ST-Map: an interactive map for discovering spatial and temporal patterns in bibliographic data

Chenyu Zuo, Yifan Xu, Linfang Ding, Liqiu Meng

Angaben zur Veröffentlichung / Publication details:

Zuo, Chenyu, Yifan Xu, Linfang Ding, and Liqiu Meng. 2024. "ST-Map: an interactive map for discovering spatial and temporal patterns in bibliographic data." Journal of Geodesy and Geoinformation Science 7 (1): 3–15. https://doi.org/10.11947/j.JGGS.2024.0102.

Journal of Geodesy and Geoinformation Science
http://jggs.chinasmp.com

Article Open Access

ST-Map: an Interactive Map for Discovering Spatial and Temporal Patterns in Bibliographic Data

ZUO Chenyu[1,2], XU Yifan[2], DING Lingfang[3], MENG Liqiu[2]

1 ETH Zurich, Zurich 8050, Switzerland;
2 Chair of Cartography and Visual Analgtics, TU Munich, Munich 80333, Germany;
3 Norwegian University of Science and Technology, Trondheim 7491, Norway

ABSTRACT

Getting insight into the spatiotemporal distribution patterns of knowledge innovation is receiving increasing attention from policymakers and economic research organizations. Many studies use bibliometric data to analyze the popularity of certain research topics, well-adopted methodologies, influential authors, and the interrelationships among research disciplines. However, the visual exploration of the patterns of research topics with an emphasis on their spatial and temporal distribution remains challenging. This study combined a Space-Time Cube (STC) and a 3D glyph to represent the complex multivariate bibliographic data. We further implemented a visual design by developing an interactive interface. The effectiveness, understandability, and engagement of ST-Map are evaluated by seven experts in geovisualization. The results suggest that it is promising to use three-dimensional visualization to show the overview and on-demand details on a single screen.

Key words: space-time cube; bibliographic data; spatiotemporal analysis; user study; interactive map

Citation: ZUO Chenyu, XU Yifan, DING Lingfang, et al. ST-Map: an Interactive Map for Discovering Spatial and Temporal Patterns in Bibliographic Data[J]. Journal of Geodesy and Geoinformation Science, 2024, 7(1): 3-15.

1 Introduction

Economic growth theory suggests that geographical knowledge spillovers and regional growth are strongly correlated[1-2]. There is a growing need for policymakers and economic research organizations to understand the spatial and temporal distribution of knowledge innovation[3]. For instance, Csomós, Vida, and Lengyel[4] analyzed high-impact collaborations among cities with the aim to reveal and finally reduce the geographic barriers; Literatures [5–6] demonstrated that the knowledge on regional innovation configurations and its intrinsic networks and linkages with industry is the foundation for regional agglomeration analysis. Moreover, the spatial distribution and temporal development of the knowledge landscape are crucial for understanding economic dynamics[7]. In particular, geographic clusters in specific industries are important for policy-making, such as launching science parks and industry zones[8-9].

Bibliographic data is frequently applied to analyze and derive knowledge innovation patterns. It contains rich metadata on scientific publications, such as title, author, publication time, and publication venue. This information is helpful to, e.g., analyze the literature landscape and identify future research opportunities[10], discover the leading universities in certain research domains[11], and analyze the relationship between knowledge innovation and business services over time[12]. The affiliation information in the bibliographic data can be used to derive the geographic information that is important for understanding the geospatial as-

* First author: ZUO Chenyu
E-mail: chenyu.zuo@csfm.ethz.ch

pects of knowledge innovation patterns. For example, Nascimento, Ávila, Taranto, and Kurozawa[13] used the geographic location of the authors to analyze the research collaborations between 60 counties over 40 years. In addition, identifying and categorizing research topics help to understand what was studied to bring an overview of studies from an aggregated view Kleminski, Kazienko, and Kajdanowicz[14]. Further studies Kemeny and Storper[15], Bottazzi and Gragnolati[16], Guido Buenstorf and Medrano[17] analyzed the temporal trends and the regional agglomeration patterns on multiple geographic scales, with different industrial domains on aggregated and detailed levels.

Besides the aforementioned work that mainly uses statistical methods for analyzing bibliographic data, many studies[18-19] employ various visualizations to represent the body of literature, trends in specific research topics, and citation networks. For example, line charts are used to show the temporal trends of cited times, histograms are used to compare the cited literature from different fields, and network graphs are widely used to show the author cooperation networks and co-citation networks[20]. Moreover, the visual representation of the geospatial distribution patterns of research domains is widely studied to understand geographic clusters of knowledge production. However, efficient visualization tools to represent the spatiotemporal patterns of the structure of research domains in both holistic and detailed views are still largely missing. Designing and developing such visualization tools face several challenges. First, it is hard to categorize the semantic information of research topics and embed them into the geographic space. Second, the temporal development of spatial patterns is often shown in a series of maps, but these separate views are hard for readers to memorize and compare the complex research structures. Third, bibliometric data is often big and complex. When visualizing these high-dimensional data, well-designed user interactions as a fundamental component are highly needed but currently often lacked to support effective data exploration and knowledge discovery. Therefore, it is highly needed to design visualization methods that can address these issues.

In this study, we proposed an interactive map ST-Map that enables the exploration of spatial, temporal, and semantic patterns in bibliographic data that reflects knowledge innovation. More specifically, we focused on the visual analysis of the variables related to knowledge domains and their spatiotemporal distribution. We adopted the space-time cube visualization method and designed a 3D symbol to encode the aggregated research domains of bibliographic data on a single screen. We further evaluated the effectiveness, understandability, and engagement of ST-Map with a set of benchmark tasks and objective questions. The findings suggest that a space-time cube with embedded symbols is a promising visualization method to provide a compact view to identifying the patterns of multivariate data.

