

Multiscale geovisual analysis of knowledge innovation patterns using big scholarly data

Chenyu Zuo, Linfang Ding, Zhuo Yang, Liqiu Meng

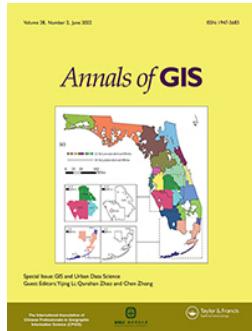
Angaben zur Veröffentlichung / Publication details:

Zuo, Chenyu, Linfang Ding, Zhuo Yang, and Liqiu Meng. 2022. "Multiscale geovisual analysis of knowledge innovation patterns using big scholarly data." *Annals of GIS* 28 (2): 197–212. <https://doi.org/10.1080/19475683.2022.2027012>.

Nutzungsbedingungen / Terms of use:

CC BY 4.0





Multiscale geovisual analysis of knowledge innovation patterns using big scholarly data

Chenyu Zuo, Linfang Ding, Zhuo Yang & Lique Meng

To cite this article: Chenyu Zuo, Linfang Ding, Zhuo Yang & Lique Meng (2022) Multiscale geovisual analysis of knowledge innovation patterns using big scholarly data, *Annals of GIS*, 28:2, 197-212, DOI: [10.1080/19475683.2022.2027012](https://doi.org/10.1080/19475683.2022.2027012)

To link to this article: <https://doi.org/10.1080/19475683.2022.2027012>



© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group, on behalf of Nanjing Normal University.



Published online: 16 Jan 2022.



Submit your article to this journal 



Article views: 1997



View related articles 



View Crossmark data 



Citing articles: 2 

Multiscale geovisual analysis of knowledge innovation patterns using big scholarly data

Chenyu Zuo^a, Linfang Ding^{ID b}, Zhuo Yang^a and Liqiu Meng^{ID a}

^aChair of Cartography and Visual Analytics, Technical University of Munich, Munich, Germany; ^bKRDB Research Centre, Faculty of Computer Science, Free University of Bozen-Bolzano, Bozen-Bolzano, Italy

ABSTRACT

Knowledge innovation is a key factor in industrial development and regional economic growth. Understanding regional knowledge innovation and its dynamic changes is one of the fundamental tasks of regional policy-makers and business decision-makers. Although many existing studies have been conducted to support in understanding knowledge innovation patterns, data-driven and intuitive visual analysis of georeferenced knowledge innovation has not been sufficiently studied. In this work, we analysed knowledge innovation by visually exploring big georeferenced scholarly data. More specifically, we first applied network analysis and statistical methods to derive key measures (e.g., the number of publications and academic collaborations) of knowledge innovation with multiple spatial scales. We then designed geovisualizations to explicitly represent the multiscale spatiotemporal patterns and relations. We integrated the analytical methods and geovisualizations into an interactive tool to facilitate stakeholders' visual learning and analysis of knowledge innovation with a spatial focus. Our work shows that geovisualizations have great potential in supporting complex geoinformation communication in knowledge innovation.

ARTICLE HISTORY

Received 18 June 2021

Accepted 1 January 2022

KEYWORDS

Geovisualization; big scholarly data; network analysis; knowledge innovation

1. Introduction

Knowledge innovation is one of the fundamental factors in regional development. Supporting knowledge innovation and collaboration becomes increasingly important in modern regional planning. Science and technology parks are good practices for the encouragement of knowledge innovation. For instance, Silicon Valley in California, Hsinchu Science Park in Taiwan, Arabianranta in Helsinki and Bio^M in Bavaria, have gained popularity and have successfully led to continuous knowledge innovation and academia-industry interactions (Henriques, Sobreiro, and Kimura 2018). This, in turn, attracts ideas, knowledge, innovation, population growth and financial investment in the regions (Carrillo et al. 2014). Therefore, fostering knowledge innovation, collaboration and academic spin-offs are key strategies in regional planning and management (Fikirkoca and Saritas 2012; Friedmann and Yong 2004). Understanding knowledge innovation patterns and regional development is essential for policy and decision-making.

The growing field of urban analytics in big geodata may bring more evidence and insights into spatiotemporal patterns of various social activities, including knowledge innovation. Analytical methods have been

proposed to measure and analyse many characteristics of a city, such as urban land functions, human mobility, life quality and place attractiveness. For example, Ye et al. (2021) identified the functions of urban areas using social media data. Brauer, Mäkinen, and Oksanen (2021) analysed the influence of cycling traffic on mobility using trajectory data. Sapena et al. (2021) predicted life quality in multiple dimensions by modeling socio-economic indicators. Among a wide range of research topics, big scholarly data are widely used to understand knowledge innovation (Hoekman, Frenken, and Van Oort 2009; Galaso and Kovářík 2021). For instance, Wang et al. (2021) used big scholarly data to analyse the correlation between academic innovation and high-tech industry and found that international collaboration has become more important for local industry in recent years. Another study proposed that functions of different cities in a polycentric region can be revealed by analysing multiscale academic collaboration (Li and Phelps 2018). The structural disparities and proximity mechanisms in academic collaboration can be measured by analysing the collaboration network (Chengliang, Qingchang, and Dezhong 2017; Li, Wei, and Wang 2015). In addition, analysing scholarly data can provide quantitative evidence to support decision-making. For

instance, a recent study, Imran and Jabeen (2020) analysed domestic and international scientific collaborations and found that international collaborations positively affect local knowledge innovation. Their findings provided some suggestions for research policies, such as establishing scholarships and investing in digital infrastructures.

While most existing studies have focused on applying statistical methods for strategic analysis and anticipating the future, effective visualizations, that leverage the human ability of visual analysis, play an important role in improving transdisciplinary knowledge among various stakeholders (Billger, Thuvander, and Wästberg 2017). As Kandt and Batty (2021) suggested, geovisual analysis and interactive interfaces are needed in smart cities and urban policy-making to facilitate communication between various stakeholders, such as investors and planners. The innovation capability of a region and its connections with other regions contributes greatly to economic development. Geovisualizations, illustration of specific topics and provision of geographic knowledge are beneficial tools for regional planning and decision-making. In general, it is challenging to design geovisualizations that represent the dynamic development of cities, especially when covering a long-time frame to gain strategic insights (Kandt and Batty 2021). Maps are often used to visualize spatial information in urban analytics (Maantay and Ziegler 2006). For example, Sobral, Galvão, and Borges (2019) used flow maps to study traffic and pedestrian movement and supported the analysis of passenger dynamics, ridership and urban service reliability. Zhao et al. (2020) used machine learning methods based on zoning strategies to analyse urban population distribution and showed the results on choropleth and heat maps. Their results had high accuracy and helped in communicating regional inequalities. 3D visualization was also adopted to provide an intuitive presentation of urban space and help the general public participate in the urban planning process (Wu, He, and Gong 2010).

Many studies have also been conducted to visualize big scholarly data for the analysis of knowledge innovation patterns. However, most have focused on developing static visualizations to show individual perspectives of publications, such as popular scientific domains, geographic publication hotspots, co-authorship and citation networks, whereas the connections among these perspectives and the relations of scientific networks in multiple scales were not adequately presented. Interactive tools for the integration of the visualizations and the display of the connections between them should be designed and developed to bridge this gap. However, there are several challenges

for designing interactive tools. First, the design of visualizations to show multivariate and multiscale spatiotemporal data is difficult. To be specific, scholarly data has multiple variables, such as domains, journals, citations and publishers (Leydesdorff, Bornmann, and Wagner 2017), and can be analysed at multiple aggregation levels, such as author level, topic level, institute level and administration level, as well as multiple geographic coverages, such as local, municipal, regional, national and international (Galaso and Kovářík 2021; Frenken, Hardeman, and Hoekman 2009; Csomós and Lengyel 2020). Second, design and development interactive tools to support dynamic analysis of multi-level and multi-coverage networks require interdisciplinary knowledge, such as knowledge innovation studies, visual analytics, cartography, and human-computer interaction. Third, there are some technical challenges related to the georeferencing of implicit spatial locations, because scholarly data is normally collected with implicit locations, for example, in the form of name and address of affiliation. When conducting geospatial analysis, there is a need to convert these names and affiliations to geographic locations relying on geocoding techniques.

This study aims to combine analytical and visualization methods to uncover knowledge innovation patterns from big scholarly data. We explored knowledge innovation in multiple spatial scales over a long time span and collaboration networks in several selected research domains. Novel geovisualization methods were applied to illustrate and reveal the complex academic networks and their geographic distributions. The visualizations were further integrated into a web-based interactive analysis tool to support strategic decision-making.

2. Related work

2.1. Network analysis of knowledge innovation

Network analysis with a set of integrated techniques is widely applied in the social sciences to explore relations and interactions among actors in societies (Borgatti et al. 2009), such as friendship networks in organizations, follower networks in social media, spread networks of news and rumours, as well as co-authorship and citation networks in academia (Tabassum et al. 2018). In academic network analysis, network theories are widely applied with multiple levels of detail ranging from individual authors, to institutes and regional and global scales.

Fine-level knowledge networks are analysed based on the characteristics of individual nodes in the network. For instance, by analysing citation networks, Imran and Jabeen (2020) and Chen, Ibekwe-sanjuan, and Hou (2010) identified the influences of journals and institutes.

Basole (2016) explored topological metrics of inter-firm collaboration networks, such as centrality and clustering coefficients, to describe the characteristics of nodes in the network. Similarly, Chen et al. (2017) calculated the metrics for directed networks, such as PageRank and connected component sizes, to describe the clusters of the network. Li et al. (2018) detected the most influential authors by analysing metrics such as number of papers, network density and centralities.

