TUM School of Engineering and Design
Technische Universität München

Map-based Dashboard for Social Environment Understanding

Chenyu Zuo

Vollständiger Abdruck der von der TUM School of Engineering and Design der Technischen Universität München zur Erlangung einer

Doktorin der Ingenieurwissenschaften (Dr.-Ing.)

genehmigten Dissertation.

Vorsitz:
Prof. Dr.ir. Walter de Vries

Prüfer*innen der Dissertation:
1. Prof. Dr.-Ing. Liqiu Meng
2. Assoc. Prof. Dr.-Ing. Linfang Ding
3. Prof. Dr.-Ing. Jochen Schiewe

Die Dissertation wurde am 14.06.2022 bei der Technischen Universität München eingereicht und durch die TUM School of Engineering and Design am 30.09.2022 angenommen.

Abstract

Social environments refer to physical and cultural surroundings, such as geographical environment, socioeconomic environment, and industrial environment. The various environments share the characteristics of space dependency and involve multiple factors with complex correlations. A holistic understanding of social environments in combination with an analytical understanding of specific aspects is a prerequisite for individuals and organizations, such as enterprises and citizens, to make their strategic and tactic decisions. The availability of a large amount of big data in recent years provides new opportunities to uncover the deep structural patterns of social environments.

However, gaining insights into the social environment from big data is very challenging. One reason is that the collection, selection, and transformation of data require not only technical skills, but also domain knowledge. Another reason is that interpreting the data is time-consuming, whereas stakeholders have limited time to understand social environments.

This thesis addresses the design and use of map-based dashboards which should support stakeholders to learn and analyze complex social environments in a short time. A map-based dashboard as a visual communication tool can present an overview of multiple factors and their spatial distributions, and enable users to explore the data to acquire insights. To design effective, efficient, and easy-to-use dashboards, the author proposed a design framework with five components - design goals, users' cognitive tasks, data, visual interface, and users' feedback. By involving users in the design process, the usability and user experience of dashboards can be significantly improved.

The proposed framework is demonstrated and evaluated using three case studies with three different map-based dashboards. These case studies cover different topics from industrial innovation environment, socioeconomic environment, to knowledge innovation environment. The studies involve heterogeneous datasets like statistical data and social network data. Based on dedicated analysis of the design goals, each map-based dashboard is designed to allow an easy perception and understanding of the complex spatiotemporal information from those data. Users' feedback from different perspectives, including 1) effectiveness and efficiency of insight acquisition, 2) visual attraction, 3) users' insight acquiring strategy, and 4) users' attitude, are collected and incorporated into the design.

In summary, this thesis enriched the visual language in learning and analyzing social environments. More specifically, the carefully designed layout connect the visualizations to jointly show an overview and sequentially guide users to acquire insights. The case studies show that except for considering the nature of social environments, users' expectation of information and their experience should be integrated into the design process to reach a better understanding of social environments.

Zusammenfassung

Das soziale Umfeld bezieht sich auf die physische und kulturelle Umgebung, wie das geografische Umfeld, das sozioökonomische Umfeld und das industrielle Umfeld. Die verschiedenen Umwelten haben die Merkmale der Raumabhängigkeit gemeinsam und umfassen zahlreiche Faktoren mit komplexen Zusammenhängen. Ein ganzheitliches Verständnis des sozialen Umfelds in Kombination mit einem analytischen Verständnis spezifischer Aspekte ist eine Voraussetzung für Einzelpersonen und Organisationen wie Unternehmen und Bürger, um strategische und taktische Entscheidungen zu treffen. Die Verfügbarkeit einer großen Menge von Big Data in den letzten Jahren bietet neue Möglichkeiten, die tiefgreifenden strukturellen Muster des sozialen Umfelds aufzudecken.

Die Gewinnung von Erkenntnissen über das soziale Umfeld aus Big Data ist jedoch eine große Herausforderung. Ein Grund dafür ist, dass die Sammlung, Auswahl und Umwandlung von Daten nicht nur technische Fähigkeiten, sondern auch Fachwissen erfordert. Ein weiterer Grund ist, dass die Interpretation der Daten zeitaufwändig ist, während die Akteure nur wenig Zeit haben, um das soziale Umfeld zu verstehen.

Diese Arbeit befasst sich mit der Entwicklung und dem Einsatz von kartenbasierten Dashboards, die es Stakeholdern ermöglichen sollen, komplexe soziale Umfelder in kurzer Zeit zu erlernen und zu analysieren. Ein kartenbasiertes Dashboard als visuelles Kommunikationsinstrument kann einen Überblick über mehrere Faktoren und ihre räumliche Verteilung geben und es den Nutzern ermöglichen, die Daten zu erkunden, um Erkenntnisse zu gewinnen. Zur Gestaltung effektiver, effizienter und benutzerfreundlicher Dashboards schlug der Autor einen Gestaltungsrahmen mit fünf Komponenten vor: Gestaltungsziele, kognitive Aufgaben der Nutzer, Daten, visuelle Schnittstelle und Nutzerfeedback. Durch die Einbeziehung der Benutzer in den Designprozess können die Benutzerfreundlichkeit und die Benutzererfahrung von Dashboards erheblich verbessert werden.

Der vorgeschlagene Rahmen wird anhand von drei Fallstudien mit drei verschiedenen kartenbasierten Dashboards demonstriert und bewertet. Diese Fall-

studien decken verschiedene Themen ab, von der industriellen Innovationsumgebung über die sozioökonomische Umgebung bis hin zur Wissensinnovationsumgebung. Die Studien beinhalten heterogene Datensätze wie statistische Daten und Daten sozialer Netzwerke. Auf der Grundlage einer gezielten Analyse der Designziele wurde jedes kartenbasierte Dashboard so gestaltet, dass es eine einfache Wahrnehmung und ein einfaches Verständnis der komplexen raum-zeitlichen Informationen aus diesen Daten ermöglicht. Das Feedback der Nutzer aus verschiedenen Perspektiven, darunter 1) Effektivität und Effizienz der Informationsgewinnung, 2) visuelle Attraktivität, 3) die Strategie der Nutzer zur Informationsgewinnung und 4) die Einstellung der Nutzer, wurde gesammelt und in das Design integriert.

Zusammenfassend lässt sich sagen, dass diese Arbeit die visuelle Sprache beim Lernen und Analysieren von sozialen Umgebungen bereichert hat. Genauer gesagt, verbindet das sorgfältig entworfene Layout die Visualisierungen, um gemeinsam einen Überblick zu zeigen und den Lesefluss der Nutzer zu leiten, um Erkenntnisse zu gewinnen. Die Fallstudien zeigen, dass Designer von kartenbasierten Dashboards nicht nur die Natur sozialer Umgebungen berücksichtigen, sondern auch die Erwartungen der Nutzer an Informationen und ihre Erfahrungen in den Designprozess integrieren sollten, um ein besseres Verständnis sozialer Umgebungen zu erreichen.

Acknowledgments

I am deeply grateful to my primary advisor, Professor Liqiu Meng, for her unwavering support, expert guidance, and through-provoking discussions throughout the entire doctoral research process. Her visionary thinking and dedication to pushing the boundaries of knowledge has been a constant source of inspiration and motivation. She supports me in exploring new scientific topics, encourages me to extend my research by teaching and supervising students, allows me to grow my organizational skills by including me in industrial projects, grant writing, and event organization. All the experiences under her supervision has become my valuable assets.

I also extend my heartfelt thanks to my mentor and advisor Professor Linfang Ding. Her brought me into the field of Geovisual Analytics, guiding me to translate the pure curiosity into research achievements. I deeply admire her passion on cartography and commitment to science. During countless hours of mentorship, her valuable insights always guide me in finding solutions. She generously offering me weekly discussion slots even when she was a young mother, and we met every Monday 10 pm via online meetings after her baby slept. Her valuable advice, suggestions and comments made a great influence in the completion of this work.

I am grateful to thesis committee member Professor Jochen Schiewe and Professor Walter de Vries. Their rigorous reviews, valuable suggestions, and insightful critiques have played a pivotal role in refining this work. Their insights not only contributed to the depth and quality of this research, but also motivates me to think deeper and formulate future research topics.

My academic journey would have been significantly more challenging without the collaborative spirit within the Chair of Cartography, which enriched my research experience. I would like to thank Ekaterina Chuprikova, Holger Kumke, Bing Liu, Andreas Divanis, Edyta Bogucka, Juliane Cron, Hao Lyu, Ruoxin Zhu, Mathias Jahnke, Christian Murphy, Peng Luo, Shengkai Wang, Stephanie Kruchen, and Fleißner Luise. They offered valuable insights and unwavering encouragement during brainstorming sessions and countless coffee breaks.

I also wish to express my appreciation to all the research participants who generously shared their time, experiences, and insights. Their contributions were instrumental in the success of this study. I acknowledge the broader cartography community, which continuously inspires and challenges us to think critically and contribute to the body of knowledge. This thesis stands on the shoulders of giants who have paved the way in this field.

Finally, I owe a profound debt of gratitude to my friends and family, whose belief in my abilities sustained me through the highs and lows of this academic pursuit. They have been my pillars of strength, always encouraging me to persevere in the face of challenges. Their support have made this journey possible.

Contents

List of Figures

List of Tables

List of abbreviations

3D Three Dimensional

ACM Association for Computing Machinery

AOI Areas of Interest

BI Business Information

DG Design Goal

GDP Gross Domestic Product

IIE Industrial Innovation Environment

KPI Key Performance Indicators

R&D Research & Development

YRD Yangtze River Delta

URL Uniform Resource Locator

UTF Unicode Transformation Format

Disclaimer

I confirm that this doctoral thesis is my own work and I have documented all sources and material used.

Munich, 10.06.2021 Chenyu Zuo

1 Introduction

This chapter provides a general introduction to this dissertation. First, the background and the motivation for this research are explained. Four research questions follow in Section 1.2 and five research tasks are outlined in Section 1.3. Finally, the structure of this thesis is described in Section 1.4.

1.1 Motivation

Social environment refers to the physical and cultural surroundings that are influenced by humans. It contains natural surroundings, infrastructure, social relationships, cultural practices, human mobility, health, etc., which have direct and indirect spatial components. For example, a knowledge innovation environment focuses on the intellectual assets, professional personnel, research institutes, and collaboration networks [Tid01], the study of which may involves complex analysis of these actors at multiple spatial and temporal scales. A thorough understanding of a specific type of social environment, especially from a spatial perspective, will help stakeholders, such as enterprises, general public, policy-makers, researchers, and domain experts, to make and evaluate critical decisions (e.g., relocate the production [AA16]) and refine the existing strategies (e.g., increase the investment [SBY+21]).

The emergence of big data brings new opportunities for stakeholders to understand social environments. For instance, census data, social surveys, remote sensing data, trajectories of transportation, and social network data can reflect social environments from different aspects. However, due to the complex data structure, understanding social environments remains effortful. The stakeholders have to take multiple factors of the complicated environments into consideration and need to understand an environment from different perspectives, for instance, identify key factors, learning the key values, finding the spatial patterns, compare factors at multiple spatial scales. Computational and data mining methods have been employed to extract meaningful information from big data, such as to reveal the dynamics of spatiotemporal distribution [BHBPW13], detect social events [ZQP+15], and identify opinions on social events [QZX16]. Moreover, advanced geostatistical methods (e.g., sampling, correlation, and prediction) are developed to discover the spatial structure and relationships [WO07].

Graphical representations provide an access to acquire insights into big data of social environments. Interactive visualizations that take advantage of human vision power and computational capabilities have been increasingly employed to facilitate the understanding of big data [ZSB+12]. With the inherent nature of thematic, spatial, and temporal characteristics, exploring big social environment data requires advanced tools to facilitate spatial analyses. According to the needs of users, a visual representation should be well-designed to e.g., give an overview of the key factors of a studied social environment and its spatiotemporal distribution patterns, and also allow the exploration of detailed information regarding users' need.