2 Related Work

2.1 Bibliometric data analysis

Bibliometrics are the study of academic publishing that use statistics to measure the scientific production (in quantity, quality, and impact) on various topics, journals, authors, and countries[21-22]. Bibliometric data usually includes metadata of scholarly publications, such as title, author, keywords, journal, publish time, citation, and abstract. With the aid of data mining technologies, insightful information of scientific production can be revealed from bibliometric data, such as network analysis, text mining, time series analysis, and spatial analysis methods. Among them, constructing and analyzing networks of co-authorships, citations, and co-occurrences can unpack intellectual structures in terms of major scholars and topics[23-25]. Many studies focus on developing statistical methods to discover the evolution of fields, knowledge gaps, knowledge clusters, networks of research fields, and social processes supporting knowledge development[26]. Moreover, analyzing bibliometric data can have social contributions in terms of discovering knowledge clusters and social geographical patterns. Some spatial aggregation and clustering methods are used to reveal such pattern[27-28].

Recent studies adopt various visualization methods to ensure that complex patterns of bibliometric data are easily understandable[29]. Bibliometric visualization is a field that empowers people to understand the body of literature with computational methods and visual analytical methods[30]. Many studies proposed interactive visualization methods to allow interpretation of the body of literatures, especially the reference linkages, in diagrams from bibliographic data. For example, some interactive tools are developed to show the emerging trends and transient patterns in scientific literatures[31], and timelines are used to show the temporal development of selected fields[32].

However, many advanced bibliometric visualizations were mainly proposed to discover research temporal trends and to find correlations between literature networks[30]. Little research was focused on designing geovisualizations to reveal spatial bibliometric distributions. In a previous study, Zuo et al.[33] developed a map-based dashboard to show the collaborations among cities, regions, and countries in different research domains. They highlighted the geospatial distribution of co-authorship networks. However, they showed the research domains in an aggregated view with a treemap, the regional bibliographic configuration remain hidden from the viewers.

2.2 Visual analysis of spatiotemporal data

Representing spatial and temporal characteristics of data can guide readers to acquire an overall structure. Aigner et al.[34] summarised 158 visualization methods for representing time-oriented. Based on their study, we further categorized the visual metaphors of time-oriented spatial data into five groups, including symbols[35], layers[36], angle[37], height[38], and abstract location[39].

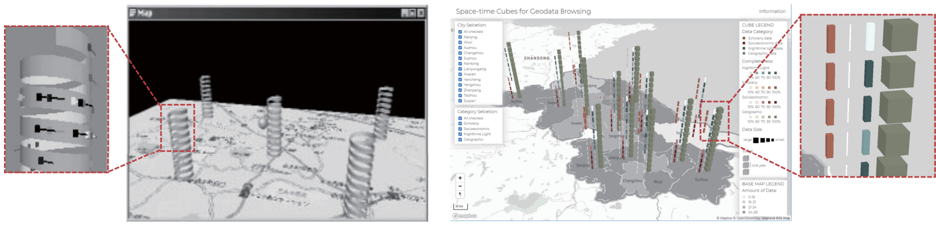
Despite many existing visualization methods, identifying data characteristics by comparing multiple map views is difficult[40-41]. It is more efficient for readers to compare the spatial and temporal patterns while they are all presented in a single map view. STC is a widely adopted visualization method that can map spatiotemporal data in a single 3D view, which is beneficial for readers to acquire an overall spatiotemporal structure[42].

STC was first proposed by Il agcrstrand[43], which extends 2D maps to 3D showing time on the z-axis. It compresses the spatial and temporal information in one single scene. STCs are increasingly used to reveal spatiotemporal patterns in diverse datasets, for example, Andrienko et al.[44] superimposed 250 trajectories of cars to show where the cars were suspended, and Peters et al.[45] adopted STC to visually analyze the spatiotemporal clustering of lightning data.

In an STC, a 3D symbol can be designed and embedded to represent future data variables. Hewagamage et al.[46] proposed a spiral STC map to show periodical spatiotemporal patterns of human activities. They represented an activity as a line, time period of activity as line length, activity category as color, and data file as icon attached to the 3D spirals, as shown in Fig. 1(a). However, it is hard for viewers to obtain an overall pattern of the represented activities because of occlusion. Zuo et al.[47] adopt the space-time cube method, as shown in Fig. 1(a), to provide a holistic overview and to enable visual queries of geo-metadata, such as geospatial coverage, topic information, data size, data type, and data quality. Applying different colors and sizes on dispersed 3D symbols to represent the data attributes reduced occlusion. We identified that the dispersed 3D symbols have more potential to represent data variables.

3 Data Description and Preprocessing

We used a dataset collected from the ACM Digital Library as our test data. The ACM Digital Library (see in https://dl.acm.org/) is a community-based repository related to computing research and practice. It collected articles from journals, mag-



(a) (b)

Figure 1 The spiral representation[46] and the cubic representation[47] in STC

Table 1 An example of metadata of an article

Item	Data
Title	Deep Supervised Quantization by Self-Organizing Map
Authors	WANG Min, ZHOU Wengang, TIAN Qi, PU Junfu, LI Houqiang
Affiliations	University of Science and Technology of China, Hefei, China; University of Science and Technology of China, Hefei, China; University of Texas at San Antonio, San Antonio, TX, USA; University of Science and Technology of China, Hefei, China; University of Science and Technology of China, Hefei, China
Published time	October 2017
Index terms	Computing methodologies, machine learning, learning paradigms, supervised learning, supervised learning by classification, information systems, information retrieval, retrieval models and ranking, top-k retrieval in databases, retrieval tasks and goals, clustering and classification
URL	`https://dl.acm.org/doi/10.1145/3283289.3283360`

azines, proceedings, and books. We collected the metadata of the articles from 1st January 2017 to 30th December 2020 by using its API. In this study, we selected China as the study area. It includes 23 385 published articles from 2017 to 2020, 48 808 authors, and 5387 institutes, accounting for 17.0%, 16.0%, and 13.8% of the globe, respectively.