The complex patterns of the relations between knowledge innovation and industry, urban structure and regional collaborations were analysed as well utilizing network analysis methods. Wang et al. (2021) studied how the cooperation network supports innovation-driven industries. They found that interregional and international innovation collaboration are important driving forces for economic development. They compared statistics, including the total number, mean, standard deviation, minimum and maximum of regional and county-level networks. Innocenti, Capone, and Lazzaretti (2020) investigated the relationship between the structure of knowledge innovation and local industry at the firm level. Their results showed that knowledge networks are correlated with regional innovative capacity. Statistics of the networks, such as scale, density and average path, were measured in their analysis. Galaso and Kovářík (2021) compared the influence of different geographical levels of embeddedness on innovation and found that regional- and country-level innovation networks exerted different influences. They further analysed the correlation between the density of the co-patenting networks and the socioeconomic indexes, e.g. gross domestic product (GDP) per capita and human capital.

In general, analysing knowledge innovation in multiple aggregation levels, such as institutes or cities, and in multiple spatial coverages, such as local or international, is important for understanding knowledge innovation patterns. In addition, location and distance in the aforementioned studies are crucial for knowledge innovation and therefore demand analysis of knowledge innovation with a spatial focus.

2.2. Geovisualization of scholarly data

Visualization of big spatiotemporal data could reveal a large amount of information at a glance or emphasize the patterns of certain key features. It is an important tool for further analysis in decision-making (Hamdi et al. 2021). Many studies have been conducted to explore innovative, aesthetic and effective visualization designs to support the identification of significant locations and time periods. Seebacher et al. (2018) overlayed designed glyphs onto a map to show the temporal changes of

certain species. Leite et al. (2020) designed donut charts which, when superimposed on a map, show economic networks. Li et al. (2020) designed glyphs to show the location and movement of different wild animals.

Various visualization methods and techniques have been developed to support understanding complex spatiotemporal trends in knowledge innovation. In the well-known scholarly analysis tool CiteSpace (Chen et al. 2006), different visualization methods were integrated for the analysis of data from multiple perspectives. For example, network visualization with weighted nodes and links was used to support cluster detection; coloured bars to visualize the temporal periods of certain keywords; flow maps to show geographic hotspots; word clouds to illustrate the popularity of research domains.

In recent years, visualization methods have been increasingly applied in network analysis (Liu et al. 2018), and maps have been widely used as an instrument to show geographic patterns in networks. Examples include colour-coded flow maps that show the connections between nodes (Hennemann, Derudder, and Taylor 2015), contour lines that represent the degree of nodes (Zhao et al. 2015) and grid maps for spatial cluster identification (Piry et al. 2016). Furthermore, the complex patterns in collaboration networks can be represented by various geovisualizations, such as 3D symbol maps (Helbig et al. 2014), flow maps (Jenny et al. 2018) and treemap visualizations (Scheibel et al. 2020). In addition, Csomós used maps to show various perspectives of spatial patterns of publications. For instance, he used proportional symbol maps to show the hotspot of scientific output areas (Csomós 2018) and designed a flow map to show international scientific collaborations (Csomós and Lengyel 2020). Moreover, spatial information is represented in other novel visualizations. For example, Liu, Derudder, and Taylor (2014) applied alluvial diagrams to analyse international connections, and Hennemann, Derudder, and Taylor (2015) employed a chord diagram to visualize intercity networks.

Although various visualizations are proposed to explore statistical attributes and spatial distributions of networks, few interactive tools are designed and developed for visual analysis of networks and connections from different perspectives. For example, tools are needed to support the overview of networks with multiple scales and aid users in comparing the differences among networks in different scientific domains. To support such functions, relevant visualizations need to be integrated and linked into an interactive visual analytical tool, which should allow users to choose the content based on their own interests. Multiple factors should be

considered in designing interactive visual analysis tools with respect to their target users, such as showing aggregated data in required geographic scales, designing the appropriate visualizations, arranging the visualizations logically and applying effective interactions (Few 2006; Zuo, Ding, and Meng 2020). Interactive tools have been designed to support various purposes, e.g., map-based interactive tools were designed to discover cyber traffic spatial patterns (Kodituwakku et al. 2020; McKenna et al. 2016) and support visual detection and monitoring of air quality (Liu et al. 2021). In a big scholarly data analysis, Bach, Pietriga, and Fekete (2014) visualized the topology of collaboration among authors over time by using a space-time cube. However, more types of geovisualizations can be integrated into an interactive tool to support visual analyses of scientific collaboration patterns.

3. Data and methodology

3.1. Big scholarly data

Big scholarly data is data about scientific resources and it is rapidly growing. It usually includes information about papers, authors, institutes, journals and conferences and reflects scientific activities, personnel, research domains and scholarly networks (Xia et al. 2017). Scholarly data can be easily accessible with the help of many online libraries and platforms, such as digital libraries like Web of Science, academic search engines like Google Scholar, and academic social media like ResearchGate. Many studies employed scholarly data to detect academic activities (Xia et al. 2017). For example, impact of authors can be predicted by analysing citations (Dong, Johnson, and Chawla 2016), and interactions among research communities can be detected by analysing co-authorships (Mercorio et al. 2019). Similarly, keywords and citations can be utilized to identify influential publications and research frontiers (Yan and Li et al. 2020). The temporal changes of the popularity of research domains can be identified by analysing the number of publications and journals (Wang et al. 2020). In addition, scholarly data can also reflect academic communities, such as university groups and research institute clusters. In this study, we employ big scholarly data to explore the patterns of knowledge innovation.

We collected scholarly data from the ACM Digital Library via web crawling technologies. The ACM Digital Library (<https://dl.acm.org/>) is a public research platform with a large collection of literature in the field of computing. According to the ACM categories, journals, proceedings, letters and magazines reflect the emerging

Table 1. An example from the collected raw data.

Item	Data
Title	Domain-aware multi-truth discovery from conflicting sources
Author	Xueling Lin Lei Chen
Affiliations	Hong Kong University of Science and Technology, Hong Kong Hong Kong University of Science and Technology, Hong Kong
Published time	1 January 2018
Domain	Information systems Information systems applications Data mining
Domain ID	10002951 10002951.10003227 10002951.10003227.10003351
URL	https://dl.acm.org/doi/10.1145/3177732.3177739

and established computing research, cutting-edge innovations, leading views and opinions on research, and up-to-date research activities, respectively [56]. In this study, we have therefore collected publications in these four categories. The collected items are from January 2017 to December 2019, with information about the title of the publications, the name and affiliation of the author(s), the publication time, the research domain and the uniform resource locator (URL) of the webpages. The collected data set contains 137,818 publication items, 305,815 authors and 39,030 institutes worldwide. Among them, 85.0% (117,150 items) are proceedings, 10.6% (14,627 items) journals, 2.6% (3533 items) letters and 1.8% (2478 items) magazines. Table 1 shows an example of a publication item from the raw data. The names and affiliations of the authors correspond in their orders. The domain IDs follow the poly-hierarchical ACM domain ontology.¹ There are 13 main domains, such as computing methodologies and human-centered computing, and each domain is further divided into subdomains.

3.2. Methodology

We proposed a geovisual analytical workflow to examine the spatiotemporal patterns of knowledge innovation. The workflow consists of four major modules, i.e., data processing, data modeling, geovisualization and the interactive tool. Figure 1 shows the modules and their major steps. Firstly, we preprocessed the scholarly data by cleaning, georeferencing the affiliations to obtain the geographic coordinates and aggregating the publications onto geographic locations with multiple scales. Secondly, the co-authorship networks were established and analysed by different methods in the modelling layer. The institutes, cites and research domains were analysed at multiple scales. Thirdly, novel visualizations were designed to show the patterns from various perspectives with a spatial focus. Finally, an interactive

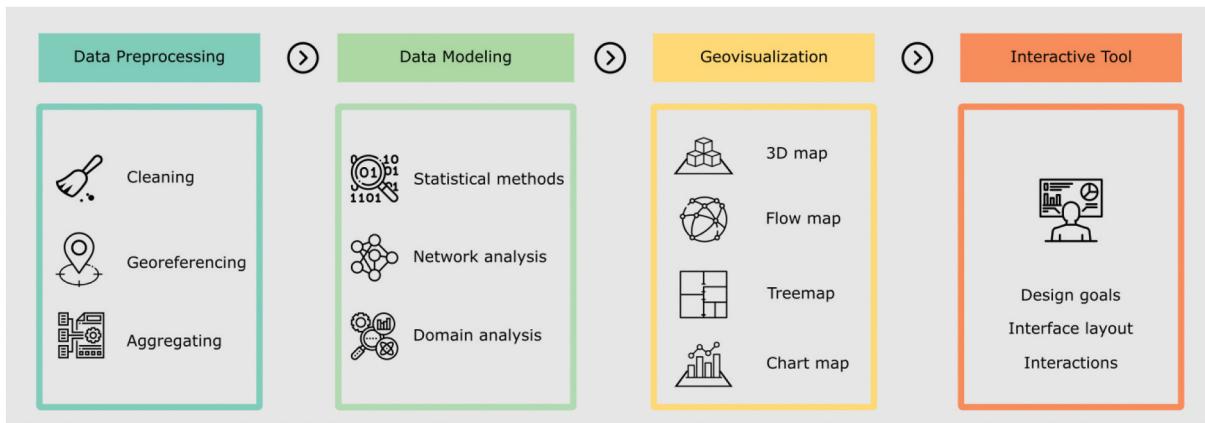


Figure 1. The visual analytical workflow chaining together four modules: data preprocessing, data modeling, geovisualization and the interactive tool.

visual interface was designed by formulating the design goals and designing accordingly the layout and the interactions.