Dashboard is one of the most common tools in data visualization, and is described as "a visual display of the most important information needed to achieve one or more objectives, consolidated and arranged on a single screen so the information can be monitored at a glance" [Few06]. In recent years, dashboards are expanding from simple monitoring to analytical guides for users [Tor70]. Facing the challenges posed by the complexity of understanding social environments, we identify map-based dashboard as a powerful tool to effectively communicate complex data at-a-glance, intuitively explain spatial information, as well as enable users to interactively analyze their required information. Moreover, this thesis on how to design map-based dashboards to facilitate learning and analyzing social environments.

This thesis is dedicated to proposing a framework of map-based dashboard design. In this framework, we first formulate design goals according to the potential users' needs and identify the corresponding cognitive tasks in understanding social environments. This is especially important for a successful design. Then we design data and visual interface. Finally, we collect users' feedback to improve visual interface and examine the design goals. We evaluate this framework using case studies of three representative social environments, including industrial innovation, socioeconomic, and knowledge innovation environments.

1.2 Research Questions

Driven by the increasing need for analytical visual interfaces and the disparate multi-factor data about social environments, this thesis aims to enrich the visual language for spatial information learning and analysis. The main research questions (RQ) are identified and summarized as follows:

> **RQ1**: Why is map-based dashboard selected to support the understanding of complex social environments?
>
> **RQ2**: What are the important elements in map-based dashboards that facilitate users in acquiring knowledge from large amount of heterogenous data?
>
> **RQ3**: How can a map-based dashboard be designed that considers users' demands, perceptions, and data characteristics?
>
> **RQ4**: How can the proposed design framework be implemented and evaluated for selected case studies?

1.3 Research Tasks

This thesis addresses multi-disciplinary research that combines data science, cartography, and human-computer interaction. The research tasks are described as follows:

- To identify information needs, users' experience, and the corresponding cognitive tasks for understanding social environments.
- To approach multiple data sources and transform them into easy understandable factors at appropriate spatial levels.
- To propose appropriate visualizations and interactions methods that can allow users to communicate and analyze social environments with a spatial focus.
- To design layout of visual interfaces that can provide an overview and a logical reading flow.
- To evaluate the usability and user experience of the proposed visual interfaces.

1.4 Thesis Structure

Following this introductory chapter, this thesis contains five further chapters.

Chapter 2 describes the foundation of visual perception and sense-making in social environment understanding. It analyzes the trend of using big data to understand social environments, introduces the use of graphicacy in representing big data, outlines the human cognition in visually understanding spatial data, and summarizes the visual analytical methods in communicating complex environmental information to the general public.

Chapter 3 introduces dashboard as an visual analytical design approach. This chapter first introduces the concept and different types of dashboards, then lays out the characteristics of map-based dashboard for analysis purposes, next outlines the visual elements that used in map-based dashboard, analyzes various evaluation methods of dashboards, and finally identifies the challenges in designing map-based dashboard.

Chapter 4 proposes a design framework of map-based dashboard for social environment understanding, including design goals, users' cognitive tasks, data, interface, and users' feedback. This chapter introduces these five components in five sections, respectively.

Chapter 5 evaluates the proposed framework in three case studies. The three case studies includes three map-based dashboards that are designed to understand industrial innovation, socioeconomic, and knowledge innovation environments. Moreover, the case studies covers various datasets, visual designs, and user feedback.

Chapter 6 concludes the thesis, describes the research transferability, and discusses the future work.

All perceiving is also thinking,
all reasoning is also intuition,
all observation is also invention.

R. Arnheim, *Art and Visual Perception*, 1954

2 Foundations for Acquiring Knowledge into Social Environment

This chapter introduces the background of using visual analytical methods to understand social environments with big data. Section 2.1 introduces the social environment and the studies using big data to reveal social environments. Section 2.2 describes the foundation of graphicacy and discusses on the sense-making of big data using visual representations. Section 2.3 outlines human cognition in understanding the visualized social environments. Section 2.4 describes the studies using visual analytical methods to support learning and reasoning social environment.

2.1 Social Environment and Big Data

Social environment is a complex system that takes geographic location as a foundation and includes human activities. Barnett and Casper [BC01] defined social environment as:

> Human social environments encompass the immediate physical surroundings, social relationships, and cultural milieus within which defined groups of people function and interact. Components of the social environment

> include built infrastructure; industrial and occupational structure; labor markets; social and economic processes; wealth; social, human, and health services; power relations; government; race relations; social inequality; cultural practices; the arts; religious institutions and practices; and beliefs about place and community.

For instance, social industrial environment describes the setting of the enterprise for industrial services, involving information such as natural resources, population, technology, and investment. People study this environment to understand, rethink, and adjust the social, economic, and environmental conditions [UTS10]. For example, there are studies conducted to improve the gender balance [DS09], manage financial services for social enterprises [SB11], and promote regional industrial innovation [LB18]. The typical stakeholders concerning societal issues are citizens, organizations, newspapers, researchers, and governments [RGH$^+$17]. Based on this finding, the author lists four types of typical individuals and organizations who are interested in social industrial environment and describes their characteristics and their information needs.

Decision-makers in enterprises. They often require holistic information from different aspects and certain information on different levels of detail. They usually have limited time to collect and analyze information. Moreover, they rely on trustworthy information sources.

General public. They require information on various topics for diverse purposes, such as career planning, education selection, and entertaining. The statistics of various factors are often needed. They focus mainly on a local scale.

Domain experts. They have strong background knowledge in social environments and require in-depth understanding. They are specialized in certain factors with their temporal dynamics. In addition, they may need to retrieve the original data and conduct further analysis.

Students. They require information on various topics at different levels of details. They have less time restrictions and might be interested not only on the processed data but also the processing methods.

An increasing number of studies have collected various data to support the understanding of social environments. We selected some recent studies that used different data sources, including traditional stationary data, statistical data, and social media data, as Table 2.1 shows. The topics cover a wide range, containing the natural environment such as water bodies and air quality, human movement such as traffic and migration, and cultural environments such as research network and sport.

Table 2.1: A few selected studies about social environment.

Study	Time	Environment types	Topic	Data type
[LCL21]	2021	Nature	Water bathymetric	Airborne Lidar data
[SHH+21]	2018	Nature	Invasive species	Animal observation records and meteorological data
[QLR+20]	2019	Nature	Air quality	Stationary air quality records
[SMP+19]	2018	Nature	Urban heat island	Stationary weather data
[JZHA22]	2022	Human dynamics	Autonomous driving	Video data
[DK20]	2020	Human dynamics	Flight connections	Origin-destination of flight data
[SHL+21]	2021	Human dynamics	Human trajectories	Mobile phone location data
[LLW+20]	2020	Human dynamics	Human trajectories	Trajectory data of athletes
[ZWL+21]	2020	Human dynamics	Population migration	Social media textual data
[BWAL20]	2020	Human activity	Sport	FIFA world cup data
[CLCY20]	2020	Human activity	Reposting behaviors	Social media textual data
[ZLJ+19]	2018	Human activity	Social events	Social media textual data
[PMSR18]	2017	Human activity	Social hot topics	Social media textual data
[GCML06]	2006	Human activity	Business environment	Company data
[KGF+22]	2022	Human activity	Tourism trends	Internet search data
[ZSC+21]	2021	Human activity	Relief supplies	Social media textual data
[LCZ+19]	2019	Socioeconomic	Socioeconomic	Economic statistical data
[LXR19]	2019	Socioeconomic	Socioeconomic	Socioeconomic statistical data
[KK22]	2022	Socioeconomic	Social network	Family tree data
[SCS21]	2021	Socioeconomic	Research network	Research grands data
[SBv+21]	2021	Public health	Medical condition	Social media textual data
[ML21]	2021	Public health	Pandemic	Statistics from public agencies
[MSFM16]	2016	Security	Cyber security	Cyber attack data

From Table 2.1, we can see that the various complex social environments can be reflected and derived from various data. Each social environment could include a wide range of topics and multiple scales. *Nature environment* involves many topics such as water body [LCL21], air quality [QLR+20], temperature [SMP+19], and animal locations [SHH+21]. In addition, the temporal changes of the factors can be measured by different intervals, such as hourly, daily, seasonally, and yearly. *Human dynamics environment* could be reflected by the flight connections [DK20] and human mobile trajectories [SHL+21], and the scale varies from individual persons to populations from a large area. *Human activity environment* contains a large range of activities but not limited to sports [LLW+20], social events [ZLJ+19], business [GCML06], tourism [KGF+22], and relief supplies [ZSC+21]. The spatial and temporal range of these topics varies from each other, and each of them includes multiple variables. For example, sport event and relief supplies happen in a relatively small temporal and spatial range, while social events and tourism are analyzed in a larger spatiotemporal range. *Socioeconomic environment* contains multiple perspectives from social and economic environments at an aggregated level [LCZ+19, LXR19] to institutes [SCS21] and families [KK22] at an individual level. *Public health environment* and *Security environment* also require analysis at different scales, such as from individuals [ML21] to regional [SBv+21, MSFM16].

Various datasets are used to analyze social environments. Traditionally, statistical data [LCZ+19, LXR19] is widely used to analyze human activities and socioeconomic environments, and stationary data [QLR+20] is commonly used in nature environment monitoring. Various sensing data can contribute to the social environments' observation. For example, the earth surface can be recorded as images and videos [JZHA22]. The location and movement of humans are recorded by sensors [SHL+21], such as their mobile phones [SHL+21]. In addition, social network data enriched the analysis of social environments from the perspective of human reactions. The emotion [ZLJ+19], connections [KK22], and individual experience [ML21] could be more revealed from the social network data. Meanwhile, the richness of data types increased the difficulty of data collection and processing. The large amount of textual and imagery data requires specific processing methods and techniques.

2.2 Graphicacy in Big Data

Information can be communicated using texts, numbers, and graphics. The corresponding skills are termed as literacy, numeracy, and graphicacy. Graphicacy, as a visual language, relies on the visuo-spatial aspect of human intelligence of interpreting maps, diagrams, and photography [Bal76]. These visual presentations can support humans in harnessing a large amount of information from the real world [CMS99]. Wilkinson summarized the taxonomy of visual attributes in four aspects as Table [Wil12] shows, and the visual attributes include form, surface, motion, and text.

Table 2.2: Visual attributes [Wil12].

Form	Surface	Motion	Text
position	color	direction	label
size	hue	speed	legend
shape	brightness	acceleration	title
polygon	saturation		data source
glyph	texture		designer
image	pattern		usage tutorial
rotation	granularity		
resolution	orientation		
	blur		
	transparency		

The graphical languages are further summarized in six categories [LD05]. Each category uses multiple visual attributes to encode information in different complexities. Table 2.3 lists the graphical languages in detail.

Table 2.3: The graphical languages [LD05].

Graphical languages	Examples	Encoding technique
Axis languages	Number line, scale	A single-position encodes information by the placement of a mark on an axis.
Apposed-position languages	Line chart, bar chart, plot chart	Information is encoded by a marked set that is positioned between two axes.
Retinal-list languages	Graphics featuring colour, shape, size, saturation, texture, orientation	Retinal properties are used to encode information. These marks are not dependent on position.
Map languages	Road map, topographic map	Information is encoded through the spatial location of the marks.
Connection languages	Tree, acyclic graph, network	Information is encoded by a set of node objects with a set of link objects.
Miscellaneous languages	Pie chart, venn diagram	Information is encoded with additional graphical techniques (e.g., angle, containment).

In addition, the visual language is enriched with interactions. We summarize the interactive operations applied in previous works at three levels, including individual visualization operations such as zooming and panning, data operations such as setting the focus on a subset or more details, and interface operations such as resetting the page and linked views. Table 2.4 shows details of the interactions.

Using visual languages, readers can perceive *insights*, which is defined as an individual observation of the data by the participant, a unit of discovery in [SND05]. These insights can be described as complex, deep, qualitative, unexpected, and relevant [Nor06, PFG08]. Some examples of insight could be "The two cities are geographically close to each other", "The most companies are located in the northeast of the city", and "The number of students is increasing".