The collected metadata includes multiple attributes, i.e., title, names of the authors, author affiliations, published time, index terms, and the URLs of the articles in the ACM Digital Library. The index term is a computing classification system used to categorize the keywords or research domains of the articles in the ACM Digital Library. Tab. 1 shows an example of the collected metadata of an article.

The data preprocessing in this study is composed of three steps, including georeferencing the affiliations, categorizing the index terms, and geographic aggregation. The georeferencing converts affiliations into geographic coordinates. We used the open-source service Nominatim (see in https://nominatim.org/) to obtain the longitudes and latitudes of the affiliations. Regarding the categorization of the index terms, we adopt the widely used ACM Computing Classification System (see in https://dl.acm.org/ccs) as a semantic vocabulary to reflect computing disciplines. This system has a hierarchical structure to describe the computer science related domains. It has 2112 index terms in 13 categories, and each category is further divided into sub-categories with a maximum of nine levels of depth. Fig. 2 shows the poly-hierarchical structure of the example article (as shown in Tab. 1) in four levels. The categorized index terms are finally aggregated to the provincial level to discover the regional differences. We chose the provincial level because the research institutes are mostly located in provincial capitals. To effectively reveal the spatial distribution, temporal evolution, and multivari-

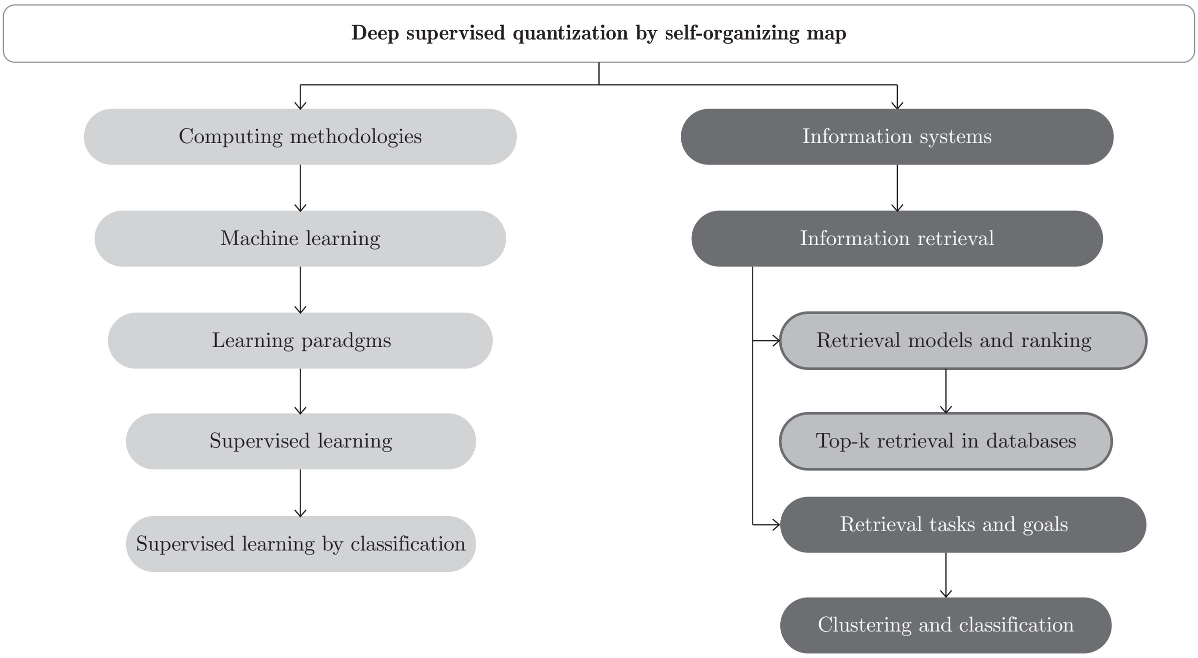


Figure 2 An example of the poly-hierarchical structure of the example article in Tab. 1

ate patterns, we aggregate the number of articles based on the 13 categories and 84 subcategories of the ACM Computing Classification System, the location of the author affiliations are aggregated to the province level, and the publication time is aggregated on a yearly basis.

4 Visualization Design Method

In this study, we design an interactive map that can show the spatial and temporal pattern of bibliometric data, represent the research domain into their natural categories, and allow viewers to filter data on their demands. As mentioned in Section 2, STC has the advantage of giving an overview of spatial and temporal distributions, we therefore adopt this method to show the distribution of bibliographic data. In addition, we integrate cartographic design methods to generalize and symbolize data. This help users to gain an holistic view and focus on important patterns without overwhelmed by too many details. This section introduces the symbol design to represent the hierarchical structure of the publications and the interactive map design.

4.1 Symbol design

Bibliographical data can be used to extract the domain information, which is labeled by index terms in the ACM digital library. The index terms can be aggregated into a hierarchical structure to outline the trend of specific domains. Moreover, revealing connections and composition of research domains is important in bibliometric analysis.

We designed a three-dimensional symbol to represent the hierarchical structure of the index terms on an aggregated level. Fig. 3 shows the symbol

