquare	check in		Hong Kong	city scale analysis		diversity	data source
swarm app	check in		Istanbul	city scale analysis		variety	data source
large city	city		Italy	city scale analysis		source	data source
major city	city		London	city scale analysis		type	data source
major European city	city		Madrid	city scale analysis		databases	data source
New York City	city		Paris	city scale analysis		reliable proxy	data visualization
cities and towns	city		Pittsburgh	city scale analysis		graphical image	data visualization
classifier increase	classification		Rotterdam	city scale analysis		new visualization method	data visualization
category	classification		Sweden	city scale analysis		visual representation	data visualization
classified model	classification		tri city region	city scale analysis		graph	data visualization
cluster	clustering		Jurong East	city scale analysis		visual programming software	data visualization
clustering analysis	clustering		Singapore	city scale analysis		visualization tool	data visualization
clustering model	clustering		testbed area	city scale analysis		interactive visualization service	data visualization
functional cluster	clustering		case study city	city scale analysis		visualization mode	data visualization
functional clusters	clustering		similar conceptualization	description		representation	data visualization
comparative analysis	comparative		ident	description		graphic	data visualization
comparative study	comparative		partial depiction	description		information visualization	data visualization
distinction	comparative		influence	effect		urban visualization	data visualization
individual difference	comparative		lasting influence	effect		geographic information systems	GIS
difference	comparative		impact	effect		geographic information data source open street map	GIS
comparison	comparative		value	effect		official GIS database	GIS
crowd sourcing marketplace	crowdsourcing		practical implication	effect		GIS	GIS
crowdsourced calibration service	crowdsourcing		show great promise	effect		system	GIS
map	digital mapping		urban center	hotspot		open source	open source data
mapping	digital mapping		city center identification	hotspot		open source data	open source data
mapping third place	digital mapping		higher spatial concentration	hotspot		osm data	open source data
mapping urban data	digital mapping		intermediate city area	hotspot		web server	open source data
maps outdoor physical activity	digital mapping		main hot spot	hotspot		web site	open source data
resulting map	digital mapping		neighborhoods center	hotspot		web sites	open source data
map	digital mapping		spot	hotspot		weblog	open source data
fine grain land use mapping	digital mapping		tourist hot spot	hotspot		Airbnb	open source data
fine grain mapping	digital mapping		center identification	hotspot		new opportunity	opportunity
digital map	digital mapping		city center	hotspot		great opportunity	opportunity
digital recording	digital mapping		new technology	ICT		possibility	opportunity
latitude longitude	geolocated data		communication	ICT		advantage	opportunity
analyzing people's geo	geolocated data		technology	ICT		cell phone network data	phone wireless network
available geotagged check	geolocated data		leverage voluntary local knowledge	knowledge		social network	phone wireless network
footprint	geolocated data		potential bias	limitation		mobile phone network	phone wireless network
geo temporal breadcrumb	geolocated data		limited accuracy	limitation		mobile phone user	phone wireless network
geolocated information	geolocated data		limited agreement	limitation		Telecom Italia mobile	phone wireless network

geolocated social media data	geolocated data		geo	location		wireless network event	phone wireless network
georeferenced photo	geolocated data		accurate location	location		mobile phone	phone wireless network
geotagged photograph	geolocated data		reliable measure	measure		wireless network	phone wireless network
geotagged social media data	geolocated data		novel method	method		photograph	photo sharing service
location information	geolocated data		methodological technique	method		photo stream	photo sharing service
locational log	geolocated data		new way	method		photographer	photo sharing service
volunteered geographic information	geolocated data		conventional method	method		photography	photo sharing service
locational data	geolocated data		new method	method		residents photograph	photo sharing service
digital footprinting	geolocated data		overall methodology	method		Instagrammability	photo sharing service
pervasive user-generated content	geolocated data		methodology	method		Instagram	photo sharing service
user-related spatio temporal data	geolocated data		way	method		Flickr	photo sharing service
geo-tags	geolocated data		administrative boundary	neighborhood		Flickr data	photo sharing service
geospatial semantics	geolocated data		administrative neighborhoods	neighborhood		Instagram	photo sharing service
location-based services	geolocated data		administrative subdivision	neighborhood		Instagramability	photo sharing service
location-based social networks	geolocated data		city partition	neighborhood		Panoramio	photo sharing service
social tagging	geolocated data		city s administrative neighborhoods	neighborhood		web site Flickr	photo sharing service
geospatial data	geolocated data		intermediate city neighborhood	neighborhood		photo-sharing services	photo sharing service
location-based social media	geolocated data		Kadiköy municipal boundary	neighborhood		geotagged photo collections	photo sharing service
volunteered geographic information (vgi)	geolocated data		neighborhood boundaries	neighborhood		urban location	place concept