Based on the characteristics of insight, Saraiya et al. [SND05] proposed eight parameters to measure insight as observation, time, domain value, hypothesis, directed versus unexpected, correctness, breadth versus depth, and category.

Observation. The actual finding of the data value.

Time. The amount of time taken to gain the insight.

Domain Value. The value, importance, or significance of the insight.

Table 2.4: The analysis of interactions used in interactive maps.

Study	Year	Graphical operation	Data operation	Interface operation
[Shn03]	1996	Overview, zoom, detail-on-demand, relate, extract	Filter	History
[DE98]	1998	Highlighting and focus, zooming and fish-eye view, change graph, animation	Drill down and hyperlinks, selection	
[Kei02]	2002	Interactive zooming, interactive distortion	Interactive projection, interactive filtering	Interactive linking and brushing
[Wil12]	2005	Zooming, panning, lens, node dragging, highlighting, animating, rotating, transforming	Filtering, categorical reordering	Linking
[YKSJ07]	2007	Reconfigure, encode	Select, filter, explore, abstract/elaborate	Reconfigure
[KI13]	2013	Navigate 3D space, change visualization styles/types, choose objects/locations in 3D, select subspace/objects/-groups	Generate data value read-out	Specify/manipulate many data exploration parameters
[SWH14]	2014	Panning and zooming	Select, abstract/elaborate	Reconfigure, control widgets, brushing and linking

Hypotheses. Some insights lead users to identify a new biologically relevant hypothesis and direction of research.

Directed versus Unexpected. Directed insights answer specific questions that users want to answer. Unexpected insights are additional exploratory or serendipitous discoveries that were not specifically being searched for.

Correctness. Some insights lead to the correction of incorrect observations resulting from misinterpretation of the visualization.

Breadth versus Depth. Breadth insights present an overview of biological processes but not much detail. Depth insights are more focused and detailed.

Category. Insights are grouped into four main categories: overview (overall distributions of gene expression), patterns (identification or comparison across data attributes), groups (identification or comparison of groups of genes), and details (focused information about specific genes).

2.3 Human Visual Cognition in Social Environment Context

Understanding visual cognition can help us to design better visual interfaces. Patterson et al. provided an overview of human cognition in brain science [PBG+14]. They mentioned that human visual cognition is derived from a dual-system in the combination of bottom-up autonomous intuition that receives retina stimulus and top-down pattern recognition and analytical reasoning. Specifically, they presented a set of six leverage points that can be exploited by visualization designers in order to measurably influence certain aspects of human cognition: (1) exogenous attention; (2) endogenous attention; (3) chunking; (4) reasoning with mental models; (5) analogical reasoning; and (6) implicit learning as shown in Figure 2.1.

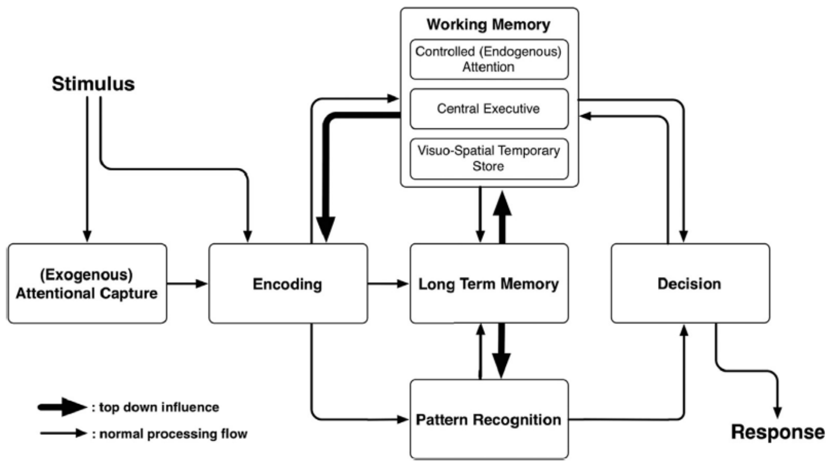


Figure 2.1: The overview of the information processing model mainly with encoding, working memory, long-term memory, pattern recognition, and decision components [PBG+14].

In the dual-process framework, the information flow proceeds from left to right in Figure 2.1. The external information enters the cognition system with encoding, which converts the visual image into neural representations. It is stored in a visuo-spatial short-term memory and does not process extensive cognitive tasks. The encoded information may activate further cognitive pattern-recognition. The encoding via attentional capture is related to the bottom-up feature detection process. The attention can be understood as focusing cognitive processing on selected stimulus while ignoring others [KE03]. The working memory is under conscious control, therefore the attention to visualization is influenced by the top-down mechanisms, and viewers normally can not look at a visualization without a task in mind. Pattern recognition, which means the recognition of statistical patterns from the stimulus, takes place in parallel with the working memory. The recognized patterns could trigger the long-term memory, which refers to the neural representation of information and knowledge. The long-term memory could in turn bias the recognition and have a large influence on the target selection in the encoding process by directing attention to the stimulus [WCR13]. The final stage is a decision, which means the process of selecting an option from the implicit pattern recognition. It can also influence the encoding process by guiding the working memory.

Visual cognition leads to sense-making. According to a previous study [YKSJ08], the sense-making process of information visualization is to provide an overview, adjust data, detect patterns, and match mental models. *Providing overview* can support users with a general impression of the data including the structure and characteristics, thus helping users to form further inquiries. *Adjustment* allows users to explore the data from different dimensions, subsets, and groups, that can help users to gain knowledge by reducing the working memory load and focusing better on reasoning. *Detecting patterns* means that users verify their own knowledge or discover new knowledge by identifying data features such as distributions, trends, outliers, and correlation. *Matching mental models* refers to a further step in that users link their observations with real-world knowledge to formulate deductions and hypotheses.

Baker et al. [BJB09] divided the sense-making tasks into six groups, including observing specific data points, looking for patterns or outliers, making inferences, comparing observed facts or patterns to one's own prior knowledge, generating hypotheses about the data, and drawing analogies from the context being observed to another context. Following their category, the author summarized the sense-making tasks proposed in previous studies in Table 2.5.

Furthermore, spatial sense-making and tasks have their own requirements. As Laha et al. [LKGA18] proposed, the tasks that are related to geo-environment, such as spatial understanding, navigation, and path following, are complex and require users to perform more tasks than just the interpretation of the visualization and relating it to their own knowledge and experience. Andrienko et al. [AAG03] extended the spatial cognitive tasks to spatio-temporal visualization exploration. They pointed out that using visualization techniques to explore spatio-temporal data requires an understanding of the characteristics of the data and the types of exploratory tasks. Expert knowledge is needed when employing visualization methods to explore spatio-temporal data. They summarized the tasks related to the change of spatio-temporal data over time:

Table 2.5: The sense-making tasks in six groups.

Author(s)	Years	Observing specific data points	Looking for patterns or outliers	Making inferences	Comparing observed facts or patterns to one's own knowledge	Generating hypotheses about the data	Drawing analogies from the context being observed to another context
[LBLS15]	2015	Search, shape description, segmentation, filtering, selection	Pattern recognition, spatial understanding, quantitative estimation, path following	Manipulation, view modification	Spatial understanding, navigation		Spatial understanding, navigation
[YKSJ08]	2007	Select, explore, reconfigure, elaborate, filter	Explore, encode, abstract		Encode, abstract, connect	Abstract	Connect
[AES05]	2005	Retrieve value, filter, compute derived value	Find extremum, characterize distribution, find anomalies, cluster	Sort, determine range			Connect
[ZF98]	1998	Locate, emphasize, identify, tabulate, trace, label, emphasize	Cluster, categorize, outline, emphasize, identify, plot, structure, map, label	Categorize, distinguish, emphasize, identify, rank, trace, switch	Associate, correlate, map, portray, generalize	Reveal, correlate, generalize, reveal	Associate, correlate, map, portray
[WL90]	1990	Locate, identify, describe	Cluster, description	Distinguish, categorize, compare within entities, compare between relations	Associate, correlate	Identify, correlate	Associate, correlate

The existential features. I.e., appearance and disappearance.

Changes of spatial properties. Location, shape or/and size, orientation, altitude, height, gradient, and volume.

Changes of thematic properties expressed through values of attributes. Qualitative changes and changes of ordinal or numeric characteristics (increase and decrease).

Visual analytics is a multidisciplinary field that combines methods in data science and information visualization to support the effective and efficient discovery of hidden patterns in a large amount of data [KAF+08]. Visualization provides intuitive understanding and may enhance humans in data exploration and insight discovery. Various visualizations are designed to present the data features, such as value, range, spatial clusters, and temporal changes. The embedded interactions allow users to focus on the data that is only used for sense-making at the moment. The interactive operations effectively relieve the working memory of remembering the data patterns for reasoning. Sacha et al. [SSS+14] proposed a process model for knowledge generation through visual analytical methods, as shown in Figure 2.2. They highlighted the sense-making process by five components, i.e., finding, insight, knowledge, hypothesis, and action. As such, visualizations are used to serve as a bridge between human knowledge acquisition and computer data processing.

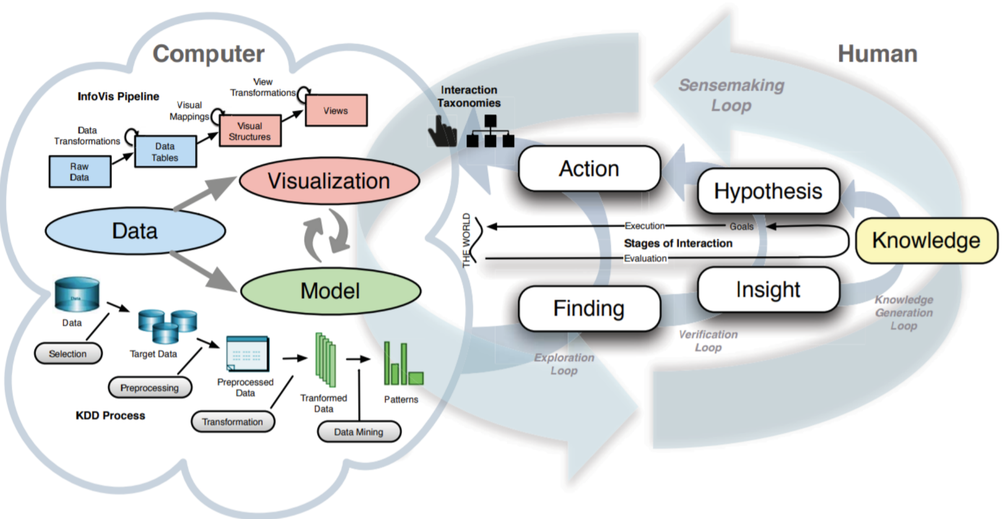


Figure 2.2: Knowledge generation model for visual analytics [SSS+14].

Users' emotions have an impact on their experience and knowledge acquisition. Intellectual work does not only occur in people's heads, but most cognition is realized by interacting with individuals, tools, eyes, ears, and computers [War19]. Except for effectiveness and efficiency, users' emotions are triggered by the visualization are part of the users' experience and affect their satisfaction, such as pleasure, fun, surprise, curiosity, exploration, aesthetics, novelty, and originality [HP12]. Especially, emotion impacts trustworthiness and credibility [MHSW19].

2.4 Visual Analysis of Social Environment

Visualizing big spatiotemporal data could guide users to an intuitive understanding of the underlining data patterns, therefore it is an important analysis tool in decision-making [HSE+22]. Many studies have been conducted to design effective, aesthetic, and innovative visualizations to support the identification of spatiotemporal patterns. The visual analysis of social environments often relies on the representations of multiple aspects using various types of visualizations, such as maps, charts, gauges, and images. We have selected a few previous examples which were designed to support social environment understanding, as shown in Table 2.6.