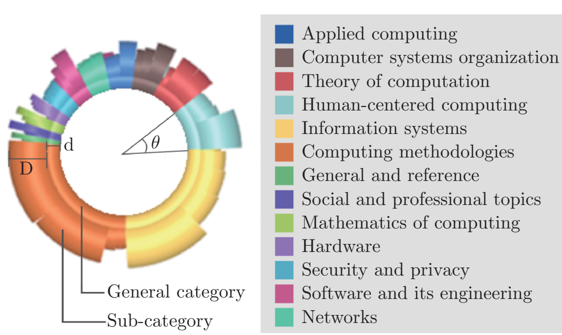


Figure 3 A three-dimensional symbol from a top view and the legend of categories

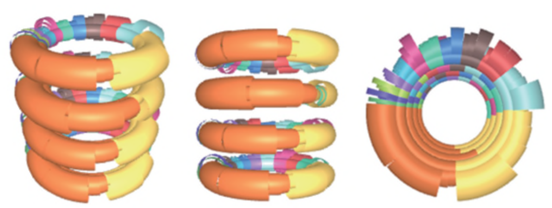


Figure 4 A design of space-time cube symbols to show the temporal patterns from a bird view, horizontal view, and top view

from a top view. We chose a cyclic symbol to reveal periodic structures. The whole symbol represents the number of publications with the author affiliations within a certain geographic area, such as a city or a region, and the legend of the categories. The symbol is composed of an inner ring and an outer ring, representing the category and sub-category of index terms. Each category is represented by one color as the legend shows. Although some colors might look similar, the users can hover over the arcs to trigger the pop-up windows to read the names of the categories.

$$\theta_i = 2\pi \times p_i \tag{1}$$

$$D_i = d \times \left(1 + \frac{f_j}{F_j}\right) \tag{2}$$

Where the central angle θ is the proportion p of a general category i; D_i is the width of the outer torus; i represents a sub-category; f represents the frequency of a sub-category; F represents the frequency of a general category. The width of the inner torus d is a constant.

To show the temporal patterns, we used the z-axis to represent the time dimension. Fig. 4 shows the vertically stacked three-dimensional symbol from a bird view, horizontal view, and top view respectively. Each ring represents the aggregated articles in one year.

4.2 Interactive interface design

We created the ST-Map in a 3D scene for users to interact. Users are allowed to explore the space-time cube by rotating, zooming, panning, selecting, filtering, clicking, and changing the viewing perspectives. The symbols show the aggregated patterns and users can retrieve more details through interactions.

Fig. 5 shows the ST-Map interface, which consists of six panels. Panel A is a location and time selector that allows users to select single or multiple regions and time periods to be shown on the map. Panel B allows users to highlight one region and learn the temporal patterns of the semantic patterns of index terms. The 3D symbol rotates horizontally to show the viewers the index terms on different positions of the ring. Panel C allows users to further select one specific time period and one index term. In ad-

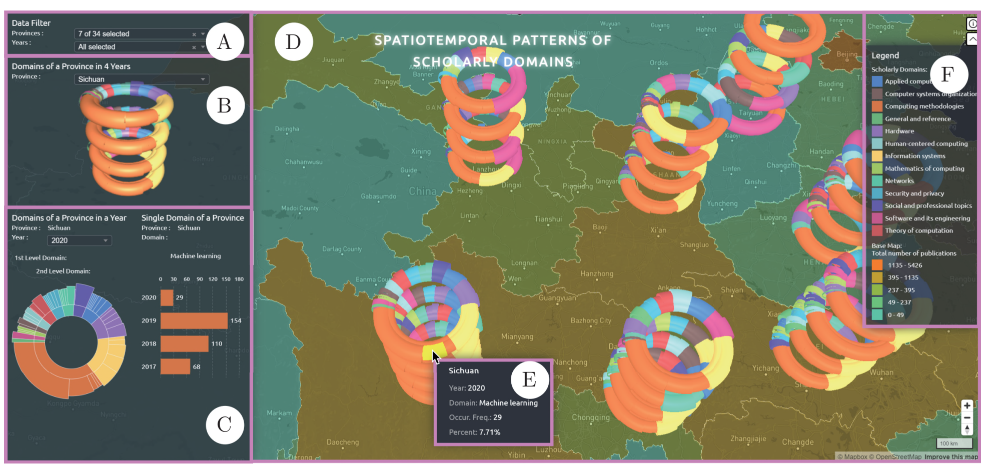


Figure 5 The visual interface of ST-Map with multiple interactive and interlinked panels

dition, it highlights the proportion of categories in a 2D donut chart and the temporal trend of a selected category in a bar chart. Panel D is a map view with embedded STC. The map color shows the number of publications in each province. The 3D rings show the proportion and number of publications in various categories. Moreover, users are allowed to zoom and rotate the map scene. Panel E is a pop-up window with detailed textual information whenever a mouse hovers over the symbols in Panel A, B, C, and D. Panel F is an information panel that provides usage tutorials on clicking the ⓘ icon at the top-right corner and the legend of the map and the symbols. Panel A, B, and C are control panels that can be hidden so that the map can be resized, zoomed, and rotated. Last but not least, the placement of the panels follows the order from general to detail in Panels A, B, and C from top to bottom. It can guide users to learn the data patterns from a general context to detailed values.

The interactive interface is implemented with open-source libraries. We used Python to process data, PostgreSQL to store data, and JavaScript for interface development. The map is developed with Leaflet, the 3D scene is developed with three.js, and the symbols are implemented with D3. The interactive map is developed can be opened with a modern web browser, such as Google Chrome, at https://gracexu108.github.io/Spatiotemporal-Patterns-of-Scholarly-Domains/dist.

5 User Study

ST-Map is implemented as a prototype to test the potential of STC in showing complex spatiotemporal data. We therefore conducted a user study to have a more in-depth understanding of the user experience rather than usability. Our user study focused on finding out how users read bibliometric information from the interactive STC and how users regard the insights they have obtained from the visual interface. Such deep reflections require the participants to have a professional background in geovisualization design. To guide the participants through ST-Map in a standard manner, we designed an ST-Map usage tutorial and four benchmark tasks. We then ask eight questions on understandability and engagement to the participants. Seven domain experts were carefully selected to participate in this survey. Their age range was between 24 to 36 with an average of 28. The participants were composed of four females and three males.

5.1 User survey design

This user study is conducted via an online survey of the ST-Map. The participants were invited via email and were recommended to perform the survey with a laptop or desktop but not a mobile device.