volunteered geographic information	geolocated data		neighborhood boundary	neighborhood		city place	place concept
volunteered geographic information	geolocated data		polynuclear neighborhood structure	neighborhood		place concepts	place concept
data sources Google Places	Google Places		precise boundary	neighborhood		site	place concept
Google Places data	Google Places		boundary	neighborhood		affluent area	place concept
Google Places social networks	Google Places		neighborhood	neighborhood		area	place concept
overall image	image		site observation	observation		preferred urban space	place concept
single image	image		site observation	observation		characterized area	place concept
image collection	image		direct observation	observation		third	place concept
prior information	information		Bing maps	open source data		third place	place concept

skilled interaction	interaction		online real estate listing	open source data		third places	place concept
urban China	place concept		city project	research		higher ethnic diversity	socio cultural factor
environment	place concept		study	research		worker	socio cultural factor
urban environment	place concept		paper	research		age group	socio cultural factor
urban space	place concept		article	research		digital age	socio cultural factor
place	place concept		work	research		gender	socio cultural factor
crowdsourced cognitive map	place identity		further research	research		cultural ecosystem services	socio cultural factor
public image	place identity		larger research project	research		housing	socio economic factor
city image	place identity		scientist	researcher		socio economic statistic	socio economic factor
cognitive mapping	place identity		large scale	scale		socio economic statistic	socio economic factor
collective identity	place identity		study scale	scale		city s business distribution	socio economic factor
collective place identity	place identity		3d sentiment map	sentiment analysis		economic trend	socio economic factor
conventional cognitive mapping method	place identity		geo sentiment analysis mechanism	sentiment analysis		livehoods project	socio economic factor
identity	place identity		geo sentiment analysis service	sentiment analysis		lower income level	socio economic factor
imageable city	place identity		2d sentiment dashboard	sentiment analysis		neighborhoods economic activity	socio economic factor
node	place identity		sentiment analysis process	sentiment analysis		e commerce	socio economic factor
people's cognitive perception	place identity		sentiment analysis mechanism	sentiment analysis		economy	socio economic factor
urban identity	place identity		sentiment analysis service	sentiment analysis		economic activity	socio economic factor
urban identity a crowd	place identity		social sentiment analysis	sentiment analysis		costs	socio economic factor
edge	place identity		social sentiment analysis engine	sentiment analysis		housing prices	socio economic factor
architectural landmark	place identity		social sentiment analysis feature	sentiment analysis		art research technique	spatial statistical technique
imageability	place identity		training sentiment classifier	sentiment analysis		kernel density estimation	spatial statistical technique
imagery	place identity		smart cities	smart city		multiple linear regression	spatial statistical technique
Kevin Lynch	place identity		social media	social media network		possible correlation	spatial statistical technique
public image	place identity		online social media website	social media network		strong spatial relationship	spatial statistical technique
points of interest	poi		social medium	social media network		weighted regression	spatial statistical technique
pois	poi		social media text data	social media network		statistic	spatial statistical technique
interests	poi		locative social media network	social media network		technique	spatial statistical technique
main tourist attraction	poi		locative social medium	social media network		pattern classification	spatial statistical technique
major attraction	poi		popular social media platform	social media network		dbscan	spatial statistical technique

poi graph	poi		smd	social media network		Grivan-Newman	spatial statistical technique
points	poi		social network data	social media network		local getis-ord	spatial statistical technique
tourist attraction	poi		social networking data	social media network		regression models	spatial statistical technique
interest	poi		text	social media network		greater dispersion	spatio-temporal analysis
attraction	poi		text message	social media network		coarse class granularity	spatio-temporal analysis
question	problem		utilizing social media	social media network		rich spatial temporal information	spatio-temporal analysis
issue	problem		geotagged social network data	social media network		peoples spatiotemporal activity pattern	spatio-temporal analysis
laborious task	problem		geotagged social network message	social media network		spatial choice	spatio-temporal analysis
research question	problem		social media data	social media network		spatial definition	spatio-temporal analysis
public sentiments	public sentiment		social networks	social media network		spatial heterogeneity	spatio-temporal analysis
aggregate record	quantitative		social media	social media network		spatial pattern	spatio-temporal analysis
increased number	quantitative		social network analysis	social media network		Spatio-temporal city characteristic	spatio-temporal analysis
larger number	quantitative		social media analytics	social media network		Spatio-temporal phenomena	spatio-temporal analysis