The analysis of previous studies show that understanding social environments requires various specific spatial data processing methods, as listed in Table 2.6. Geolocation extraction are required in many studies [SHL+21, SHL+21, HZC+20] to explore the spatial components of the social phenomena. Spatial clustering and aggregation methods are applied to analyze the social environments at different scales [BWAL20, DK20, QLR+20, LXR19, SHH+21, MSFM16, GCML06]. With the increasing use of social media data, human reaction and emotion to certain social events and their spatial distribution are studied using natural language processing methods [SBv+21, CLCY20, ZLJ+19, PMSR18]. Last but not least, statistical methods are widely used in analyzing social environments at a large spatial and temporal scale [QLR+20, LCZ+19, SMP+19, MSFM16].

Various cartographic visualizations are designed to show the spatial distributions and guide users to verify or discover knowledge. Choropleth maps [SBv+21, BWAL20, GCML06] and isarithmic maps [JZHA22] are widely used to show the spatial patterns of the aggregated single variate at a small scale. Symbol maps [LXR19, SHH+21, SMP+19, PMSR18, MSFM16] are designed to show the spatial distribution of multiple variates on an aggregated level. While dot maps [LLW+20, DK20, QLR+20, LCZ+19] are designed to show the spatial distribution of individual events and objects. Tables, charts, gauges, and other visualization methods extend the map representations, therefore they are often jointly used in many aforementioned studies. In

Table 2.6: The analysis of studies using visual analytical methods in social environment understanding.

Study	Year	Data processing methods	Visualization methods	Knowledge types
[JZHA22]	2022	Using deep learning to predict the vehicles' motion	Isarithmic map, video, bar chart, table	Identify spatial distribution, predict driving conditions, observe individual trip
[SHL+21]	2021	Using movement extraction models to extract movement trajectories and clusters	Flow map, bar chart, clock chart	Identify spatial distribution at the city level, identify spatial and temporal clusters, correlate movement with population, detect anomaly
[SBv+21]	2021	Using bio-psycho-social model to extract symptoms, drug usage, and emotions	Choropleth map, radar chart, word cloud	Discover spatial patterns
[BWAL20]	2020	Spatial and temporal aggregation	Choropleth map, tree map, bar chart	Identify spatial distribution, find extermum
[DK20]	2020	Spatial and temporal aggregation	Flow map, dot map, bar chart, line chart, scatter plot	Identify, compare, and correlate the spatiotemporal patterns
[LLW+20]	2020	Anomaly detection	Dot map, scatter plot, bar chart, area chart	Identify the anomaly of sportsman and locate them
[CLCY20]	2020	Using natural language processing methods to extract semantic features and network analysis methods to analyze the following network	Treemap, sequence chart, word cloud, table	Event detection, sentiment, semantics interpretation
[HZC+20]	2019	Trajectory extraction	Flow map, isarithmic map, sequence chart	Identify spatial hotspots, regular commuting routes, and temporal patterns
[QLR+20]	2019	Air quality categorizing, temporal interpolation, spatial and temporal aggregation	Dot map, gauge, line chart, bar chart, word cloud, scatter plot	Identify spatial and temporal patterns, compare air quality among cities, correlate air quality and other factors
[LCZ+19]	2019	Co-occurrence pattern extraction	Dot map, small multiple of maps, sequence chart, bar chart	Identify spatiotemporal co-occurrence patterns
[LXR19]	2019	Multi-dimensional spatiotemporal data clustering	Symbol map, bar chart, tree map	Compare spatial and temporal patterns in different dimensions
[SHH+21]	2018	Using machine learning methods to extract important factors, spatiotemporal aggregation	Symbol map, parallel coordinates, line chart, bar chart	Spatiotemporal distribution, correlation
[SMP+19]	2018	Classification of urban heat island events	Symbol map, calendar view, clock chart, gauge	Identify spatiotemporal patterns
[ZLJ+19]	2018	Natural language processing for event extraction	Dot map, space-time cube, line chart, word cloud	Identify spatiotemporal distribution and sentimental patterns
[PMSR18]	2017	Spatiotemporal and topic aggregation, correlation between topics	Symbol map, matrix, table	Identify spatiotemporal distribution and correlation
[MSFM16]	2016	Statistical methods for spatiotemporal aggregation	Symbol map, calendar view, table, bar chart, gauge	Spatiotemporal distribution, identify anomalies
[GCML06]	2006	Multivariate clustering	Choropleth map, matrix, parallel, coordinate, self-organized map	Identify spatiotemporal distribution with different variables

addition, maps are juxtaposed, e.g., in the form of small multiples, to guide readers to compare and relate spatial patterns [APP11, GMH+06].

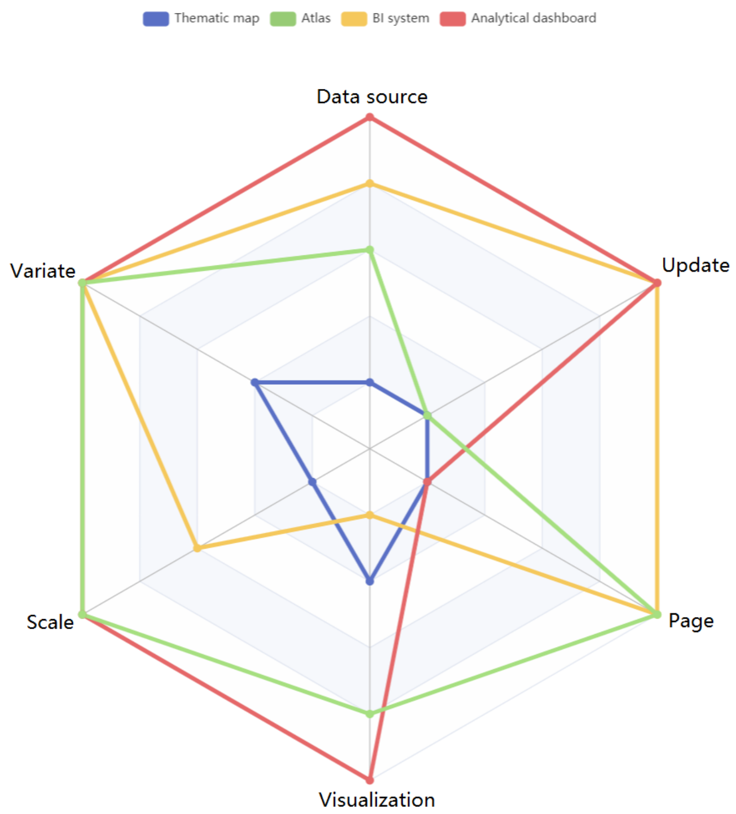


Figure 2.3: The traits of thematic map, atlas, BI system, and analytical dashboard.

In addition to individual visualizations, visual interfaces are designed to synthesize information from multiple factors, spatial scales, and analysis methods. Analytical dashboards, business information (BI) systems, thematic maps, and atlases are widely used visual interfaces to communicate complex information. More specifically, we present the traits of analytical dashboard, BI system, thematic map, and atlas in six dimensions - data source from single to multiple, variate from single to multiple, spatial scale from single to multiple, visualization from simple to complex, page from single to multiple, and update frequency from slow to quick in Figure 2.3. Analytical dashboards and BI systems have the advantage of multiple views and they are able to include heterogeneous data, whereas atlases and thematic maps show fewer types of data. Most atlases, BI systems, and analytical dashboards often show the information of multiple variates, while a large number of thematic maps show a limited number of variates. Analytical dashboards and atlases could cover multiple spatial scales, while BI systems and thematic maps have less ability. Most maps, BI systems, and dashboards have only one page, while atlases contain multiple pages of maps under one topic. Last but not least, analytical dashboards and BI systems can have a much higher update frequency than atlases and thematic maps.

Moreover, we identified the most suitable purposes of BI system, thematic map, and atlas by following the map-use cube [MK97]. The cube covers the four main purposes of maps as explore, analyze, synthesize, and presentation, and we illustrate the purposes of the four visual interfaces in Figure 2.4. Analytical dashboards and BI systems serve for analytical and synthesis purposes, with high interaction and both known and unknown information. Analytical dashboards are targeted to the general public, while BI systems are designed mainly for the experts. In contrast, thematic maps and atlas present known information with low interaction. Most of the thematic maps are targeted to the general public, while atlases show organized information to experts. Many previous studies focus on data mining and representing while neglecting user experience with the interfaces. The potential of integrating user studies into the design process can improve the usefulness, ease of use, and effectiveness and efficiency of knowledge discovery.

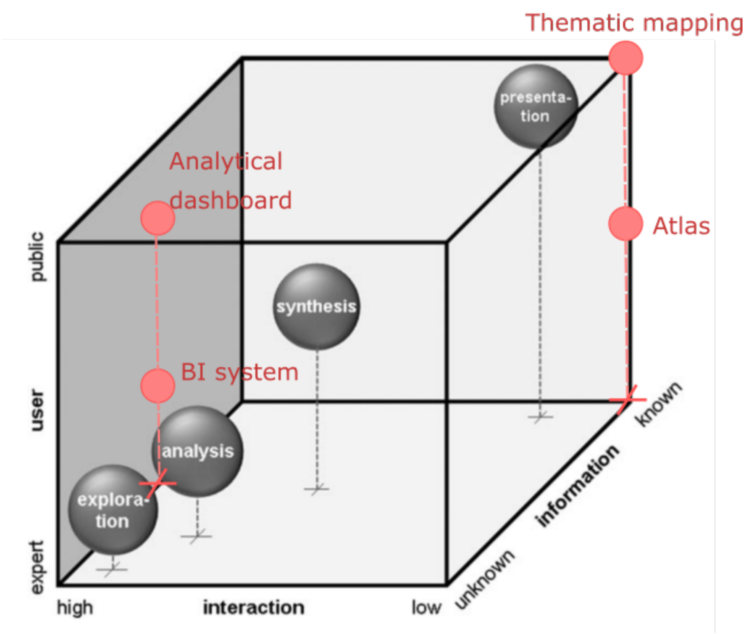


Figure 2.4: Illustration of the purposes of analytical dashboard, BI system, thematic map, and atlas in the map-use cube.

The social environments commonly involve multiple factors, attributes, dimensions, and spatial scales and coverages. It is challenging to logically represent the complex aspects of social environments on an at-a-glance visual interface and with an easily understandable style. When people are short on time, it is especially hard for them to summarize the key information from various information visualizations. Map-based dashboard is a special type of analytical dashboard and has the potential to represent the spatial information of social environments and guide users to learn about social environments and analyze their spatial information.

Everything points to the conclusion that the phrase 'the language of art' is more than a loose metaphor, that even to describe the visible world in images we need a developed system of schemata.

E. H. Gombrich *Art and Illusion*, 1959 (p.76)

3

Map-based Dashboard as a Visual Analytical Design Approach

This chapter is dedicated to the concept and design essentials of map-based dashboard for visual analytical purposes. Section 3.1 introduces the definition, typical use cases, and types of dashboards. Section 3.2 introduces the foundations of map-based dashboards, their features, and the design space. Section 3.3 and 3.4 describe the interface design and evaluation of map-based dashboard, respectively. Section 3.5 outlines the challenges in designing map-based dashboard for the social environment understanding.

3.1 Definition and Aims of Dashboard

Dashboard is a multimedia presentation style that concisely combines texts, images, charts, maps, videos, and gauges to allow users' instant perception. The interactions on dashboards, such as selecting, filtering, searching, arranging, or drilling down, would additionally empower users with the flexibility to view and explore information effectively [JDLL19]. Few [Few06] described dashboard as "a visual display of the most important information needed to achieve one or more objectives, consolidated and arranged on a single screen so the information can be monitored at a glance".

Sarikaya et al. [SCB+19] defined dashboard as "a distinct area of visualization that offers impactful directions for future research." Multiple views do not necessarily improve the efficiency of understanding the visualized data, but they can significantly broaden users' perspectives on the data and reduced the chances of missing important data features within a given time [vdEvW13].

Dashboard is a distinct area of visualization. As mentioned in [SCB+19], there are diverse dashboards designed for many purposes, and they are expanding from simply monitoring to leading users analyze. Their wide use brought challenges to the designers of presenting the overview and details, representing the increasing data in an easy-to-understand manner, and organizing the disparate information on an at-a-glance view. We identify the dashboard as a cross-disciplinary study area, as shown in Figure 3.1. It involves data science, human-computer interaction, information visualization, and cognition science.