The introduction part of the survey includes a short textual description of the whole procedure and a tutorial on the interface. The four benchmark tasks are predefined to test the interface in empowering users to acquire insights from bibliographic data. The four tasks require users to acquire certain information from the interface, and they are designed with increasing information complexity. To conduct the benchmark tasks, the participants were required to verify four statements by exploring the interface. There was no time limitation for the participants to perform the tasks. Tab. 2 lists the statements and the corresponding operations that can verify them. They were asked to verify the statements and mark them with one of the four options: True, False, Not Given, or Not Sure.

After performing the benchmark tasks, the survey measured the user experience in two types, i.e., understandability and engagement, as shown in Tab. 3. The participants can rate these two measures on five scales from Strongly Agree to Strongly Disagree. Following each rating, the participants were asked to write the reasons in detail.

5.2 Results

We analyzed the survey results in three parts. The effectiveness of the symbol and map design for insight acquisition is analyzed via the calculation of the success rate of the benchmark tasks. The

Table 2 The four statements and operations on the interface

Item	Statement	Operation
T1	In 2017, the proportion of the domain "Computer Systems Organization" in Beijing was 3.81%.	It can be verified by first selecting the year and location in Panel A, and then checking the pop-up window by hovering the ring symbols either in Panel B, Panel C, or Panel D.
T2	The occurrence frequency of the domain "Applied Computing" in Shanghai increased yearly from 2017 to 2020.	It can be verified by first selecting the location and then checking the width of the arcs either in Panel B or D.
T3	In 2020, the domain "Information Systems" was more popular in the western provinces than elsewhere among the eight default provinces, i.e., Beijing, Shanghai, Jiangsu, Guangdong, Gansu, Yunnan, Hubei, and Heilongjiang.	It can be verified by first selecting the eight places and then comparing the arcs.
T4	In the foreseeable future, among Anhui, Jiangsu, Shanghai, and Zhejiang, the domain "Computing Methodologies" is more likely to take the dominant position than "Theory of Computation".	It can be verified by first selecting the four places and then summarizing the trend of the arcs.

Table 3 The metrics of understandability (U1-U4) and engagement (E1-E4)

Item	Statement
U1	I can get an overview of the data theme and the spatiotemporal distributions at first glance.
U2	I can understand the meaning of the 4-ring-stack symbol of a province.
U3	I can quickly find the interaction tool that I need.
U4	I think the layout of the interface is easy to understand.
E1	I think the interface visual design is visually appealing.
E2	I think interacting with the interface is entertaining.
E3	I trust the interface when doing the tasks.
E4	I am confident with my answers to the tasks.

understandability and engagement of the interface were analyzed through the rating of their metrics respectively.

Fig. 6 shows the success rate of the benchmark tasks. Except for Task 3, the success rate of benchmark tasks was high. The success rate from Task 1 to Task 4 does not show a decreasing trend as the information complexity increases. It might indicate that the symbol design and map design can in general support spatiotemporal insight acquisition. We did not receive the answer "Not Given" from any participants, which suggests that the participants believed the statements could be verified by using the ST-Map interface. Specifically, Task 1 required the participants to identify the value of a certain place at a certain time. Task 2 required the participants to identify the temporal trend of an attribute of a certain place. Both Task 1 and Task 2 were completed with a relatively high success rate. It might indicate that the interface can effectively support the users in finding values and identifying semantic and temporal patterns. Task 3 asked the participants to summarize a spatiotemporal distribution of attributes in multiple places. Task 4 asked the participants to predict and compare the spatiotemporal trend of an attribute in multiple places. Both Task 3 and Task 4 involved multiple spaces and time periods, however, the success rate of Task 3 was much lower than Task 4. One possible explanation is that the mental effort of Task 3 was much

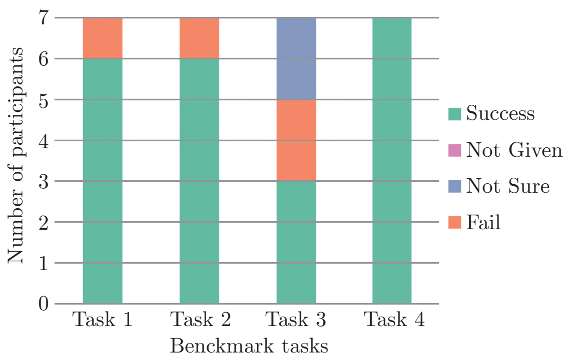


Figure 6 The success rate of the benchmark tasks

higher than Task 4 in practice. Because Task 3 involved eight places, whereas Task 4 requires four places. Two participants reported with "Not Sure" to Task 3 which indicates that the participants had difficulties conducting this task. Another possible explanation is that the participants learned how to identify the complex spatiotemporal patterns from Task 3, which might lead to a better performance in Task 4.

We examined four aspects that related to understandability, i.e., spatiotemporal patterns (U1), three-dimensional symbols (U2), user interaction (U3), and the layout of the interface (U4) as shown in Tab. 3. Fig. 7 shows the rating of the four metrics from the participants. The results suggest that applying 3D symbols in space-time cubes can help people acquire an overview of a complex dataset (U1). Surprisingly, the stacked ring symbol can be understood easily as shown by all the seven participants who agreed with U2. In addition, the user interactions, including selection, mouse hovering, clicking, zooming, and panning, are easy to understand as suggested by the results of U3. However, as reported by U4, the layout is not optimal to be understood easily. Furthermore, in-depth answers from the participants help the authors to understand better the design issues. In summary, the STC as a visualization method using the z-axis to represent time is understandable. The layout of the interface provides contextual information for users to connect the visual symbols to bibliographic data. In ST-Map, the layout should be improved according to the users' feedback.