3.2.1. Data preprocessing

Before data visualization and analysis, the scholarly data needs to be cleaned, georeferenced and aggregated. In this work, data cleaning included data restructuring and text encoding. The data was restructured into four types of entities, including paper, author, institute and domain, and these entities were connected with extra tables (as Figure 2 shows). Text encoding converts non-English letters and mathematical symbols into the UTF-8 encoding system. Next, we used the Nominatim² open-source geocoding service to georeference the names and addresses of affiliations into longitudes and latitudes and then project the coordinates of the affiliations onto multiple administrative divisions, such as cities, regions, or counties. The administration boundary data were collected from the <http://nnu.geodata.cn:8008/geodata> open data portal. Finally, the number of publications, co-authorship, and domain information were

aggregated into the multiple levels of administrations. All of the preprocessed data was stored in the PostgreSQL database.

3.2.2. Data modeling

Social network research is dedicated to studying relationships between connected actors (van Steen 2010). An actor could be a person, an organization, a webpage, or a region. In a network, the actors are represented as nodes and their relationships are represented as links or edges. Network analysis may focus on different levels, e.g., node level, group level, or network level. According to Li, Wei, and Wang (2015), an academic network is a complex network with some unique features, such as a heavy tail in the degree distribution and a high clustering coefficient (Yang and Yang 2008).

In this study, we measure knowledge innovation by the number of scientific publications and research collaborations. We established academic networks at two representative levels, i.e., institute level and the city level, and we analysed them locally, regionally and internationally. In the academic network at the institute level,

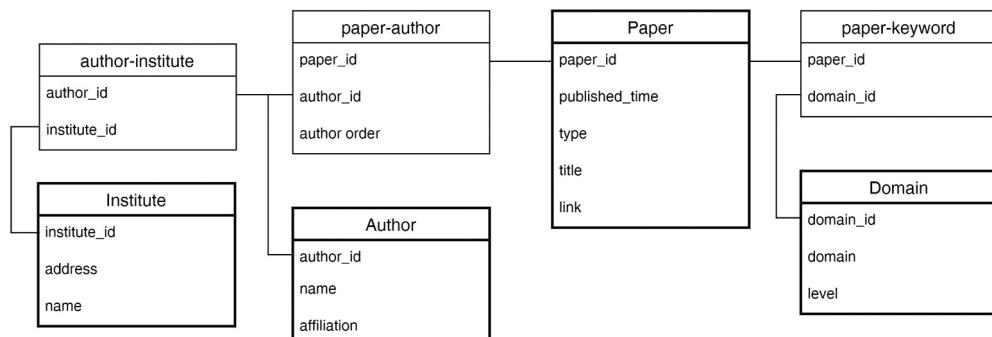


Figure 2. The data schema of the restructured scholarly data.

nodes represent institutes where authors' affiliations and weights of nodes represent the number of publications of an institute. The links represent co-authorships between two institutes, and the weights of links represent the number of joint publications. For instance, if a publication had several authors with their affiliated institutes, each distinct affiliation of the publication was represented as one node, and this publication in turn contributed to the weight of each node. Links were established between every two involved affiliations, and each pair of authors with different affiliations contributed to the weight of the link. Similarly, the network was also established at the city level, with each city had publications as a node and the intercity co-authorship as links. In addition, the weights of the nodes were calculated with integer counting, which means the affiliated institute or city receives full credit for each author of a publication. We consider the networks to be undirected because the communication among the collaborated authors could be bidirectional.

To explore knowledge innovation patterns and to further support decision-making in regional analytics, analysis methods on the network level were applied. We selected scale, node, link, degree centrality and gravity centre to measure the academic networks. In addition, the geographic locations of the nodes were calculated, and the spatial distributions were analysed. The metric scale measures the number of nodes in a network, indicating the number of institutes or cities involved in the publications. The weights of nodes and links reflect the number of publications and collaborations. We further calculated the median and standard deviation of the weights to measure the statistical distributions of the publications and the collaborations at the institute and city levels. The degree centrality of network was calculated as the number of links a node has. We used it at the city-level network to measure the activeness of cities in intercity collaborations. The gravity centre is the weighted centre of the nodes, indicating the theoretical central location of a network. We

calculated it at the city-level network to analyse the spatial distribution of the knowledge innovation in the region. It is calculated as

$$Longitude_{GC} = \frac{\sum_{i=0}^n Longitude_i \cdot Weight_i}{\sum_{i=0}^n Weight_i}$$

$$Latitude_{GC} = \frac{\sum_{i=0}^n Latitude_i \cdot Weight_i}{\sum_{i=0}^n Weight_i}$$

where GC is the gravity centre, $Longitude_{GC}$ and $Latitude_{GC}$ indicate the coordinate of the gravity centre, i is a node in a network, $Longitude_i$ and $Latitude_i$ indicate the coordinate of Node i , and $Weight_i$ is the number of publication of Node i .

3.2.3. Geovisualization

To support an intuitive and in-depth analysis of the characteristics of the academic networks, we applied geovisualization methods to reveal the patterns in different aggregated levels and spatial coverages. In particular, maps were used extensively to support exploring and understanding spatial patterns and relationships. We identified four important types of spatial information in knowledge innovation: the location and number of publications, the location and number of collaborations, the proportion of publications in different domains, and the spatial distribution of specific domains. In this study, we selected and applied geovisualizations to show these features.

Figure 3 illustrates four types of selected visualization methods, i.e., a 3D symbol map, a flow map, a treemap and a chart map dedicated to the representation of these four types of data features. We applied the 3D symbol map to show the spatial locations of institutes and their number of publications, that can reflect the innovation ability. As Figure 3(a) shows, each bar represents an institute, and the height of the bars represents the number of publications in the institute. The academic network at the city level reflects regional innovation and collaboration

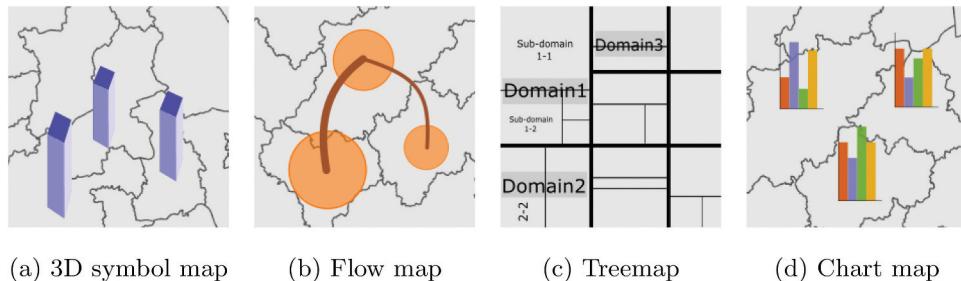


Figure 3. The visualization methods that show the patterns of knowledge innovation with spatial information.

capabilities. We adopted the flow map to show the nodes and the links of the network at the city level. As Figure 3(b) shows, each city is represented as a circle, and the number of publications of city is represented by the size of the circle. The arcs represent the collaborations among cities, and the width of arcs represents the number of co-authorships. To reflect the hierarchical structure of research domains and their relative proportions, we utilized the treemap as shown in Figure 3(c). Each square represents a domain or a subdomain. The entire square is divided into several regions. We visualized two levels of domains in the treemap: The first-level domains are represented by the squares with bold frames, and the subdomains by squares with light frames. The chart map was used to show the spatial distribution of several domains, so that the spatial distribution of individual domains can be revealed and compared. Figure 3(d) shows a chart map representing the number of publications of four subdomains in cities. The colour of bars represents a specific academic domain, and the height of bars shows the number of publications.

3.2.4. Interactive tool

With the support of visual and analytical functions, interactive tools can help decision-makers improve their understanding of knowledge innovation in various applications. Based on the aforementioned studies in Section 2.1, we formulated the design goals as follows:

A focus on spatial patterns. The tool should provide visualizations of the knowledge innovation network with geographic locations, so that users can identify spatial clusters and compare different areas intuitively.

Multi-level knowledge innovation exploration. The tool should support the visualization of networks at multiple levels, such as at the institute level and city level.

Multi-coverage knowledge innovation exploration. The tool should aid users to explore and relate between local, regional and international knowledge innovation networks.

Research domain analysis. The tool should enable users to analyze the popularity of research domains and their spatial patterns.

Easy to use. The interface should be easy to understand and provide simple but effective interactions for users.

Driven by these goals, we designed the interactive tool with an integration of the visualization methods mentioned in Section 3.2.3. Figure 4 shows the interface layout of the conceptual design, which consists of eight panels. The title panel describes the name of the

Title		Information
Area Selector		Map Selector
Treemap	Local coverage map	Domestic coverage map
		International coverage map

Figure 4. The interface layout of the interactive tool.

tool. The information panel shows the data source and designers. There are two selector panels in the middle of the interface, , allowing users to switch areas and types of maps. The treemap panel shows the popular research domains of a specific region. The remaining three map panels show the networks in the corresponding region on 3D symbol maps, flow maps and chart maps with local, domestic and international coverage, respectively.