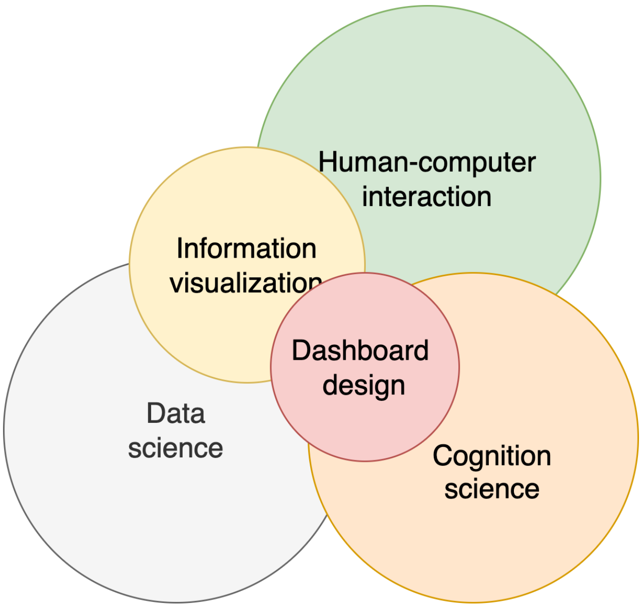
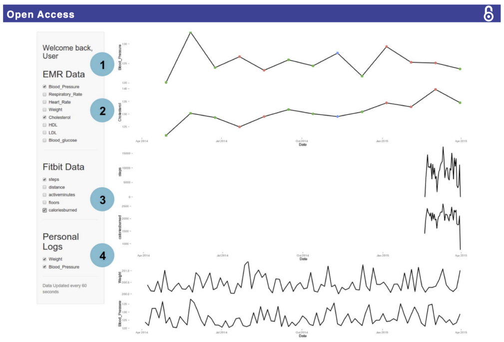
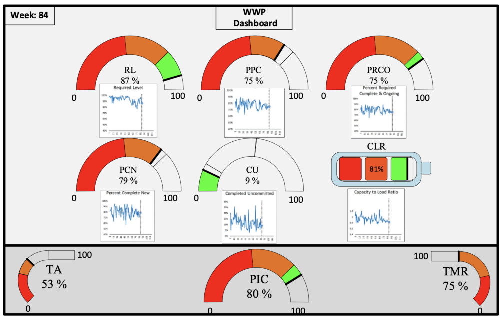


Figure 3.1: Dashboard design and the related domains.

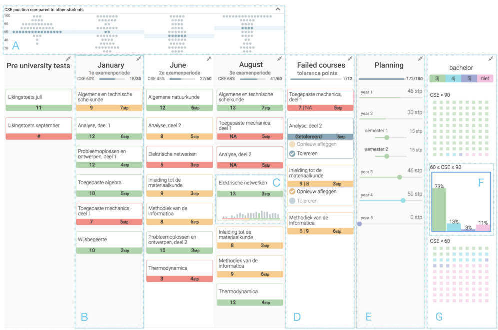
Taking advantage of presenting the overview, communicating the multiple perspectives, dashboards support various stakeholders, such as industry, non-profit organizations, government agencies, medical practitioners, scholars, and citizens, to acquire useful insights [ZDM20, FHWR20]. We selected six typical examples of dashboards shown in Figure 3.2. Badgeley et al. [BSG+16] designed a dashboard, as shown in Figure 3.2a to monitor the real-time clinic metrics of patients. Hamzeh et al. [HES+20] designed a dashboard to show the important metrics in construction projects, as shown in Figure 3.2b. Dashboards are also widely used in education, such as assisting teachers to manage classes as shown in Figure 3.2c [MGC+18] and help students to track their courses as shown in Figure 3.2d [Moo15]. Moreover, dashboards are extended to support the visual analysis of complex topics. Figure 3.2e shows a dashboard designed for analyzing hierarchical data and showing their underlining patterns in multiple views [HCP+20]. Figure 3.2f shows a dashboard presenting the pandemic spatiotemporal patterns and key indicators [MWR+20].



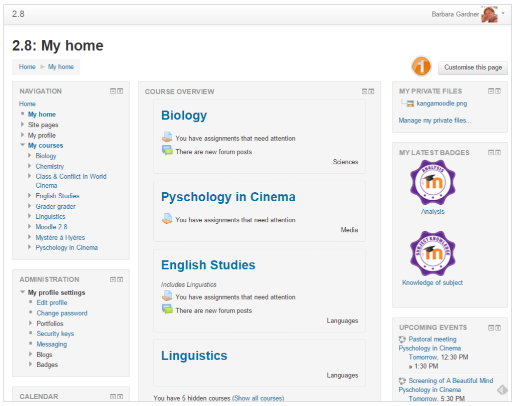
(a) Clinical monitoring dashboard [BSG+16]



(b) Project management dashboard [HES+20]

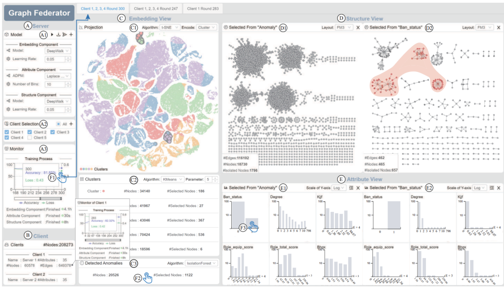


(c) Course management dashboard (teacher) [MGC+18]

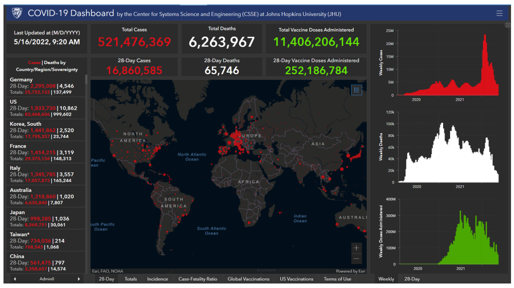


(d) Course management dashboard (student) [Moo15]

Figure 3.2: The interfaces of some analytical map-based dashboards. (Cont.)



(e) Data analysis dashboard [HCP+20]



(f) Covid dashboard [MWR+20]

Figure 3.2: The interfaces of some analytical map-based dashboards.

Depending on the information needs of stakeholders, dashboards can take one of three main types: operational, analytical, and strategic [Few06]:

Operational dashboard. The operational dashboards aim to monitor the situation with a high temporal resolution and use dynamic visualizations to show the changes in detail (e.g., Figure 3.2a and Figure 3.2f). The visualizations reflect the development of events in the near past and keep updating over time. The visualizations can inform users of a series of states, guide users to learn the trend of the facts, and drive users to compare differences between states. For instance, a sport monitoring system included a visual interface for people to learn their sports status [Li19].

Strategic dashboard. The strategic dashboards provide an overview of the most general information and require fewer update and interactions to support decision-making, which requires a collection of information from different dimensions. With the help of the visual metaphors, these dashboards convey information from different sources and formats. The application examples are personal course management [Moo15] as shown in Figure 3.2b and project management [HES+20] as shown in Figure 3.2d.

Analytical dashboard. The analytical dashboards present patterns at multiple level of details, such as a high abstract level or details (e.g., Figure 3.2c and Figure 3.2e). Enabled by the fast development of web technology, the analytical dashboards are often designed as an interactive tool. They integrate big data to show complex topics and allow users to interactively explore the visual representations to acquire insights. Examples can be found in dashboards designed to analyze taxi trips [FPV+13] and crime distributions [OLP+16].

Dashboards are widely designed to support people to understand various issues. Facing the increasing need from the public to understand social topics from a spacial perspective, the dashboards that dedicated to represent spatial complex social issues are not sufficiently studied. The research on map-based dashboards for the visual analysis of social environments is still in its beginning stage.

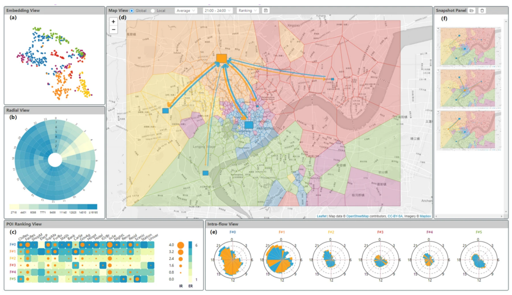
3.2 Foundations of Map-based Dashboard

Map-based dashboards, as one of the widely used web-based visualizations, provide an at-a-glance overview of knowledge and insight into a large number of geographic features and other key information [ZDM20]. They are designed and implemented for purposes such as city monitoring [ZMTK17], spatial data exploration [YS20, HZC+20], spatial data pattern analysis [LCZ+19, LXR19], geographical environment analysis [DWC+20, JCS+19], and geoinformation mining [PMSR18, MSFM16].

A collection of geodata in various topics, e.g., natural resources, urban infrastructures, and social events, can be visualized on map-based dashboards with various information visualization methods. Specifically, maps are used in these dashboards as an instrument, together with other visual representations such as charts, images, and texts, to allow users to examine spatial distributions, compare regional differences, and perform further analysis. Important interactions, such as searching, filtering, view linking, importing, and exporting, are embedded in map-based dashboards to help users explore details. The arranged visualizations provide an at-a-glance overview of geospatial information and that supports users to visually learn and analyze [KLM15]. The well-designed map-based dashboards have the potential to support people to gather information and making strategic decisions [HVdB11].

Map-based dashboards are designed for different scenarios requires different knowledge gaining processes with a spatial focus, such as finding data values, identifying relations, and detecting anomalies, map-based dashboards highlight the geographic information. For instance, Liu et al. [LVG+21] designed a map-based dashboard AQEyes to support visual anomaly detection of air quality data, as shown in Figure 3.3a. They designed a map that allows users to select locations, a calendar view to show the weekly and monthly patterns, and two line charts to show the progression of various factors in air quality. Another map-based dashboard is designed in [LZL+21] to support the visual exploration of urban functional zones as shown in Figure 3.3b. They included six panels to show the overview, weekly pattern, ranking, spatial distribution of intra-flow, and users' search records, respectively. Würstle et al. [WSPC20] designed a single-panel dashboard to present the energy consumption in urban areas, and pop-up windows are available to show the historical energy consumption data if a building is selected, as shown in Figure 3.3c. Map-based dashboards are further applied in pandemic monitoring. For instance, Figure 3.3d shows a dashboard using maps, lines, and tables to show the geographic distribution along with the temporal progressions. Furthermore, Luo et al. [LYL+19] designed a dashboard FBVA to support visual analysis of crowd mobility, as Figure 3.3e shows. Leite et al. [LAS+20] designed a dashboard Hermes, as shown in Figure 3.3f, to support the visual exploration of economic networks. In general, each panel shows an individual dimension of the data, and the panels are linked to collectively show the geographic environment.

(a) Air quality analysis [LVG+21]

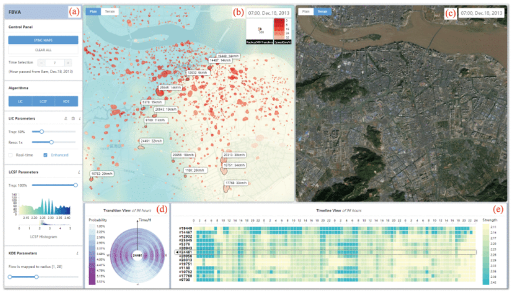


(b) Urban functional zones analysis [LZL+21]

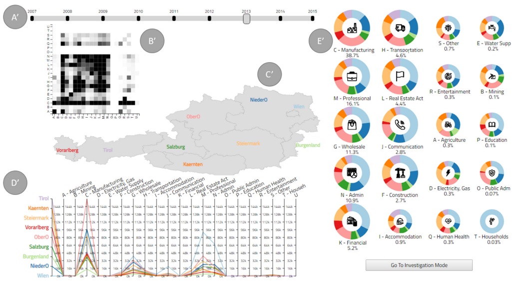
Figure 3.3: The interfaces of selected analytical map-based dashboards. (Cont.)