Four metrics related to user engagement were asked (as Tab. 3 shows), i.e., visual attractiveness (E1), entertainment (E2), trustworthiness (E3), and confidence (E4). In addition to the ratings, the participants were asked to explain their answers in detail. Fig. 8 shows the rating results from the survey. E1 shows that six participants thought the visual design was appealing. They reported the reason as, "the interface is clean, the color palette is appealing, and the arrangement of visualizations is natural from overview to detail." One of the participants disagreed and this person stated as "the dark background color caused uncomfortableness to the eyes and prefer light background color like grey." The results of E2 show five participants held positive on the entertainment of the ST-Map. The reasons were stated as "much valuable information is provided and the multidimensional spatiotemporal information can be perceived with a few clean operations." Two participants held negative impressions because they experienced that the interaction

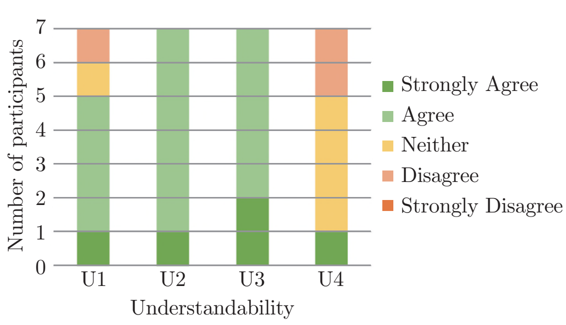


Figure 7 The results of the user understandability

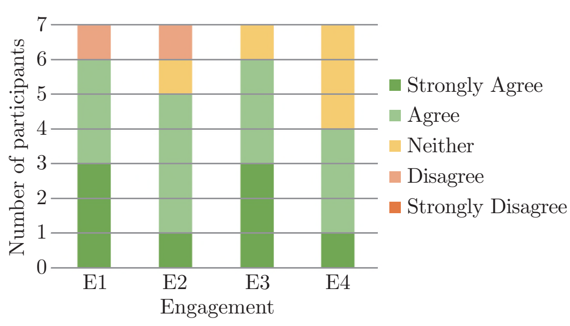


Figure 8 The results of the engagement

with the map was slow and hard to find the cities. We found that the support for knowledge discovery increased entertainment, while the difficulty of user interactions (like hover-over) reduced the entertainment level. It might be sometimes difficult to hover over a small arc of the 3D symbol. It suggests that error-free user interactions do not necessarily contribute to user entertainment, but deficient interactions decrease user entertainment. The results of E3 show that the trustworthiness of the interface is in general good, with six participants trusting and one having a neutral opinion. The people who held positive opinions have stated that "the interface gave me much data support and the information the interface presents is consistent with my perception." The person reported neutral because "I only used part of the interface and 'I don't know if I should trust the interface since I do not have any (bibliometric analysis) background knowledge.' ". The comments show that trustworthiness is gained through using the interface to acquire meaningful insights, rather than simply viewing visual elements (like map style, color, and text). The results of E4 show that the confidence level is lower than the other metrics. The positive comments are "I found all the graphic clues, I can exactly find the answers to the questions, and the interface is present in a simple way." The reasons that lead to neutral opinions are "I am not sure if I understood the task correctly and hard to understand Task 3." It might indicate that the high mental effort can overwhelm the users and decrease their confidence in the interface.

6 Discussion

This study designed the ST-Map that incorporates a space-time cube with coordinated views, which allows an effective exploration of multivariate bibliographic data. ST-Map shows the hierarchical structure and composition of the literature along with their spatiotemporal distributions. In particular, the 3D stacked ring symbol embedded in a space-time cube encoded the spatial distribution and temporal trend of research domains along with the hierarchical structure and proportions of research fields. Compared to a previous study that provided spatial information, proportion, and hierarchical information in separate views[33], ST-Map represents the bibliographic data in a compact view, avoids excessive visual clutter, and provides detailed information on demand. The usability and user experience are evaluated by a preliminary user study with benchmark tasks and in-depth questions. The results show that encoding spatiotemporal multivariate data in a 3D environment can guide users to gain an overview, find patterns, and compare values easily. Although the authors assumed the complexity of the 3D visualization might overwhelm the users, which is partly confirmed by the response of U4. Nevertheless, the participants can understand the symbol and find the needed information by completing three of the benchmark tasks and strongly agree with U1, U2, and U3.

However, there is some potential space to lift the usability of ST-Map. Occlusion is a common issue in 3D visualization because some objects are hidden by those in the foreground, which might lead to a loss of the overview. ST-Map provides user interactions to reduce occlusion. More user interactions could be introduced to reduce the occlusion. For example, changing the unselected arcs to semi-transparent colors. Moreover, some of the arcs with the same color are not vertically aligned, which might cause difficulties for users to compare them. Although users can use the bar chart in Panel C to compare the temporal trend of a selected index, the 3D symbols should allow users to realign the arcs. In addition, the symbols could utilize more visual variables to represent even more data features, for example, using the radius of the ring to show the number of articles, allowing users to view the monthly trend in addition to the yearly trend, extending the 3D symbol to represent the nine-level depth of the index terms, and providing different aggregation methods for users to explore spatial patterns. In the future, the data preparation module of the ST-Map should provide more user-friendly functions, such as allowing users to upload their own datasets, supplying various aggregation methods, and ordering methods for location searching.

We evaluate this work using ST-Map as a prototype to show the potential of the space-time cube method. It is important to obtain the first round of feedback from a small group of experts, their valuable feedback will serve as suggestions for future improvement. After the improvement, a user study on understandability and engagement will be carried out with a larger group.

7 Conclusion and Outlook

This study explores the possibility of extending the space-time cube with embedded symbols to represent the semantic information of multivariate bibliographic data. ST-Map shows the spatial and temporal distribution of multiple research domains and their interrelations. The use of interactive visual interfaces brings understanding and engagement in knowledge discovery in bibliographic data. The method could be extended to show the overview of patterns of knowledge spillover, entrepreneurship, economy, market, funding, and their linkages. Supporting visual pattern discovery is at the interface of information visualization, data science, human-computer interaction, cartography, and cognition science. Proposing ST-Map helps us to identify the potential of the space-time cube and expose some of the design issues. Especially, we conducted a preliminary user study with in-depth reflections of seven domain experts. It shows that the STC method is promising for discovering complex geospatial insights, although finding the 3D scene requires a high mental effort at first sight.