4. Case study

In this section, we applied the analysis and visualization methods on a case study to explore the knowledge innovation patterns reflected by the ACM Digital Library data. More specifically, we analysed the patterns of knowledge innovation from the perspectives of academic organizations, spatial relations and influential research domains at multiple scales. In addition, an interactive tool was designed and developed for users to explore knowledge innovation patterns in a test area.

4.1. Test area

We chose the Yangtze River Delta (YRD) as our test area. It is located in the east of China and covers more than 358,000 square kilometres (about 3.7% of China). It is one of the economically leading regions of China, including four provinces and containing 41 cities. With about 16% of the Chinese population, the YRD contributes about 24% of Chinese GDP (China statistical yearbook 2020). Figure 5 shows the location of the YRD and the administrative boundaries of the provinces and cities.

Although the YRD had an economic boom during recent decades, its economy is spatially unbalanced. YRD regional economic development is highly correlated with the industrial structure, and the collaborations between cities are largely related to factors such as spatial distance, population and the local economies (Ye et al. 2019). Therefore, analysing the structure of knowledge innovation of the YRD can benefit policy-



Figure 5. The administrative boundaries and location of the Yangtze River Delta (YRD).

makers and planners, e.g., helping them understand the local spatial spillover effect (Wu et al. 2017) and supporting formulating research policy (Imran and Jabeen 2020; Diercks, Larsen, and Steward 2019).

4.2. Analysis of academic institutes

Analysis of individual research institutes and their collaborations is crucial for understanding sustainable innovation capability. We analysed the spatial distribution of

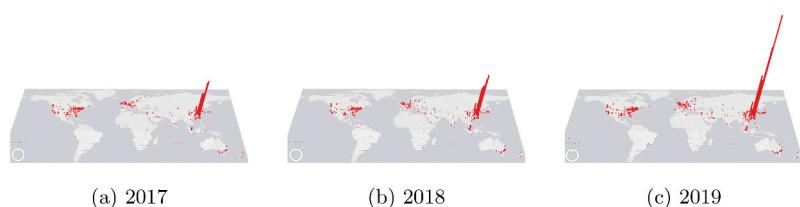


Figure 6. Spatiotemporal distribution of institutes that have joint publications with the YRD region.

institutes and the number of their publications. In this analysis, we selected worldwide institutes that have joint publications with institutes located within the YRD region. The spatial patterns were visualized in the 3D symbol maps as shown in Figure 6. We found that the number of publications have increased significantly from 2017 to 2019, especially between 2018 and 2019. Several spatial clusters can be identified that have joint publications with the YRD region and that are mainly distributed in East China, Northern America and Western Europe. Besides, there was a large amount of collaboration with Singapore and Japan.

The statistics of the publications and their affiliated institutes are shown in Table 2. The statistical values also confirm the significant increasing trend of the publications and the collaborated institutes that have collaborated with the YRD from 2017 to 2019. Compared with inter-regional joint publications, the increase of intra-regional publications within the YRD was more significant.

Table 2. The statistics of the publications and collaborations of the YRD from 2017 to 2019.

Statistical items	Year		
	2017	2018	2019
Number of internal joint publications within the YRD	611	851	1885
Number of local institutes involved in internal joint publications within the YRD	231	330	475
Number of domestic joint publications	679	802	1227
Number of domestic institutes involved in joint publications outside the YRD	192	239	343
Number of international joint publications	497	561	826
Number of international institutes involved in joint publications	377	419	469

4.3. Analysis of spatial relations

Regional and international collaborations reflect the collaboration network from different perspectives and are important for resource integration in the YRD region. Li, Wei, and Wang (2015) have found that the scientific network in China meets scale-free properties. In this section, we analyse the characteristics of the intra-regional academic network based on prefectural cities and the international collaborations of the YRD. We calculated the nodes and links on the aggregated level as mentioned in Section 3.2.2. We then calculated and analysed the attributes of the networks. In addition, we visualized the spatial information of the networks on flow maps as mentioned in Section 3.2.3.

Figure 7 shows the spatial and temporal intra-regional network patterns from 2017 to 2019. We can see that the networks were well connected in the centre while surrounded by some isolated nodes. In addition, not all cities were involved in the networks. A few cities are very active in publications: Shanghai, Hangzhou, Nanjing and Hefei. The scale of the networks did not show an obvious growth, but the number of publications and the connections in the active cities had grown greatly. This suggests that few new research institutes appeared between 2017 and 2018, but that the academic activities in the existing institutes increased greatly.

Table 3 shows the statistics of the city-level network within the YRD. It confirms that the increase of publications and collaborations varied in different cities. The number of cities involved in publication increased from 2017 to 2019. The average number of publications was much higher than the median and the standard

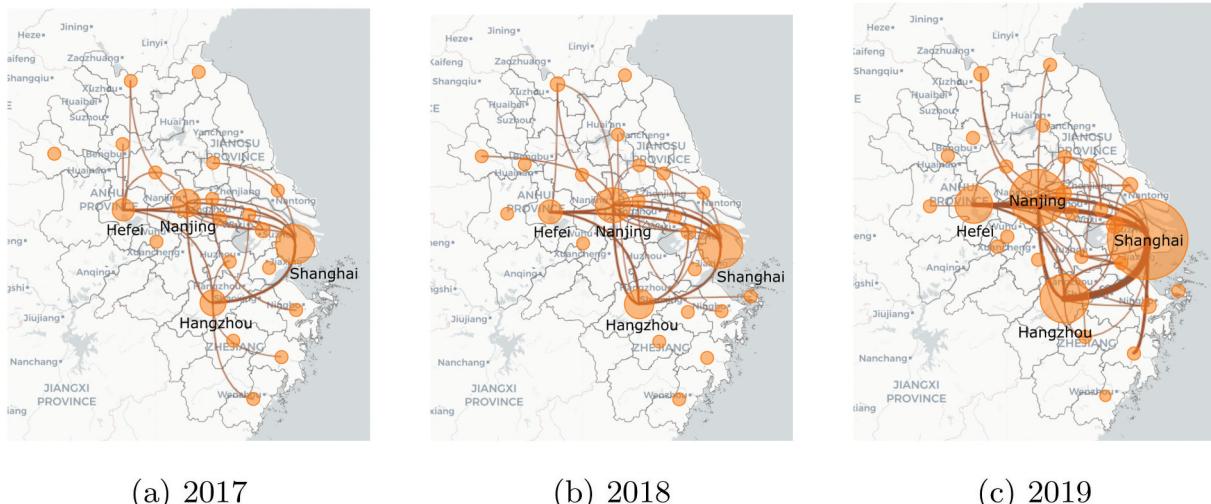


Figure 7. Flow maps showing the number of publications and co-authorship connections between cities in the YRD.

Table 3. Statistics of the academic networks at the city level within the YRD from 2017 to 2019.

Statistical items	Year		
	2017	2018	2019
The number of the cities involved in publication	22	25	28
The average number of publications in the cities	154.1	173.5	317.8
The median number of publications in the cities	10.5	12	14.5
The standard deviation of number of publications of the cities	324.7	383.8	760.3
The number of city pairs involved in collaborations	27	24	39
The average publications in city-pairs	7.4	11.5	15.0
The median publications in each city-pairs	4	7	5
The standard deviation of publications in city pairs	8.3	12.0	25.0

deviation values were very high, which means that a few cities had the most publications. However, the intercity

collaborations decreased in 2018 and increased in 2019. Similarly, most intercity collaborations were between a few popular city pairs.

Figure 8 shows the histogram of the degrees of the city-level network. It shows the distribution of the number of collaborations between cities within the YRD. We can see that the intercity collaborations grew. Some of the cities reached out to have more than 10 collaborated cities in 2019, whereas most cities had small-scale intercity collaborations.

By applying the same approach, we visualized the spatial and temporal patterns of international collaboration with the YRD from 2017 to 2019 in Figure 9. It shows that the number of international collaboration countries

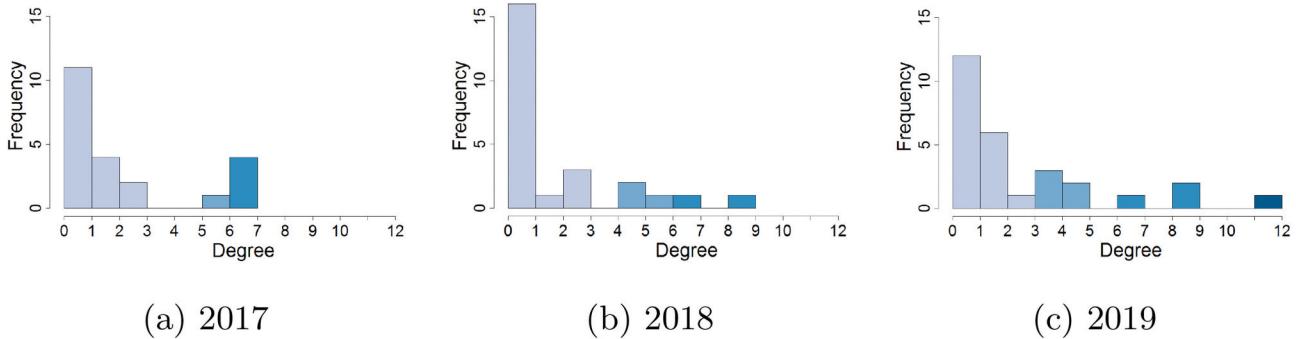


Figure 8. Histograms illustrating the number of collaborations between cities in the YRD.