(c) Energy consumption analysis [WSPC20]

(d) Pandemic analysis [fDPC21]



(e) Mobility analysis [LYL+19]



(f) Economic network analysis [LAS+20]

Figure 3.3: The interfaces of selected analytical map-based dashboards.

We concluded map-based dashboard with five metrics: at-a-glance, spatial focus, data-driven, customized, and reusable.

At-a-glance: Map-based dashboard offers an at-a-glance view of a certain topic on a fixed layout. Its visual interface provides target users an overview of the topic and the key factors.

Spatial focus: The maps are the main elements on the dashboard interfaces and guide users' attention to the spatial patterns. Other visualizations, such as bar charts and gauges, provide statistical information as a response to spatial queries.

Data-driven: To represent the features of various data, visual analytical methods are integrated into map-based dashboards for the purpose of supporting the understanding of spatial information. In addition, it shows the data features, such as multiple perspectives, multiple variables, and multiple granularites in an easily understandable manner.

Customized: Map-based dashboard is a result of a design that considers multiple constrains from target users, topics, design purpose, and data features.

Reusable: The layout helps users to establish a reading habit. Through the arranged visualizations, users are able to locate their areas of interest straightforwardly when they revisit map-based dashboards.

Designers of map-based dashboards must take mindful consideration to maintain these metrics and reflect them in the design space.

Data transformation. The raw data often has a large size and is difficult to interpret. Therefore the designers should choose an appropriate method to transform the raw data into understandable indexes, such as, choosing the suitable spatial and temporal coverages, aggregating the data at various levels, and calculating indexes and statistics.

Visualization. The designers should balance the complex data patterns, supporting insight acquisition, at-a-glance view, and easy understanding by choosing the appropriate visualizations.

Interaction. The most common interactions for analytical dashboards are data filtering and linked view. The designers should provide flexible and sufficient interaction possibilities without overwhelming the users.

Description. The descriptions of data sources, data processing methods, visualization methods, and author information are an important means of transparency for users to understand the topic.

Layout. There are various ways to place the visual elements, i.e., visualizations, titles, information groups, and descriptions. Designing the layout influences the reading orders and further reading strategies of users. A good layout could guide users understand the topic and avoid missing important information.

Color palette. Color design is an important factor for the preservation of a consistent and understandable interface. In addition, good color design can entertain and motivate users.

Of the above dimensions, layout weighs particularly much for analytical dashboard design. However, it has not yet been systematically addressed in previous studies.

3.3 Interface Design of Map-based Dashboard

The layout and visual hierarchy of visual elements strongly influence the effectiveness and efficiency of knowledge discovery. It commonly includes the placement of textual description information (such as titles, legends, data sources, and designers) and the arrangement of the visualizations. Previous studies showed that the visual attention of dashboard users is distributed unequally in different positions [Few06, ZDM20]. Figure 3.4 shows the users' attention when they view the map-based dashboards. The dark orange area indicate intense attention from users.

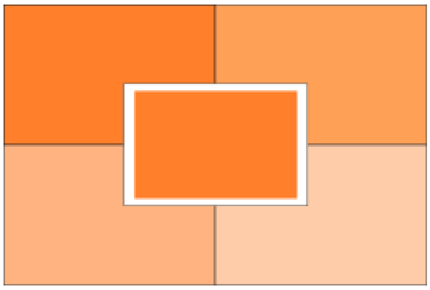


Figure 3.4: Users' attention on different regions of dashboards.

We have identified five typical types of layouts from previous studies. The map-based dashboards in [LVG+21, fDPC21, WZD+21, LAS+20, BWAL20] placed maps at the left or the top-left, as Figure 3.5a shows. It is one of the most common map-based dashboard layout styles. Users read these dashboards from left to right and perceive the spatial information at first. The map-based dashboards designed in [CLCY20, LWCZ20, LZL+21] placed maps in the middle or top middle, as Figure 3.5b shows. Users read these dashboards from left to right, and they firstly customize the map from the left panels and then view the visualizations in the middle and right side. The map-based dashboards in [QLR+20, NB19] placed maps at the right or top-right side of the interface, as shown in Figure 3.5c. These dashboards are rare, and users are guided to focus more on the visualizations placed on the left side than the map on the right side. The map-based dashboards in [WSPC20, APM20, OLP+16] used maps as the background, as shown in Figure 3.5d. These dashboards often show fewer data attributes or dimensions as the geographic information is shown on a single map. Last but not least, the map-based dashboards in [HZC+20, JCS+19, LYL+19, DK20] juxtaposed multiple maps to support visual comparison and correlation of multiple variables, as shown in Figure 3.5e. These juxtaposed maps often show data from the same geographic location but from different times or data sources.

To achieve a high effectiveness of map-based dashboards, the layout should place the important information on the areas that receive the most visual attention from

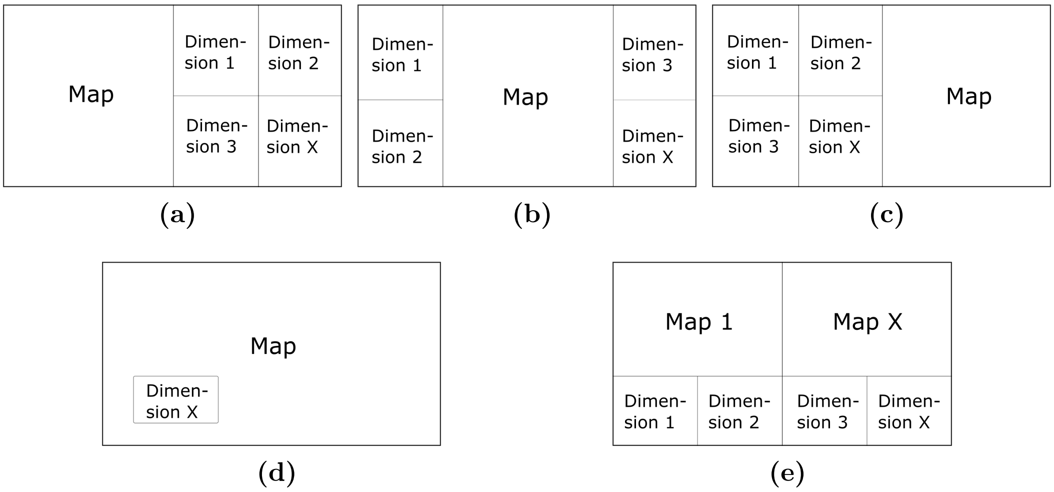


Figure 3.5: Different styles of map-based dashboard layout.

users, so that the users would not miss the important information. Moreover, the arrangement of the visualizations along the reading order should follow a logical order, such as from general to detail or from small to large scale. It can help users firstly to have an overview and then dive into the details.

The visualization and interaction methods also play an important role in the effectiveness and efficiency of map-based dashboard. Figure 3.6 provides a comparative overview of purposes, visualizations, and interactions from 35 selected map-based dashboards. The length of the axis represents the proportion of a feature among all the selected dashboards. These dashboards serve the purposes of data analysis, data management, decision making, monitoring, and learning. In most of the map-based dashboards, maps are used to present spatial information and are coupled with other types of visualizations such as time sliders to show the spatiotemporal distribution. The basic charts are often used to present statistics, such as table, line chart, and bar chart. Some innovative elements are included in the design, such as glyph, parallel coordinates, calendar view, cartogram, sankey diagram, to support some complicated data features like hierarchy and correlation. In addition, the interactions such as filtering, selecting, highlighting, and zooming are commonly used interactions for analytical purposes than for other purposes.

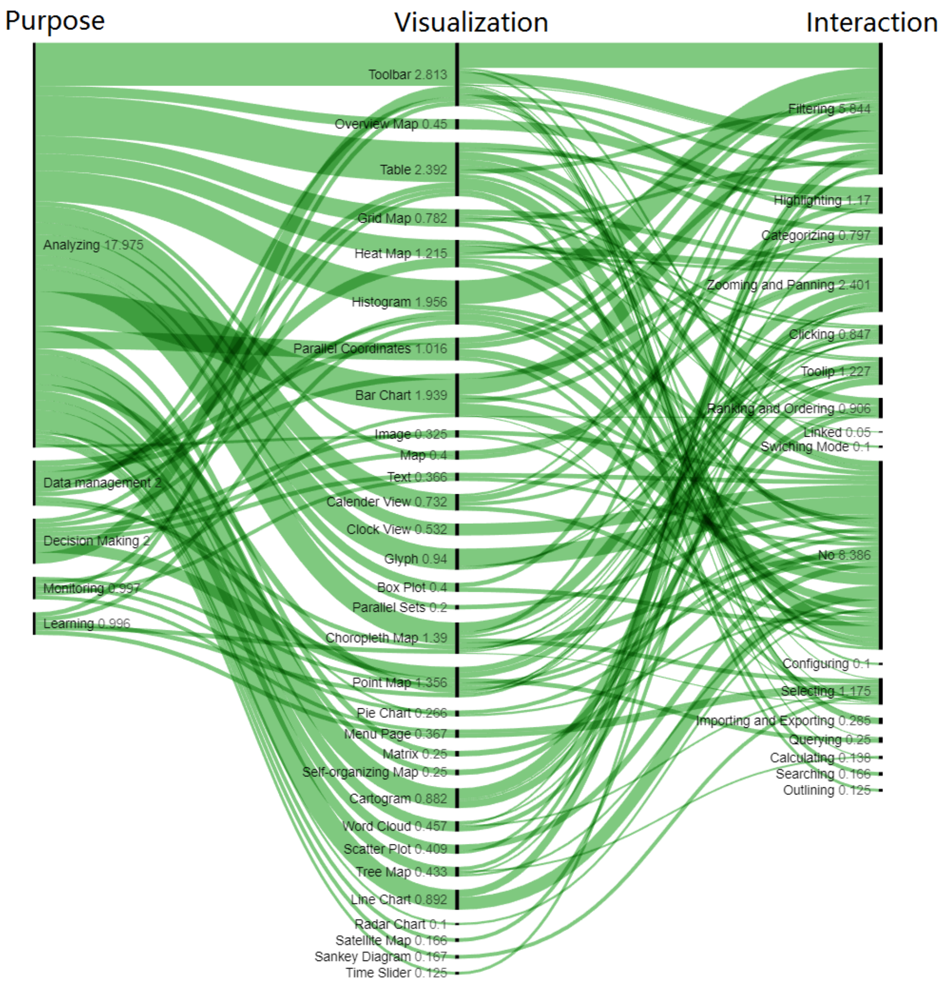


Figure 3.6: The parallel set chart illustrates the correlations among the purpose, visualizations, and interactions of the samples of map-based dashboards.

3.4 Evaluation of Dashboard

Whether dashboards are useful and effective depends highly on user evaluation [Rot13]. User studies are often conducted to evaluate various cartographic visualizations. Numerous studies have focused on investigating the influence of specific design elements on interactive maps.

Usability tests are widely adopted to evaluate the abilities of visual representations and guide designers to improve the visual interface design. Popular methods for evaluating map-based dashboards include survey, interview, think-aloud, eye-tracking, and think aloud method [BGW01]. Effectiveness and efficiency are commonly assessed in various studies. Table 3.1 listed several previous usability studies of visualizations. Among them, [ACG14, IHBD18] evaluated the effectiveness of charts in supporting users in understanding data value and its distribution. They compared accuracy and response time by asking users to conduct benchmark tasks with individual visualizations, such as identification of the data range and comparison. Some studies adopted this method to evaluate the usability of interactive maps [BHW^{+}21, BJ18, CH20]. Moreover, [BJ18, CH20] analyzed eye movement data to assess users' visual attention while conducting the tasks. In general, the benchmark tasks in these studies commonly include low- and medium-level cognitive tasks, such as identifying locations and selecting geographic features. The usability of visual interface can be improved based on these findings.

Table 3.1: The examples of evaluating usability of geovisualizations.