This study shows the potential of extending visual analytical methods to a 3D environment, which can give the visualization more viewing perspectives and interactions to encode and elaborate the multivariate data. Future experiments could explore more 3D symbols, such as 3D radar charts or 3D trees, to represent multivariate spatiotemporal data. However, the challenge remains in designing sophisticated visualization methods to show complex datasets while ensuring good understandability and engagement. Furthermore, this study exposes that user experience is more related to users' prior knowledge and usability, rather than the visual design style. A large-scale user study will be carried out to investigate how visual design can better support multivariate data exploration.

Acknowledgments

The authors acknowledge all the participants in the experiment. We are grateful for their valuable contribution.

References

[1] GROSSMAN G M, HELPMAN E. Trade, knowledge spillovers, and growth[J]. European Economic Review, 1991, 35(2-3): 517-526.

[2] DÖRING T, SCHNELLENBACH J. What do we know about geographical knowledge spillovers and regional growth: A survey of the literature[J]. Regional Studies, 2006, 40(3): 375-395.

[3] CSOMÓS G. On the challenges ahead of spatial scientometrics focusing on the city level[J]. Aslib Journal of Information Management, 2020, 72(1): 67-87.

[4] CSOMÓS G, VIDA Z V, LENGYEL B. Exploring the changing geographical pattern of international scientific collaborations through the prism of cities[J]. PLoS One, 2020, 15(11): e0242468.

[5] BORT S, OEHME M, ZOCK F. Regional networks, alliance portfolio configuration, and innovation performance[M]//Understanding the Relationship Between Networks and Technology, Creativity and Innovation. Leeds: Emerald Group Publishing Limited, 2014: 229-256.

[6] ORLANDO M, VERBA M, WEILER S. Universities, agglomeration, and regional innovation[J]. Review of Regional Studies, 2019, 49(3): 407-427.

[7] AUTANT-BERNARD C, LESAGE J P. Quantifying knowledge spillovers using spatial econometric models[J]. Journal of Regional Science, 2011, 51(3): 471-496. DOI: 10.1111/j.1467-9787.2010.00705.x.

[8] HOLL A, PETERS B, RAMMER C. Local knowledge spillovers and innovation persistence of firms[J]. Economics of Innovation and New Technology, 2023, 32(6): 826-850. DOI: 10.1080/10438599.2022.2036609.

[9] BUZARD K, CARLINO G A, HUNT R M, et al. Localized knowledge spillovers: evidence from the spatial clustering of R&D labs and patent citations[J]. Regional Science and Urban Economics, 2020, 81: 103490.

[10] TIBERIUS V, SCHWARZER H, ROIG-DOBÓN S. Radical innovations: between established knowledge and future research opportunities[J]. Journal of Innovation & Knowledge, 2021, 6(3): 145-153.

[11] CANCINO C A, MERIGÓ J M, CORONADO F C. A bibliometric analysis of leading universities in innovation research[J]. Journal of Innovation & Knowledge, 2017, 2(3): 106-124.

[12] J-FIGUEIREDO R, NETO J V, QUELHAS O L G, et al. Knowledge Intensive Business Services (KIBS): bibliometric analysis and their different behaviors in the scientific literature: topic 16–innovation and services[J]. RAI Revista de Administração e Inovação, 2017, 14(3): 216-225.

[13] NASCIMENTO R F, ÁVILA M F, TARANTO O P, et al. Agglomeration in fluidized bed: bibliometric analysis, a review, and future perspectives[J]. Powder Technology, 2022, 406: 117597.

[14] KLEMINSKI R, KAZIENKO P, KAJDANOWICZ T. Analysis of direct citation, co-citation and bibliographic coupling in scientific topic identification[J]. Journal of Information Science, 2022, 48(3): 349-373. DOI: 10.1177/0165551520962775.

[15] KEMENY T, STORPER M. Is specialization good for regional economic development?[J]. Regional Studies, 2015, 49(6): 1003-1018.

[16] BOTTAZZI G, GRAGNOLATI U. Cities and clusters: economy-wide and sector-specific effects in corporate location[J]. Regional Studies, 2015, 49(1): 113-129.

[17] BUENSTORF G, FRITSCH M, MEDRANO L F. Regional knowledge, organizational capabilities and the emergence of the west German laser systems industry, 1975—2005[J]. Regional Studies, 2015, 49(1): 59-75.

[18] GUI Qinchang, DU Debin, LIU Chengliang. The changing geography of scientific knowledge production: evidence from the metropolitan area level[J]. Applied Spatial Analysis and Policy, 2023, 17(1): 157–174.

[19] SUN Huaping, EDZIAH B K, KPORSU A K, et al. Energy efficiency: the role of technological innovation and knowledge spillover[J]. Technological Forecasting and Social Change, 2021, 167: 120659.

[20] ZHAO Yang, WANG Lin, ZHANG Yaming. Research thematic and emerging trends of contextual cues: a bibliometrics and visualization approach[J]. Library Hi Tech, 2021, 39(2): 462-487.

[21] NINKOV A, FRANK J R, MAGGIO L A. Bibliometrics: methods for studying academic publishing[J]. Perspectives on Medical Education, 2022, 11(3): 173-176.

[22] DONTHU N, KUMAR S, MUKHERJEE D, et al. How to conduct a bibliometric analysis: an overview and guidelines[J]. Journal of business research, 2021, 133: 285-296.

[23] MOLONTAY R, NAGY M. Twenty years of network science: a bibliographic and co-authorship network analysis[M]//ÇAKIRTAŞ M, OZDEMIR M K. Big Data and Social Media Analytics: Trending Applications. Cham: Springer, 2021: 1-24. DOI: 10.1007/978-3-030-67044-3_1.