quantitative ing park use	quantitative		diverse socio cultural factor	socio cultural factor		spatio-temporal analysis	spatio-temporal analysis
quantitative comparison	quantitative		higher ethnic diversity	socio cultural factor		spatial analysis	spatio-temporal analysis
quantitative understanding	quantitative		socio cultural factor	socio cultural factor		day	spatio-temporal analysis
quantitative	quantitative		traditional social character	socio cultural factor		year	spatio-temporal analysis
number	quantitative		diverse socio cultural factor	socio cultural factor		date data	spatio-temporal analysis
specifying time	spatio-temporal analysis		activity intensity	urban activity		public administration	urban management
strong spatial concentration	spatio-temporal analysis		activity pattern	urban activity		designers planners	urban management
concentration	spatio-temporal analysis		city exhibits urban activity	urban activity		city planner	urban management
temporal coverage	spatio-temporal analysis		intermediate city s urban activity area	urban activity		relevant plaza	urban pos
time	spatio-temporal analysis		urban activity	urban activity		successful plaza	urban pos
time series analysis	spatio-temporal analysis		urban function	urban activity		successful public space	urban pos
timely way	spatio-temporal analysis		urban economic activities	urban activity		successful urban plaza	urban pos
time	spatio-temporal analysis		urban activities	urban activity		square	urban pos
extensive survey	survey		appropriate decision making	urban decision support		recreational site	urban pos
questionnaire response	survey		insight	urban decision support		recreational area	urban pos
surveys interview	survey		ground truth	urban decision support		leisure	urban pos

visitor survey	survey		provision	urban decision support		urban park	urban pos
interview	survey		strong support	urban decision support		cultural relics park	urban pos
subject	topic		urban decision making	urban decision support		large urban park	urban pos
historic city center	touristic attraction		urban design decision	urban decision support		national park	urban pos
cultural heritage	touristic attraction		politic	urban decision support		neighborhood park	urban pos
historic urban landscape	touristic attraction		decision	urban decision support		park context	urban pos
popular bus tour	touristic attraction		decision making	urban decision support		park location	urban pos
tourist	touristic attraction		policy	urban decision support		park size	urban pos
tourism geography	touristic attraction		evidence	urban decision support		park typology	urban pos
uncovering tourist	touristic attraction		urban design decision support	urban decision support		park user	urban pos
touristic attractions	touristic attraction		urban planning	urban decision support		park usage	urban pos
trail	track		dynamic urban area	urban dynamic		plaza	urban pos
destination constraint	travel		dynamics	urban dynamic		public space	urban pos
bus stop	travel		human dynamic	urban dynamic		recreational demand	urban pos
bus station	travel		fine grain land use taxonomy	urban functioning		urban greenspace	urban pos
high quality travel	travel		coarse grain authoritative data	urban functioning		park attributes	urban pos
human travel	travel		urban data	urban functioning		park	urban pos
intra city travel	travel		residential area	urban functioning		services comprise	urban service
vacation planning	travel		local resident	urban functioning		service	urban service
transport	travel		clear topology	urban functioning		everyday life	urban social life
public transportation	travel		better understanding	urban functioning		life	urban social life
travel itinerary	travel		understand	urban functioning		social activity	urban social life
transit time	travel		land use	urban functioning		social breadcrumbs	urban social life
transportation need	travel		character	urban functioning		social dynamic	urban social life
travel itineraries	travel		factor	urban functioning		social life	urban social life
travel web site	travel		characteristic	urban functioning		life habit	urban social life
travelers origin	travel		condition	urban functioning		social spaces	urban social life
Twitterability	Twitter		state	urban functioning		everyday landscape	urban social life
Twitter	Twitter		street networks	urban functioning		livable places	urban social life
activity	urban activity		city characteristics	urban functioning		big project livable places	urban social life
function	urban activity		gentrification	urban functioning		individual user	user
land use	urban activity		urban functions	urban functioning		person	user
restaurant	urban activity		local authority	urban management		humans	user
active print	urban activity		functional organization	urban management		greater cross movement	user activity
active urban area	urban activity		city government	urban management		urban sports movement	user activity
activity area	urban activity		place design	urban management		outdoor physical activity	user activity
fitness applications	user activity		passive track	user activity		human mobility	user activity
sports activity	user activity		similar trait	user activity		park visit	visit
personal fitness data	user activity		usage	user activity		approximate visitation rate	visit
run	user activity		human activity	user activity		empirical visitation rate	visit
urban jogger	user activity		human behavior	user activity		entrance	visit
path	user activity		user activity patterns	user activity		entrance fee	visit
digital trace	user activity		people's behavior	user activity		higher visitation rate	visit
visitation rate	visit		visitor	visit		entry	visit
visitor attraction	visit						