Authors	Purpose	Instrument	Task	Participant Type	Participant Number	Analysis attributes
Albers et al. (2014) [ACG14]	Compare the effectiveness of various visualizations (such as line graph, box plot, and color stock chart) in supporting time series data understanding	Online survey	Identify maxima, minima, range, average, speed, and outliers	General public	582	The accuracy of the tasks
Indratmo et al. (2018) [IHBD18]	Compare the efficacy of various bar charts in supporting comparison	In-lab survey	Single-attribute and overall-attribute comparison	Students	30	The completion time, error rate, and perceived difficulty
Bartling et al. (2021) [BHW+21]	Assessing usability of mobile maps with different base map styles and zoom functions	Online survey	Identify location, select point feature(s)	General public	50	How successful, comfortable and confident the participants were using mobile maps
Bogucka and Jahnke (2018) [BJ18]	Assessing the effectiveness, efficiency, and satisfaction of space-time cube	In-lab eye-tracking	Identify locations and timespans	General public	21	The accuracy and response time of the tasks, the searching strategies, the reported issues by the participants
Cybulski and Horbiński (2020) [CH20]	Assessing the user experience of different map layouts	In-lab eye-tracking and interaction logs	Identify locations and route	General public	40	The respond time, number of fixation, and saccadic amplitude

Map-based dashboards are designed for visual reasoning, therefore an insight-based evaluation facilitates the investigation of the effectiveness and efficiency of the knowledge discovery. Nevertheless, evaluating visual data analysis and reasoning process is challenging because the knowledge acquisition through dashboards varies with user backgrounds, data complexity, experiment tasks, visualization methods, and application scenarios [vW13, LBI+12]. Table 3.2 listed some examples of evaluating how the visual analytical tools help users acquire complex knowledge. In the studies [HJS+18, PHŘ+19, MOC21, OGF14], high-level cognitive tasks, such as comparison, correlation, estimation, summary, and prediction, were included in their benchmark tasks. In the studies [CLZ+17, GOR20], users were asked to freely explore the visual interface and report the knowledge they had discovered while think aloud, and the authors categorized the reported knowledge and analyzed the frequency. Several studies [PHŘ+19, OGF14, GOR20] applied the eye-tracking method to analyze visual attention by measuring the duration and sequence of fixation. In addition, users' subjective preferences and feedback are valuable for improving the designs [HJS+18, GOR20].

Table 3.2: The examples of evaluating visual analytical tools in supporting high-level cognitive tasks.

Authors	Purpose	Instrument	Task	Participant		Analysis
				Type	Number	
Herman et al. (2018) [HJS+18]	Assessing the effectiveness of static and interactive 3D terrain maps	Online survey	Identify locations, select point features, correlation estimation, comparison	Experts and laypersons	76	The correctness, response time, subjective preferences
Popelka et al. (2019) [PHŘ+19]	Assessing the usability of map-based visual analytic tools	In-lab eye-tracking	Utility tasks like select map layers, display legend, knowledge discovery tasks like estimation and summary	Experienced users	17	The sequence, fixation, revisit of eye movement data
Mosca et al. (2020) [MOC21]	Assessing effectiveness of interaction in supporting reasoning with visualizations	Online survey	Prediction	General public	530	The accuracy of the tasks
Opach et al. (2014) [OGF14]	Examining the viewing strategies while exploring spatiotemporal patterns with multi-component animated maps	In-lab eye-tracking	Watching the animation and then answer questions related to identify spatiotemporal patterns, comparison and correlation	Students	23	The accuracy of tasks, the duration and order of fixations, and the percentage of fixation duration per map
Choe et al. (2017) [CLZ+17]	Assessing what insights can users gain from a visual analytical tool	In-lab think-aloud	Free exploration and discover knowledge	Experienced users	11	The frequency of reported insights coded in several categories, i.e., recall, detail, comparison, trend, value judgement, distribution, correlation, outlier, summary, and prediction
Golebiowska et al. (2020) [GOR20]	Assessing the learnability of multiple view geovisualization tool	In-lab eye-tracking, think aloud, and interaction logs	Free exploration and discover knowledge	Experienced users	29	The attention of views measured by fixation, the type and frequency of the use of interactions, and the subjective preferences

Although user studies can help designers adjust the dashboard design and ensure the usefulness of the dashboards, it is difficult for designers to approach a large number of participants for the experiments. As shown in the above Table 3.1 and Table 3.2, most studies have around 20 to 50 participants.

3.5 Challenges in Designing Map-based Dashboard

Social environments consist of various data in multiple dimensions. Using visual languages to represent the social environments can potentially help diverse users to gain an intuitive understanding. Map-based dashboard as one of the instruments helps users to access data, provides an overview, and guides users to obtain insights. The sense-making model described by Baker et al. [BJB09] included four components, world, data, visual representation, and people. It means people understanding the world by viewing the representations and the data. Noteworthy, human cognition influences the insight acquisition process and we address this component and integrate it in the sense-making model as shown in Figure 3.7.

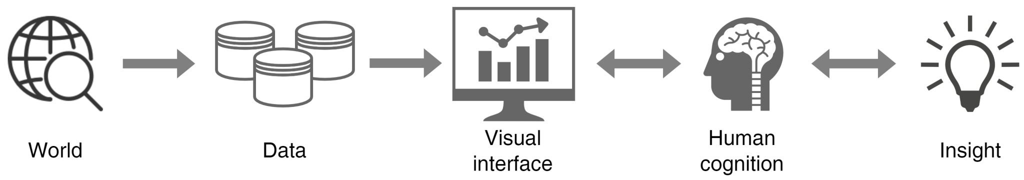


Figure 3.7: The extended sense-making model of humans understanding the world with visual interfaces and the insight is influenced by human cognition.

Although visualizing spatiotemporal data has a long tradition, several challenges remain in designing map-based dashboards for understanding social environments.

Data design. The social environments commonly involve many factors. Selecting and acquiring relevant datasets from various sources and portals, fitting spatial and temporal coverages, transforming raw data into understandable parameters by applying appropriate processing methods, and keeping this information transparent to users require both domain knowledge and data handling skills.

Organizing visual vocabularies. Applying appropriate visualization methods for various data is difficult. In addition, organizing the visualizations on an at-a-glance screen and guiding users' attention to these visualizations in a logical order have to be considered has not been addressed in previous studies.

Provide easy accessibility. Users need not only the self-explanatory visualizations but also interactions such as data selection and filtering to dive into the underlying data and data aggregation to gain knowledge at a large scale. Identifying the most relevant data dimensions, connecting them by interactions, and keeping interactions easily reachable to the users need to be balanced by the designers.

Involve human cognition in the design process. User evaluation can ensure the usefulness of map-based dashboards. Designing experiments to explore users' information needs and visual viewing behavior and derive cues for improving dashboard design is an area of research that is far from well understood.

We never look just at one thing;? we are always looking at the relation between things and ourselves.

J. Berger *Ways of Seeing*, 1972 (p. 9)

4

Design Methodology of User-oriented Map-based Dashboard

To ensure the ability of supporting spatial analysis, we extend the user-center design method [RRM15] as a map-based dashboards design framework. The framework is composed of five phases, including *design goals, users' cognitive tasks, data, interface, and feedback*, as shown in Figure 4.1. First, formulating clear design goals can help dashboard designers to define the scope. In the *design goals* phase, the dashboard designers should understand the nature of certain social environment and outline its required information, then highlight the information according to the users' experience in the environment and visual interfaces. Second, translating design goals into specific cognition tasks can guide designers to select and process corresponding data and design visual interfaces. Analyzing the *users' cognitive tasks* can help the designers to prioritize the relevance the data, analysis methods, and visualization methods to the target groups. Third, defining data space should be designed align with the *users' cognitive tasks.* In the *data* phase, appropriate data would be collected, prepared, and transformed according to the previous phases. Forth, providing interactive visual analytical interfaces can facilitate a broad audience to acquire insights. In the *Interface* phase, the visualization methods, interactions, layout, and cartographic methods are jointly representing the characteristics of the processed data from the last phase. Last but not least, the previous phases should be evaluated and improved by collecting feedback from domain experts and potential users. In the phase *feedback*, the usability and user experience, such as efficiency and viewing strategy, would be collected. Integrating the five phases as a loop, map-based dashboard would be designed under the social environment context.

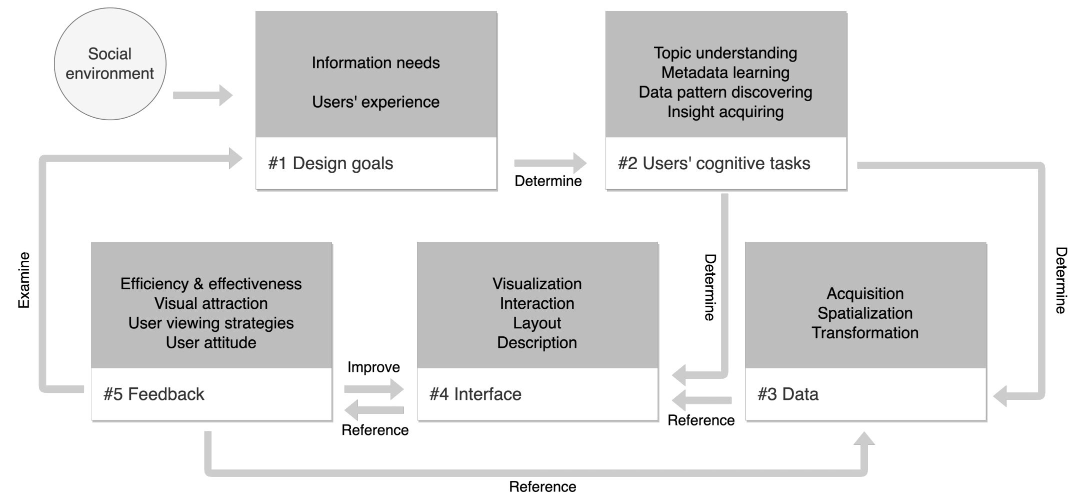


Figure 4.1: The design framework of map-based dashboard to understand social environment.

4.1 Design Goals

Social environments contain a wide range of topics, including geographic environment, the interaction between humans and geographic environment, and activities in human society. It is a mixture of landscape, natural resources, economic activities, and cultural atmosphere. Each topic exhibits different data sources, spatial and temporal coverages, and key factors. The stakeholders such as policy-makers, investors, and citizens are interested in learning and analyzing the social environment to support their decisions in various tasks. Thus, understanding the stakeholders and their information needs in social environments is the basis of designing a dashboard.

We conducted a survey on June 2018 to formulate the design goals of understanding social environments. We recruited 18 stakeholders of social environments to ask about their current information source, their needed information, their experience of dashboards, and their attitudes through questionnaires. Among the experiment participants, 16 were entrepreneurs from 25 to 60 years old, and they were involved in the decision-making process in private firms. Most of them collected such as economy, population, finance, and talents data from open government data portals.

The local and regional spatial distributions of these factors are important for them in their strategy-making process.

We identify four typical characteristics of the users from the survey. (1) They need the information on social environments to examine their decisions, although they are not experts in the social environments. They have to learn the scope and key factors of certain social environments. (2) Some of them are not experienced in social environment analysis. For instance, they might not be aware of, e.g. the average value of certain factors and the typical modeling methods. (3) They are familiar with the low graphicacy visualization methods (such as line charts and bar charts), but might not have used the high graphicacy visualizations (such as parallel coordinates and treemap). (4) They are limited in time and motivation to acquire knowledge from dashboards.

Facing the typical characteristics, we propose six design goals (DG) to guide the design process of analytical map-based dashboards. The DGs fall into two categories, namely information design goals and user experience design goals. DG1, DG2, and DG3 ensure the users are aware of the topics, the key factors, and the spatial features.

DG1. At-a-glance view. It is important for the stakeholders to have an overview of a social environment and values of key factors. These information should support them in answering essential questions, for instance, which factors should be considered in a social environment, what are the values of key factors, what are the relations between key factors, and what are the spatial distributions. In addition, descriptive information and interaction functions should be easy to find on the interfaces.