[24] HEVEY D. Network analysis: a brief overview and tutorial[J]. Health Psychology and Behavioral Medicine, 2018, 6(1): 301-328. DOI: 10.1080/21642850.2018.1521283.

[25] LOZANO S, CALZADA-INFANTE L, ADENSO-DÍAZ B, et al. Complex network analysis of keywords co-occurrence in the recent efficiency analysis literature[J]. Scientometrics, 2019, 120(2): 609-629.

[26] LIM W M, KUMAR S. Guidelines for interpreting the results of bibliometric analysis: a sensemaking approach[J]. Global Business and Organizational Excellence, 2024, 43(2): 17-26.

[27] ZHUANG Liang, YE Chao, LIESKE S N. Intertwining globality and locality: bibliometric analysis based on the top geography annual conferences in America and China[J]. Scientometrics, 2020, 122(2): 1075-1096.

[28] IBÁÑEZ J J, BREVIK E C, CERDÀ A. Geodiversity and geoheritage: detecting scientific and geographic biases and gaps through a bibliometric study[J]. Science of The Total Environment, 2019, 659: 1032-1044.

[29] KRAKER P, SCHRAMM M, KITTEL C, et al. Openknowledgemaps/Headstart: Headstart 7[EB/OL]. (2021-11-25)[2024-01-02]. https://doi.org/10.5281/zenodo.5726914.

[30] CHEN Chaomei, SONG Min. Visualizing a field of research: A methodology of systematic scientometric reviews[J]. PLoS One, 14(10): e0223994. DOI: 10.1371/journal.pone.0223994.

[31] CHEN Chaomei. Citespace II: detecting and visualizing emerging trends and transient patterns in scientific literature[J]. Journal of the American Society for Information Science and Technology, 2006, 57(3): 359-377.

[32] SEBASTIAN Y, SIEW E G, ORIMAYE S O.

Emerging approaches in literature-based discovery: techniques and performance review[J]. The Knowledge Engineering Review, 2017, 32: e12.

[33] ZUO Chenyu, DING Linfang, YANG Zhuo, et al. Multiscale geovisual analysis of knowledge innovation patterns using big scholarly data[J]. Annals of GIS, 2022, 28(2): 197-212.

[34] AIGNER W, MIKSCH S, SCHUMANN H, et al. Guiding the selection of visualization techniques[M]//AIGNER W, MIKSCH S, SCHUMANN H, et al. Visualization of Time-Oriented Data. London: Springer, 2023: 193-210. DOI: 10.1007/978-1-4471-7527-8_7.

[35] MCNABB L, LARAMEE R S. Multivariate maps—a glyph-placement algorithm to support multivariate geospatial visualization[J]. Information, 2019, 10(10): 302.

[36] WEI Zhonghui, GU Xiaohe, SUN Qian, et al. Analysis of the spatial and temporal pattern of changes in abandoned farmland based on long time series of remote sensing data[J]. Remote Sensing, 2021, 13(13): 2549.

[37] ZHU Yingzhen, YU Jifang, WU Jiangqin. Chro-Ring: a time-oriented visual approach to represent writer's history[J]. The Visual Computer, 2016, 32(9): 1133-1149.

[38] ANDRIENKO G, ANDRIENKO N, SCHUMANN H, et al. Visualization of trajectory attributes in space–time cube and trajectory wall[M]//BUCHROITHNER M, PRECHTEL N, BURGHARDT D. Cartography from Pole to Pole. Berlin, Heidelberg: Springer, 2014: 157-163. DOI: 10.1007/978-3-642-32618-9_11.

[39] BU Chuan, ZHANG Quanjie, WANG Qianwen, et al. Sinestream: improving the readability of streamgraphs by minimizing sine illusion effects[J]. IEEE Transactions on Visualization and Computer Graphics, 2021, 27(2): 1634-1643.

[40] GRIFFIN A L, MACEACHREN A M, HARDISTY F, et al. A comparison of animated maps with static small-multiple maps for visually identifying space-time clusters[J]. Annals of the Association of American Geographers, 2006, 96(4): 740-753. DOI: 10.1111/j.1467-8306.2006.00514.x.

[41] BOYANDIN I, BERTINI E, LALANNE D. A qualitative study on the exploration of temporal changes in flow maps with animation and small-multiples[J]. Computer Graphics Forum, 2012, 31(3pt2): 1005-1014. DOI: 10.1111/j.1467-8659.2012.03093.x.

[42] KRISTENSSON P O, DAHLBACK N, ANUNDI D, et al. An evaluation of space time cube representation of spatiotemporal patterns[J]. IEEE Transactions on Visualization and Computer Graphics, 2009, 15(4): 696-702.

[43] HÄGERSTRAAND T. What about people in regional science?[J]. Papers in Regional Science, 1970, 24(1): 7-21.

[44] ANDRIENKO N, ANDRIENKO G, RINZIVILLO S. Exploiting spatial abstraction in predictive analytics of vehicle traffic[J]. ISPRS International Journal of Geo-Information, 2015, 4(2): 591-606.

[45] PETERS S, BETZ H D, MENG Liqiu. Visual analysis of lightning data using space–time-cube[M]//BUCHROITHNER M, PRECHTEL N, BURGHARDT D. Cartography from Pole to Pole. Berlin, Heidelberg: Springer, 2014: 165-176.

[46] HEWAGAMAGE K P, HIRAKAWA M, ICHIKAWA T. Interactive visualization of spatiotemporal patterns using spirals on a geographical map[C]//Proceedings of 1999 IEEE Symposium on Visual Languages. Tokyo, Japan: IEEE, 1999: 296-303.

[47] ZUO Chenyu, GAO Mengyao, DING Linfang, et al. Space-time cube for visual queries over metadata of heterogeneous geodata[J]. KN-Journal of Cartography and Geographic Information, 2022, 72(1): 29-39.