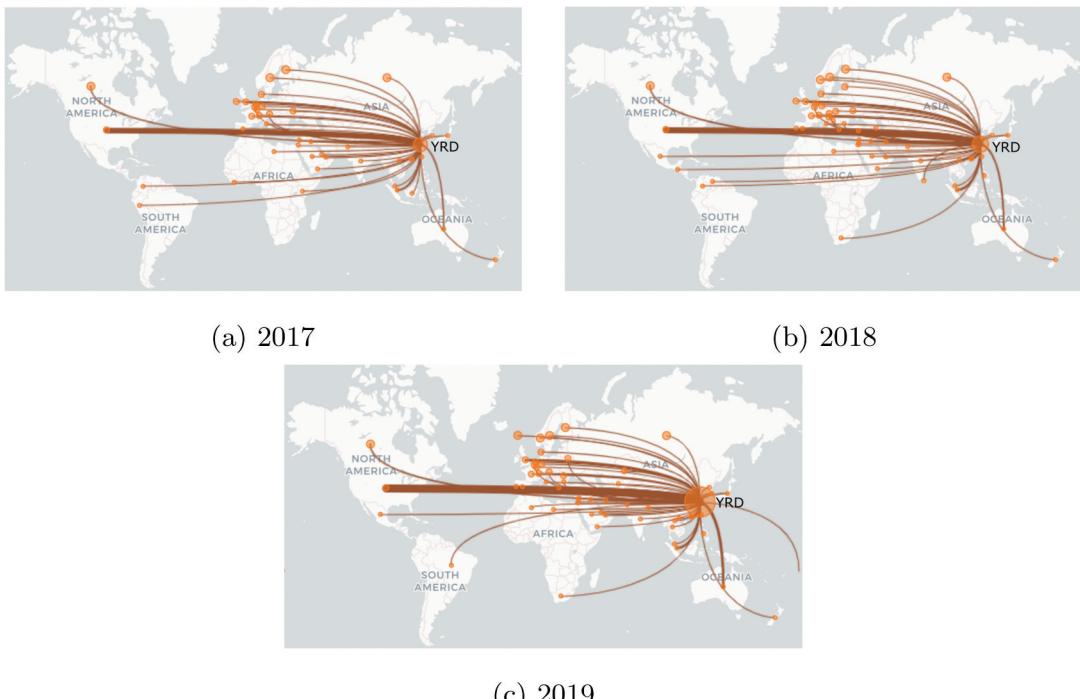


Figure 9. Flow maps of international collaborations of the YRD.

showed a great increase, whereas the number of publications of each country did not show much change from 2017 to 2019.

4.4. Analysis of influential domains

Regional competitiveness of knowledge innovation can be partially reflected in the popular research domains of computing and information technology. We first visualized the proportion of the research domains and subdomains, and then selected four representative subdomains for deeper analysis with visual representations of their spatial distributions.

We applied the treemap visualization method to show the proportion of various domains. In Figure 10, two levels of the domains from the publications during the years 2017 to 2019 were aggregated and visualized. The size of the squares represents the relative number of publications in each domain. The treemap is divided into 13 sections representing the 13 domains, each further divided into various subdomains. The subdomains within each domain have similar colours. We can see in the top level of Figure 10 that the two largest domains in

YRD were Computing methodologies and Information systems, whereas the two largest worldwide domains were Computing methodologies and Human-centred computing.

We selected four representative subdomains, artificial intelligence (AI), human-computer interaction (HCI), life and medical sciences (LiMS), and machine learning (ML) from the popular subdomains for further spatial analysis. We showed their spatiotemporal distributions at the city level using bar charts (shown in Figure 11). In Figure 11, we can see that the number of publications in AI and ML increased exponentially in cities like Shanghai, Nanjing, Hefei, and Hangzhou from 2017 to 2019, while the publications related to HCI and LiMS increased relatively slowly. In addition, a growing number of cities had publications in these four domains.

Gravity centre was used in a previous study on spatial scientific activity assessment (Ye et al. 2019), which showed that the YRD economic gravity centre had a northwards movement from 2003 to 2015. We calculated the weighted centres of publications of the four selected subdomains and marked the trajectories of the centroids from 2017 to 2019 in Figure 12. We can see that the centroids generally moved to the north in 2018

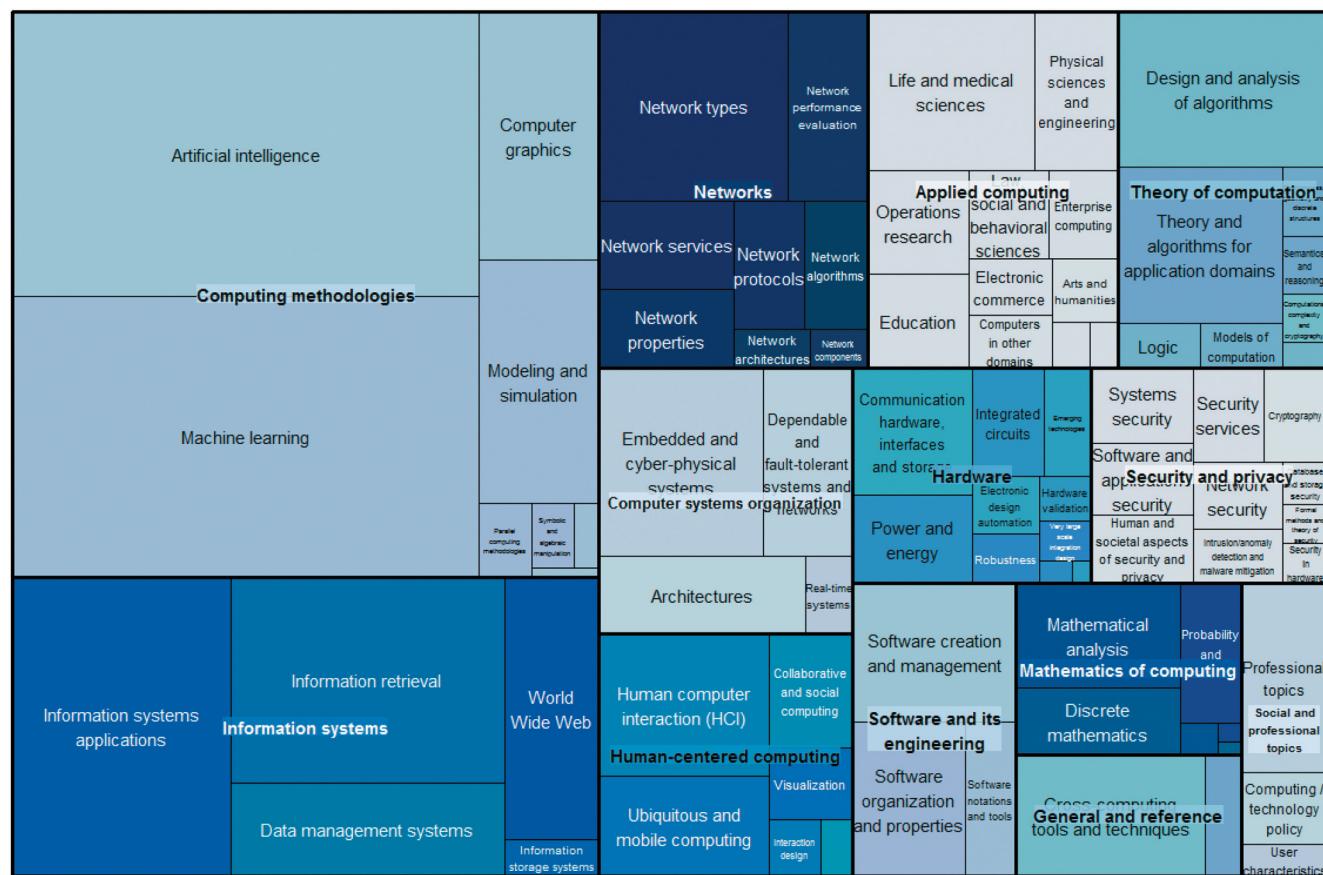


Figure 10. A treemap showing the popular domains of the ACM publications in the YRD from 2017 to 2019.

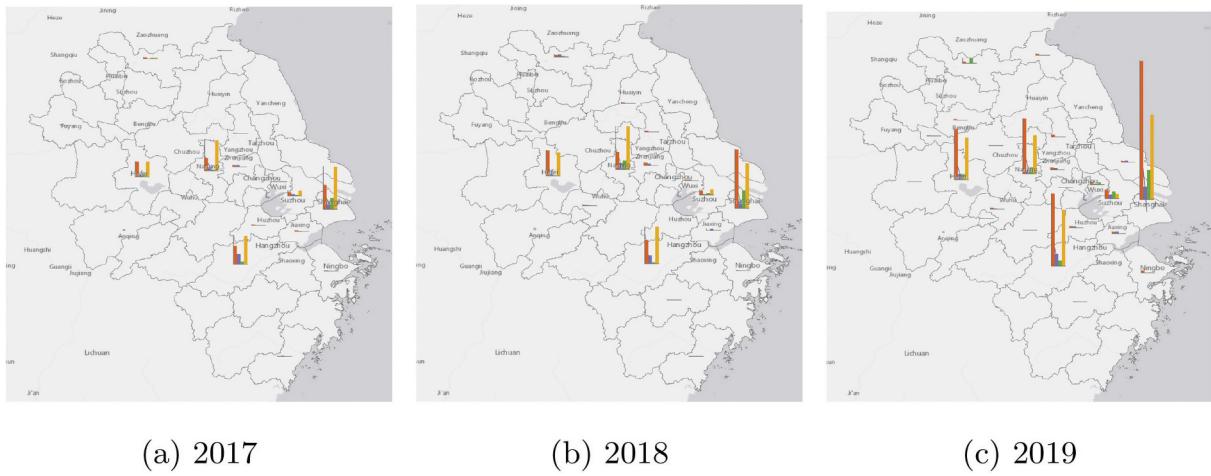


Figure 11. Bar charts showing the spatial distribution of the four selected subdomains in the YRD. AI in red, HCl in purple, LiMS in green and ML in yellow.

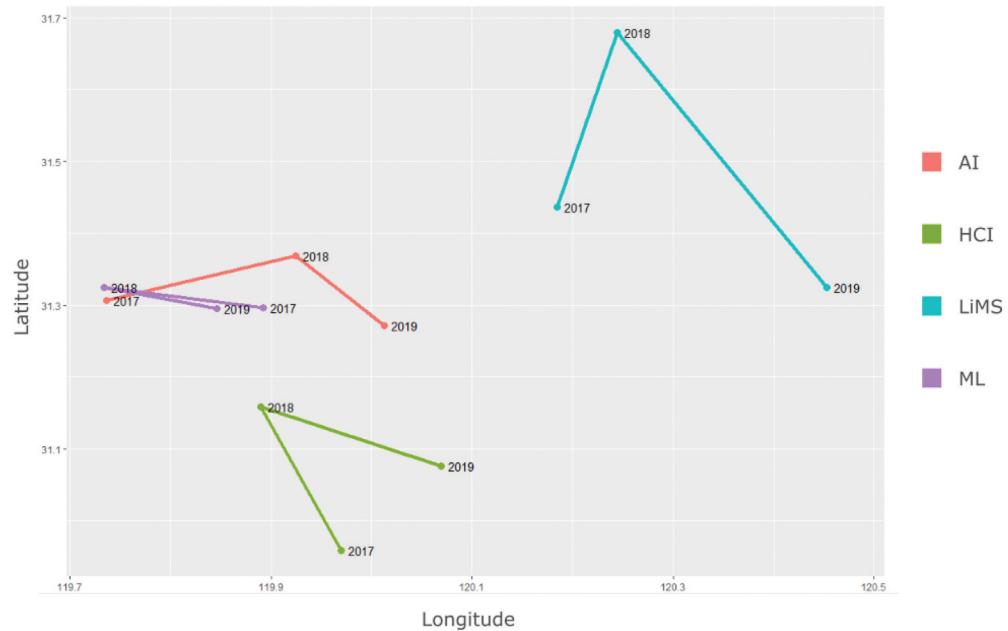


Figure 12. The movement of the centroids in the subdomains of AL, HCl, LiMS and ML from 2017 to 2019 in the YRD.

and to the south in 2019. The centroids in the subdomains of ML and HCl moved westwards in 2018 and eastwards in 2019, whereas AI and LiMS moved eastwards in both years. However, the movement trends may be better observed with a larger data set.

4.5. An interactive visual analysis tool

We developed an interactive tool, integrating geovisualizations and providing the stakeholders a structured overview. The tool aims to support users' understanding

and analysis of the knowledge innovation patterns in terms of spatial distribution of the institutes, the collaborations, and the popular domains. As described in Section 3.2.3 and shown in Figure 13, the tool is composed of eight panels, including a title panel, a information panel, an area selector, a map selector, the research domains visualized in treemaps, and the spatial distribution of the academic network visualized in flow maps in the YRD region, in Chinese domestic area, and the international coverage. The users follow the natural reading order from top to bottom (knowing

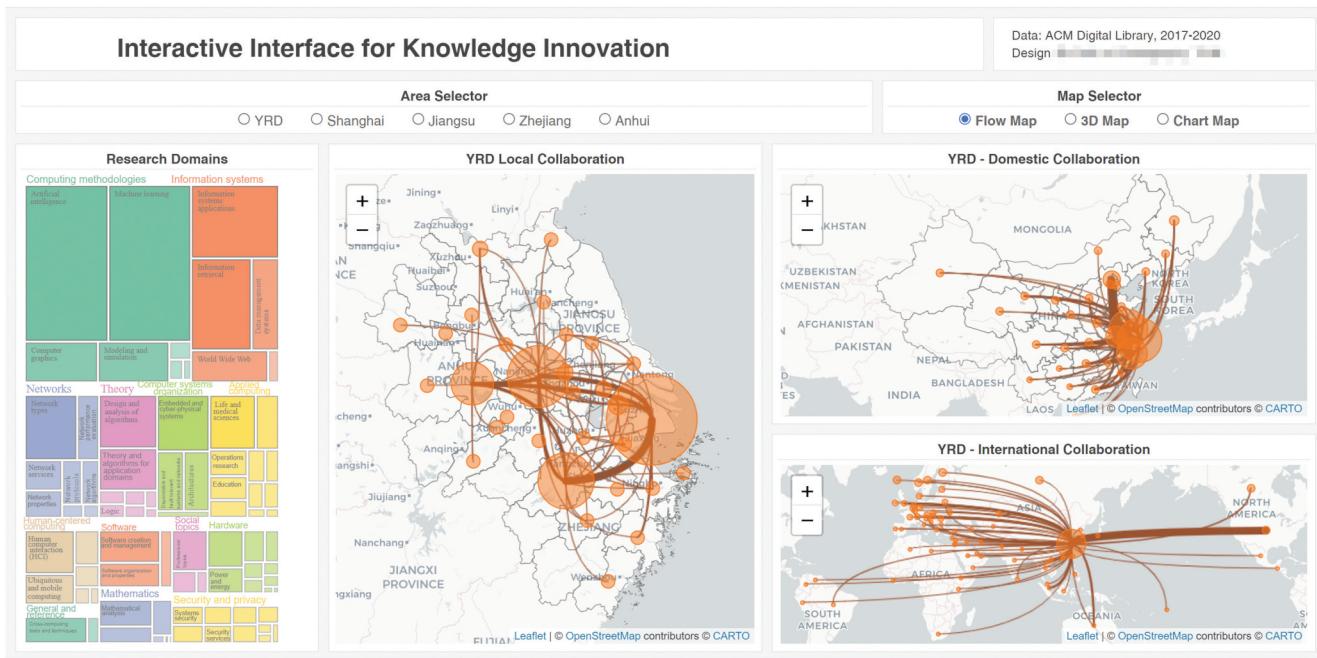


Figure 13. Visual interface of the analysis tool.

the general background to the details) and from left to right (from the innovation domains to their spatial distributions). Additionally, users could select focus regions by clicking on the area selector. If a region is selected, the maps will adjust the content to the selected region. Users could select the flow map, 3D symbol map, or chart map to show the knowledge innovation in multiple perspectives. The interactive tool is available at <http://129.187.45.33/KnowledgeDash/>. However, the interactions are not completely available in the online version. The Area Selector is functional, but the Map Selector has not yet been integrated to the current version.

5. Conclusion

This study proposes a geovisual analytical method coupling network analysis with geovisualizations for multi-scale analysis of big scholarly data. We designed and developed an interactive visual analysis tool to support users in discovering comprehensive patterns of knowledge innovation. The spatial clusters and centroid movements were analysed and visualized at different levels, from the institute and city levels, and in different local, regional, and international coverages. The interactive tool was designed and developed to help stakeholders investigate knowledge innovation from different perspectives.

This methodology is applied to analyse academic networks in the Yangtze River Delta region using scholarly data from 2017 to 2019 collected from the ACM library.

Our analytical results showed an increasing trend in the number of publications and research collaborations in the region. We further found that there are large spatial variations in the distribution of the publications and their growth. Specifically, most publications were from a few hotspot cities such as Shanghai, Nanjing, Hefei and Hangzhou, and there are more collaborations among the hotspot cities. Inside the hotspot cities, a large part of the increase in publications originated from a few particularly active institutes. On the contrary, there were few initiatives and research collaborations in under-developed cities. For instance, the cities, Anqing, Huangshan, Quzhou and Lishui should be better integrated into the regional academic network. Moreover, the temporal trend varies in different domains. Some worldwide popular domains (AI and ML) demonstrated rapid increase in the YRD as well, whereas other popular domains grew relatively slowly. In line with the previous studies (Wang et al. 2021), international collaborations significantly impact regional innovation, such as human-centred computing. Finding such domains that are highly valued worldwide but underdeveloped in the YRD can encourage further investigation into the causes.

Despite the aforementioned findings, we are aware of some limitations and formulated future recommendations from this study. The scholarly data from the ACM digital library used in our study is focused on science-based innovation. This may not comprehensively and adequately reflect knowledge innovation patterns. According to

Håkansson and Waluszewski (2007), knowledge and innovation are embedded in various recourses. To obtain a holistic picture of scientific developments, further data sources on research projects and innovation patents in various languages are required. In addition, as Castaneda and Cuellar (2020) stated, the process of innovation is facilitated by knowledge sharing and innovation. Distinguishing knowledge sharing and innovation from the publications and analyzing their patterns respectively might provide a deeper understanding of regional knowledge innovation development. Text mining methods can be integrated to separate the knowledge types. In the future, more analytical indexes, such as centrality, clustering coefficient and core-edge index, can be calculated and visualized for power users to explore more patterns about knowledge innovation. Last but not least, a usability test of the interactive tool should be carried to improve its usefulness.

Notes

1. <https://dl.acm.org/ccs>
2. <https://nominatim.org>

Acknowledgements

This work is funded by the project “A Visual Computing Platform for the Industrial Innovation Environment in Yangtze River Delta” supported by the Jiangsu Industrial Technology Research Institute (JITRI). We also thank Yangtze River Delta Science Data Center, National Earth System Science Data Sharing Infrastructure, National Science & Technology Infrastructure of China (<http://nnu.geodata.cn:8008/>). We thank the two anonymous reviewers whose comments and suggestions helped improve and clarify this manuscript.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Linfang Ding  <http://orcid.org/0000-0002-3707-5845>
 Liqiu Meng  <http://orcid.org/0000-0001-8787-3418>

References

About ACM publications. Updated in 2020. Accessed on 06 2021. <https://www.acm.org/publications/about-publications>

Bach, B., E. Pietriga, and J.-D. Fekete. 2014. “Visualizing Dynamic Networks with Matrix Cubes.” in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI ’14, Association for Computing Machinery, New York, NY, USA, 877–886. doi:[10.1145/2556288.2557010](https://doi.org/10.1145/2556288.2557010).

Basole, R. C. 2016. “Topological Analysis and Visualization of Interfirm Collaboration Networks in the Electronics Industry.” *Decision Support Systems* 83: 22–31. doi:[10.1016/j.dss.2015.12.005](https://doi.org/10.1016/j.dss.2015.12.005).

Billger, M., L. Thuvander, and B. S. Wästberg. 2017. “In Search of Visualization Challenges: The Development and Implementation of Visualization Tools for Supporting Dialogue in Urban Planning Processes.” *Environment and Planning B: Urban Analytics and City Science* 44 (6): 1012–1035. doi:[10.1177/0265813516657341](https://doi.org/10.1177/0265813516657341).

Borgatti, S. P., A. Mehra, D. J. Brass, and G. Labianca. 2009. “Network Analysis in the Social Sciences.” *science* 323 (5916): 892–895. doi:[10.1126/science.1165821](https://doi.org/10.1126/science.1165821).

Brauer, A., V. Mäkinen, and J. Oksanen. 2021. “Characterizing Cycling Traffic Fluency Using Big Mobile Activity Tracking Data, Computers.” *Environment and Urban Systems* 85: 101553. doi:[10.1016/j.compenvurbsys.2020.101553](https://doi.org/10.1016/j.compenvurbsys.2020.101553).

Carrillo, F. J., T. Yigitcanlar, B. Garca, and A. Lönnqvist. 2014. *Knowledge and the City: Concepts, Applications and Trends of Knowledge-based Urban Development*. New York and London: Routledge.

Castaneda, D. I., and S. Cuellar. 2020. “Knowledge Sharing and Innovation: A Systematic Review.” *Knowledge and Process Management* 27 (3): 159–173. doi:[10.1002/kpm.1637](https://doi.org/10.1002/kpm.1637).

Chen, C., F. Ibekwe-sanjuan, and J. Hou. 2010. “The Structure and Dynamics of Cocitation Clusters: A Multiple-perspective Cocitation Analysis.” *Journal of the American Society for Information Science and Technology* 61 (7): 1386–1409. doi:[10.1002/asi.21309](https://doi.org/10.1002/asi.21309).

Chen, C. 2006. “Citespace II: Detecting and Visualizing Emerging Trends and Transient Patterns in Scientific Literature.” *Journal of the American Society for Information Science and Technology* 57 (3): 359–377. doi:[10.1002/asi.20317](https://doi.org/10.1002/asi.20317).

Chen, Y., C. Ding, J. Hu, R. Chen, P. Hui, and X. Fu. 2017. “Building and Analyzing a Global Co-authorship Network Using Google Scholar Data.” in: Proceedings of the 26th International Conference on World Wide Web Companion, WWW ’17 Companion, International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 1219–1224. doi:[10.1145/3041021.3053056](https://doi.org/10.1145/3041021.3053056).

China statistical yearbook (2020). URL <http://www.stats.gov.cn/tjsj/ndsj/2020/indexch.htm>

Csomós, G., and B. Lengyel. 2020. “Mapping the Efficiency of International Scientific Collaboration between Cities Worldwide.” *Journal of Information Science* 46 (4): 575–578. doi:[10.1177/0165551519842128](https://doi.org/10.1177/0165551519842128).

Csomós, G. 2018. “A Spatial Scientometric Analysis of the Publication Output of Cities Worldwide.” *Journal of Informetrics* 12 (2): 547–566. doi:[10.1016/j.joi.2018.05.003](https://doi.org/10.1016/j.joi.2018.05.003).

Diercks, G., H. Larsen, and F. Steward. 2019. “Transformative Innovation Policy: Addressing Variety in an Emerging Policy Paradigm.” *Research Policy* 48 (4): 880–894. new Frontiers in Science, Technology and Innovation Research from SPRU’s 50th Anniversary Conference. doi:[10.1016/j.respol.2018.10.028](https://doi.org/10.1016/j.respol.2018.10.028).

Dong, Y., R. A. Johnson, and N. V. Chawla. 2016. “Can Scientific Impact Be Predicted?” *IEEE Transactions on Big Data* 2 (1): 18–30. doi:[10.1109/TBDA.2016.2521657](https://doi.org/10.1109/TBDA.2016.2521657).

Few, S. 2006. *Information Dashboard Design: The Effective Visual Communication of Data*. Vol. 2. Sebastopol, CA: O’reilly.

Fikirkoca, A., and O. Saritas. 2012. "Foresight for Science Parks: The Case of Ankara University." *Technology Analysis & Strategic Management* 24 (10): 1071–1085. doi:[10.1080/09537325.2012.723688](https://doi.org/10.1080/09537325.2012.723688).

Frenken, K., S. Hardeman, and J. Hoekman. 2009. "Spatial Scientometrics: Towards a Cumulative Research Program." *Journal of Informetrics* 3 (3): 222–232. science of Science: Conceptualizations and Models of Science. doi:[10.1016/j.joi.2009.03.005](https://doi.org/10.1016/j.joi.2009.03.005).

Friedmann, J., and L. Yong. 2004. "Planning for the Global Cities: The Endogenous Development Model." *Urban Planning Forum* 4: 1.

Galaso, P., and J. Kovářík. 2021. "Collaboration Networks, Geography and Innovation: Local and National Embeddedness." *Papers in Regional Science* 100 (2): 349–377. doi:[10.1111/pirs.12578](https://doi.org/10.1111/pirs.12578).

Håkansson, H., and A. Waluszewski. 2007. *Knowledge and Innovation in Business and Industry: The Importance of Using Others*. Vol. 5. London and New York: Routledge.

Hamdi, A., K. Shaban, A. Erradi, A. Mohamed, S. K. Rumi, and F. D. Salim. 2021. "Spatiotemporal Data Mining: A Survey on Challenges and Open Problems." *Artificial Intelligence Review* 1–48. doi:[10.1007/s10462-021-09994-y](https://doi.org/10.1007/s10462-021-09994-y).

Helbig, C., H.-S. Bauer, K. Rink, V. Wulfmeyer, M. Frank, and O. Kolditz. 2014. "Concept and Workflow for 3d Visualization of Atmospheric Data in a Virtual Reality Environment for Analytical Approaches." *Environmental Earth Sciences* 72 (10): 3767–3780. doi:[10.1007/s12665-014-3136-6](https://doi.org/10.1007/s12665-014-3136-6).

Hennemann, S., B. Derudder, and P. J. Taylor. 2015. "Cutting the Gordian Knot of Visualizing Dense Spatial Networks: The Case of the World City Network, 2013." *Environment and Planning A: Economy and Space* 47 (6): 1332–1340. doi:[10.1177/0308518X15594909](https://doi.org/10.1177/0308518X15594909).

Henriques, I. C., V. A. Sobreiro, and H. Kimura. "Science and Technology Park: Future Challenges." *Technology in Society* 53 (2018) 144–160. doi:[10.1016/j.techsoc.2018.01.009](https://doi.org/10.1016/j.techsoc.2018.01.009).

Hoekman, J., K. Frenken, and F. Van Oort. 2009. "The Geography of Collaborative Knowledge Production in Europe." *The Annals of Regional Science* 43 (3): 721–738. doi:[10.1007/s00168-008-0252-9](https://doi.org/10.1007/s00168-008-0252-9).

Imran, M., and M. Jabeen. 2020. "Visual Exploration of Scientific Literature to Formulate Research Policy: A Case of Gis Scholarly Communication in Pakistan during 2000–2019." *Journal of Geography in Higher Education* 1–20. doi:[10.1080/03098265.2020.1836480](https://doi.org/10.1080/03098265.2020.1836480).

Innocenti, N., F. Capone, and L. Lazzaretti. 2020. "Knowledge Networks and Industrial Structure for Regional Innovation: An Analysis of Patents Collaborations in Italy." *Papers in Regional Science* 99 (1): 55–72. doi:[10.1111/pirs.12478](https://doi.org/10.1111/pirs.12478).

Jenny, B., D. M. Stephen, I. Muehlenhaus, B. E. Marston, R. Sharma, E. Zhang, and H. Jenny. 2018. "Design Principles for Origin-destination Flow Maps." *Cartography and Geographic Information Science* 45 (1): 62–75. doi:[10.1080/15230406.2016.1262280](https://doi.org/10.1080/15230406.2016.1262280).

Kandt, J., and M. Batty. 2021. "Smart Cities, Big Data and Urban Policy: Towards Urban Analytics for the Long Run." *Cities* 109: 102992. doi:[10.1016/j.cities.2020.102992](https://doi.org/10.1016/j.cities.2020.102992).

Kodituwakk, H. A. D. E., A. Keller, J. Gregor, and C. Bajaj. 2020. "Insight2: A Modular Visual Analysis Platform for Network Situational Awareness in Large-scale Networks." *Electronics* 9 (10). doi:[10.3390/electronics9101747](https://doi.org/10.3390/electronics9101747).

Leite, R. A., A. Arleo, J. Sorger, T. Gschwandtner, and S. Miksch. 2020. "Hermes: Guidance-enriched Visual Analytics for Economic Network Exploration." *Visual Informatics* 4 (4): 11–22. doi:[10.1016/j.visinf.2020.09.006](https://doi.org/10.1016/j.visinf.2020.09.006).

Leydesdorff, L., L. Bornmann, and C. S. Wagner. 2017. "Generating Clustered Journal Maps: An Automated System for Hierarchical Classification." *Scientometrics* 110 (3): 1601–1614. doi:[10.1007/s11192-016-2226-5](https://doi.org/10.1007/s11192-016-2226-5).

Li, D., Y. D. Wei, and T. Wang. 2015. "Spatial and Temporal Evolution of Urban Innovation Network in China." *Habitat International* 49: 484–496. doi:[10.1016/j.habitatint.2015.05.031](https://doi.org/10.1016/j.habitatint.2015.05.031).

Li, N., J. Kramer, P. Gordon, and A. Agogino. 2018. "Co-author Network Analysis of Human-centered Design for Development." *Design Science* 4: e10. doi:[10.1017/dsj.2018.1](https://doi.org/10.1017/dsj.2018.1).

Li, W., M. Funk, J. Eikelboom, and A. Brombacher. 2020. "Visual Exploration of Movement Relatedness for Multi-species Ecology Analysis." *arXiv Preprint arXiv:2001.11163*.

Li, Y., and N. Phelps. 2018. "Megalopolis Unbound: Knowledge Collaboration and Functional Polycentricity within and beyond the Yangtze River Delta Region in China, 2014." *Urban Studies* 55 (2): 443–460. doi:[10.1177/0042098016656971](https://doi.org/10.1177/0042098016656971).

Liu, C., Q. Gui, D. Duan, and Yin, M., et al. 2017. "Structural Heterogeneity and Proximity Mechanism of Global Scientific Collaboration Network Based on Co-authored Papers." *Acta Geographica Sinica* 72 (4): 737–752.

Liu, D., K. Veeramachaneni, A. Geiger, V. O. Li, and H. Qu. 2021. "Aqeyes: Visual Analytics for Anomaly Detection and Examination of Air Quality Data." *arXiv Preprint arXiv:2103.12910*.

Liu, J., T. Tang, W. Wang, B. Xu, X. Kong, and F. Xia. 2018. "A Survey of Scholarly Data Visualization." *IEEE Access* 6: 19205–19221. doi:[10.1109/ACCESS.2018.2815030](https://doi.org/10.1109/ACCESS.2018.2815030).

Liu, X., B. Derudder, and P. Taylor. 2014. "Mapping the Evolution of Hierarchical and Regional Tendencies in the World City Network, 2000–2010." *Computers, Environment and Urban Systems* 43: 51–66. doi:[10.1016/j.compenvurbsys.2013.10.004](https://doi.org/10.1016/j.compenvurbsys.2013.10.004).

Maantay, J., and J. Ziegler. 2006. *GIS for the Urban Environment*. Redlands, CA: Esri Press.

McKenna, S., D. Staheli, C. Fulcher, and M. Meyer. 2016. "Bubblenet: A Cyber Security Dashboard for Visualizing Patterns." *Computer Graphics Forum* 35 (3): 281–290. doi:[10.1111/cgf.12904](https://doi.org/10.1111/cgf.12904).

Mercorio, F., M. Mezzanica, V. Moscato, A. Picariello, and G. Sperli. 2019. "Dico: A Graph-db Framework for Community Detection on Big Scholarly Data." *IEEE Transactions on Emerging Topics in Computing* 1–1. doi:[10.1109/TETC.2019.2952765](https://doi.org/10.1109/TETC.2019.2952765).

Piry, S., M.-P. Chapuis, B. Gauffre, J. Papaix, A. Cruaud, K. Berthier, and S. Dray. 2016. "Mapping Averaged Pairwise Information (Mapi): A New Exploratory Tool to Uncover Spatial Structure." *Methods in Ecology and Evolution* 7 (12): 1463–1475. doi:[10.1111/2041-210X.12616](https://doi.org/10.1111/2041-210X.12616).

Sapena, M., M. Wurm, H. Taubenböck, D. Tuia, and L. A. Ruiz. 2021. "Estimating Quality of Life Dimensions from Urban Spatial Pattern Metrics, Computers." *Environment and Urban Systems* 85: 101549. doi:[10.1016/j.compenvurbsys.2020.101549](https://doi.org/10.1016/j.compenvurbsys.2020.101549).

Scheibel, W., M. Trapp, D. Limberger, and J. Döllner. 2020. "A Taxonomy of Treemap Visualization Techniques." *VISIGRAPP* 27-29, Feb, 2020 3: IVAPP Malta, 273–280.

Seebacher, D., J. Häuäler, M. Hundt, M. Stein, H. Müller, U. Engelke, and D. Keim. 2018. "Visual Analysis of Spatio-temporal Event Predictions: Investigating the Spread Dynamics of Invasive Species." *IEEE Transactions on Big Data* 7 (3): 497–509.

Sobral, T., T. Galvão, and J. Borges. 2019. "Visualization of Urban Mobility Data from Intelligent Transportation Systems." *Sensors* 19 (2). doi:10.3390/s19020332.

Tabassum, S., F. S. F. Pereira, S. Fernandes, and J. Gama. 2018. "Social Network Analysis: An Overview." *WIREs Data Mining and Knowledge Discovery* 8 (5): e1256. doi:10.1002/widm.1256.

van Steen, M. 2010. *Graph Theory and Complex Networks: An Introduction*. Pisa, Italy: Maarten van Steen. URL <https://www.distributed-systems.net/index.php/books/gtcn/>

Wang, T., H. Kroll, L. Wang, and X. Zheng. 2021. "How S&T Connectivity Supports Innovation-driven Development: An Analysis of China's Cooperation Networks in High and New Technology Fields." *Asian Journal of Technology Innovation* 29 (1): 134–157. doi:10.1080/19761597.2020.1792783.

Wang, Z., D. Ma, R. Pang, F. Xie, J. Zhang, and D. Sun. 2020. "Research Progress and Development Trend of Social Media Big Data (Smbd): Knowledge Mapping Analysis Based on Citespace." *ISPRS International Journal of Geo-Information* 9 (11). doi:10.3390/ijgi9110632.

Wu, C., Y. D. Wei, X. Huang, and B. Chen. 2017. "Economic Transition, Spatial Development and Urban Land Use Efficiency in the Yangtze River Delta, China." *Habitat International* 63: 67–78. doi:10.1016/j.habitatint.2017.03.012.

Wu, H., Z. He, and J. Gong. 2010. "A Virtual Globe-based 3d Visualization and Interactive Framework for Public Participation in Urban Planning Processes, Computers." *Environment and Urban Systems* 34 (4): 291–298. geospatial Cyberinfrastructure. doi:10.1016/j.compenvurbsys.2009.12.001.

Xia, F., W. Wang, T. M. Bekele, and H. Liu. 2017. "Big Scholarly Data: A Survey." *IEEE Transactions on Big Data* 3 (1): 18–35. doi:10.1109/TBDA.2016.2641460.

Yan, C., and H. Li. 2020. "Knowledge Mapping of Virtual Academic Communities: A Bibliometric Study Using Visual Analysis." *Advances in Science, Technology and Engineering Systems Journal* 5 (5): 493–499.

Yang, Y., and H. Yang. 2008. "Complex Network-based Time Series Analysis." *Physica A: Statistical Mechanics and Its Applications* 387 (5): 1381–1386. doi:10.1016/j.physa.2007.10.055.

Ye, C., F. Zhang, L. Mu, Y. Gao, and Y. Liu. 2021. "Urban Function Recognition by Integrating Social Media and Street-level Imagery." *Environment and Planning B: Urban Analytics and City Science* 2399808320935467. doi:10.1177/2399808320935467.

Ye, C., J. Zhu, S. Li, S. Yang, and M. Chen. 2019. "Assessment and Analysis of Regional Economic Collaborative Development within an Urban Agglomeration: Yangtze River Delta as a Case Study." *Habitat International* 83: 20–29. doi:10.1016/j.habitatint.2018.10.010.

Zhao, G., M. Yang, F. A. Montiel Ishino, C. Zhang, and F. Williams. 2020. "Urban Population Distribution Mapping with Multisource Geospatial Data Based on Zonal Strategy." *ISPRS International Journal of Geo-Information* 9 (11). doi:10.3390/ijgi9110654.

Zhao, M., K. Wu, X. Liu, and D. Ben. 2015. "A Novel Method for Approximating Intercity Networks: An Empirical Comparison for Validating the City Networks in Two Chinese City-regions." *Journal of Geographical Sciences* 25 (3): 337–354. doi:10.1007/s11442-015-1172-0.

Zuo, C., L. Ding, and L. Meng. 2020. "A Feasibility Study of Map-based Dashboard for Spatiotemporal Knowledge Acquisition and Analysis." *ISPRS International Journal of Geo-Information* 9 (11): 636. doi:10.3390/ijgi9110636.