DG2. Focus on spatial patterns. Spatial context, location, and patterns are the fundamental information in the decision-making and policy-making process. Displaying the patterns based on their geographic location can support users gain an overview. Maps can be used to show spatial distributions, and they can be connected with various visualizations to show patterns such as ranking and correlation based on geographic locations.

DG3. Show multiple granularities. Showing the social environment information on different spatial scales, such as local, regional, and global levels, can support stakeholders to learn their interested areas and have an overarching view of a social environment. Therefore, multiple aggregation levels and corresponding visualizations should be designed to support users to learn, compare, and relate key factors.

However, making various datasets and adequate data visualizations accessible to stakeholders are not enough. Interactive visual tools should be designed to support stakeholders to understand the complex social environments easier. We specified these requirements into three design goals to support a good user experience.

DG4. Support visual reasoning. Compared to textual reports and tables, visualization could quickly provide users with reasoning possibilities for various analysis tasks, such as finding values, identifying ranges, comparing places, and correlating factors. Users should be able to use the dashboard interface for social environment reasoning without intensive training.

DG5. Support user interaction. Different stakeholders have different interests in social environments. Some of them may focus only on several specific cities, and some on one or two categories of social environments. In addition, users need to examine their assumptions on a higher or lower aggregation level. Therefore, users should have the possibilities to select, search, filter, switch, and highlight data.

DG6. Easy to use. The stakeholders usually have very different profiles and their knowledge about dashboard differs. Thus, the interface should be self-explanatory, allowing users to interpret the information in a short time. The arrangement of elements on the interface should follow the common reading flow. The interface should include the least amount and the most efficient interactions to reach a better user experience.

4.2 Cognitive Tasks

Translating design goals into specific cognition tasks can guide designers to select and process corresponding data and design visual interfaces. According to the previous studies [BJB09, AES05, AAG03, Tuf16], the main tasks for visual analysis of data include observing specific data points, observing data ranges, comparing multiple datasets, identifying spatial distributions, identifying outliers, correlating with previous knowledge, generating hypotheses, and anticipating real-world context. Cognitive tasks for understanding social environments refer learning the scope and gaining knowledge. In this section, we identify and summarize the cognitive tasks into three basic information learning tasks and eleven knowledge acquisition tasks.

Firstly, the basic information is needed for users to get acquainted with the subject field, including the description of dashboard topic, information of data and publisher, and user instructions of interface.

Learning the topic. Description of the topic allows users to understand the general purpose and content of map-based dashboard. It includes a concise title of the map-based dashboard, an introduction of the background, a description of the panels.

Understanding the background. Some users need to know data source and designers' information to learn the background of map-based dashboards. Keeping these information transparent can help create trust in dashboards. In addition, some power users need to know the data processing methods to gain a thorough understanding.

Knowing the function of dashboards. The interface should provide self-explanatory visualizations, interactions, and a logical layout for users to view and interact with. Sufficient usage instructions about the visualizations and interactions, for instance, data selection, searching, and 3D scene rotation, should be provided to users.

After learning the basic information of dashboards, users need to view the visualizations to acquire knowledge from social environments. Following eleven essential knowledge discovery tasks are applicable to social environments understanding.

Locate places. This is the fundamental task in a spatial analysis. Users have to find the location of certain places on maps, identify their borders and sizes, and learn the surrounding places. Location is used as an anchor point for users to find and compare values.

Find values. This task requires finding values of given variates, for instance, how much population a specific city had in 2015, and how many scientific articles were published in a specific country in 2018. This task is often performed while monitoring KPIs and serves as a subtask of many other tasks such as computing values and finding extremum.

Compute metrics. This task calculates datasets with aggregation functions, such as count, average, and sum. It answers questions such as how many accumulated visitors are in a city in the last six months, how many projects are internationally collaborated in a specific city in 2015. In addition, more advanced computing, such as spatial aggregation, interpolation, topological calculation, and graph analysis can be conducted.

Compare. This task requires to compare of several given data cases. In the social environment, values can be compared between multiple areas in a certain time period or between different time periods in a certain area. It answers questions such as whether City A has a higher number of students than City B. The proportion of a variate in a higher aggregated level is as well often calculated and compared, for instance, the share of the first industrial gross take in the gross domestic product in a city in 2020. Compare is a subtask of finding extremum and sorting.

Find extremum. This task requires finding N top/bottom values among a series of values, for instance, the city with the highest number of graduate students. Considering the spatial extremum, the task requires finding the spatial borders and the related objects, for instance, the east-most city in a specific area and the closest gas station near a specific street. The extreme values provide a reference for users to understand the meaning of values.

Sort. This task requires ranking a set of data cases, which normally includes more data cases than finding only the extreme values. For instance, ordering the cities by the number of the research institute or by the number of immigrants.

Categorize. This task is an extended task from sorting, and it requires finding similarities among data cases in a given dataset and splitting them into several groups. The groups can be numeric or non-numeric, dealing with the questions such as how many different types of land-use areas are in a city. This task can support users to find patterns.

Characterize distribution. This task requires generating the distribution of a given series of values, such as the age distribution of a region. It requires a deeper understanding of the dataset than finding extremum. In addition, spatial clusters can be identified by observing the visualization on maps, such as hotspots of certain weather conditions in a specific region. This task can support the understanding of a dataset on a higher abstract level.

Find outliers. Identifying outliers of a given dataset is relevant to finding distribution, which focuses more on the unexpected/exceptional values, such as the exceptional relationship between GDP and average residence area, and outliers of cities with large enterprises. Identifying outliers often triggers in-depth thinking of users and association with personal knowledge and experiences.

Correlate. Mutual influences can be calculated between two variates, such as using variate A as a unit to calculate the portion of variate B on each unit of variate A. For example, the average income is strongly correlated with the gross regional product, and the economic growth might not be influenced strongly by the infrastructure in a specific region.

Generate hypothesis. Based on the understanding of the discovered patterns and insights from data and background knowledge, users can generate hypotheses. For instance, a specific city can provide more skilled workers for a specific company.

In general, a dashboard should reflect all these three basic information learning tasks and selectively focus on one or several knowledge acquisition tasks.

4.3 Data Design

Data design, including the data acquisition, data selection, and data modeling, is conducted based on the design goals and cognitive tasks. In addition, it also influences the subsequent visual design.

As mentioned in Section 2.1, many data sources and modeling methods are available to reveal insights into social environments, such as statistical data and social network data. Many open data portals provide data in topics such as agriculture, economy, education, population, and public safety. For instance, the European data portal data.europa.eu [1], the German open government data portal GOVDATA [2], and Singapore's public data portal [3]. There are also much open statistical data provided by the statistical offices, such as the Eurostat [4], Vienna Yearbook [5], and OECD Health Statistics [6]. In addition, some social networks provide partial data to the public, e.g., Twitter data [7] and Uber data [8].

[1] https://data.europa.eu/
[2] https://www.govdata.de/
[3] https://data.gov.sg/
[4] https://ec.europa.eu/eurostat
[5] https://www.oeaw.ac.at/vid/publications/serial-publications/vienna-yearbook-of-population-research
[6] https://www.oecd.org/
[7] https://developer.twitter.com/en/docs/labs/tweets-and-users/quick-start/get-tweets
[8] https://eng.uber.com/category/articles/uberdata/

In this section, we illustrate the statistical data and social network data with two examples. The first one uses open government data to reflect the socioeconomic environment, and the second one uses big scholarly data to reveal the knowledge innovation environment.

Open government data. The test datasets are socioeconomic data at the municipality level provided by the Yangtze River Delta Science Data Center [9]. The datasets originated from yearbook data. The datasets cover various topics, including gross domestic product (GDP), public investment, industrial output, population, and employment.

In this study, we select four representative categories of socioeconomic environment, including economy, inhabitant, infrastructure, and R&D, and further identified 22 related factors in these four categories. Table 4.1 shows an example of the raw data in city Nanjing. In addition, we collect the administrative boundary data from the National Catalogue Service for Geographic Information [10].

[9]http://nnu.geodata.cn:8008/
[10]https://www.webmap.cn

Table 4.1: The selected factors of IIE and an original data example of Nanjing.

Category	Factor	Value	Unit
Economy	Gross regional product	12820.4	100 million Yuan
	Gross of the first industry	273.42	100 million Yuan
	Gross of the second industry	4721.61	100 million Yuan
	Gross of the tertiary industry	7825.37	100 million Yuan
	Enterprise above designated size	2556	-
	Foreign capital	38.53	100 million US dollar
Inhabitant	Population	696.94	10 thousand
	Number of employees	462.6	10 thousand
	Budget revenue	1470.02	100 million Yuan
	Budget expenditure	1532.72	100 million Yuan
	Average citizen disposable income	52916	Yuan
	Housing area per capita	43.4	Square meter
Infrastructure	Number of buses	9246	-
	Total road length	10632	Kilometer
	Total urban green space	92202	Hectares
	Water supply	133978	10 thousand tons
	Electricity supply	606.4	10 billion kWh
	Gas supply	129494	10 thousand cubic meters
R&D	Number of the granted patents	99020	-
	Number of undergraduate students	85.68	10 thousand
	Number of Public library books	19644	Thousand
	Number of higher education institute	53	-
	S&T budget expenditure	80.54	100 million Yuan
	Number of expert	10.23	10 thousand

We preprocess the data in several steps. Firstly, we clean the data by converting the local unit of measurements to international standard units, and we adjust the digits of values to a unified precision. Secondly, the data is aggregated into specific geographic areas, time intervals, and IIE categories. Third, multiple indexes reflecting different perspectives of IIE are calculated, such as minimum, maximum, average, mean, and standard deviation, to show the data distribution. To support further comparison between categories, the data are normalized as Equation (4.1) shows, and the Pearson's coefficient between two categories are calculated as Equation (4.2) shows.

$$x_i = \frac{x_i - x_{min}}{x_{max} - x_{min}} \tag{4.1}$$

where x is a value from a group of numbers, i is the number of factors ($i = 1, 2, ...n$).

$$r = \frac{\sum(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum(x_i - \overline{x})^2 \sum(y_i - \overline{y})^2}} \tag{4.2}$$

where x and y stand for values from two groups, respectively.

The land-use type and area are considered in the spatial interpolation model. As previous studies show, there is a strong positive correlation between land-use subcategories and GDP values, such as the GDP value on the built-up area is higher than barren land. A few studies were conducted to estimate the GDP at a finer level of granularity based on land-use data [GRXY04, GOR+07]. We use a model [HJZ+14], as Equation (4.3) shows, to calculate GDP grid using census countywide GDP data and land-use data.

$$GDP_j = \sum_{i=1}^{n} \left(g_i \times L_{ij} \right) + B_j \tag{4.3}$$

where GDP_j is the GDP value of the j-th county, g_i is the GDP density of the i-th land-use subcategory, L_{ij} is the area of the i-th land-use subcategory of the j-th county and B_j is the intercept. The values of g_i can be calculated by the least square method by applying the method to all the counties. Then the gridded GDP values can be calculated by multiplying the area of land-use types and the GDP density g_i.

Big scholarly data. In this study, we use the scientific publications and joint publication networks to reflect the knowledge innovation environment. We establish the academic network at two representative scales: the institution level and the city level. At the institution level, the nodes represent institutions where the authors affiliate, and the weights of nodes represent the number of publications of an institution. The links represent the co-authorships between two institutions, and the weights of links represent the number of joint publications. For instance, if a publication has several authors, each affiliation of the publication is represented as one node, and this publication contributes weight for each node; the links are established between every two involved affiliations, and each pair of authors with different affiliations contributes to the weight of the link. Similarly, the network is also established at the city level, with each city as a node and the intercity co-authorship as links. In this study, we consider the networks to be undirected because the communication among the collaborated authors could be bidirectional.

We collect this test data from the public webpages of the ACM Digital Library via web crawling technologies. The ACM Digital Library (`https://dl.acm.org/`) is a research platform with a large collection of literature in the field of computing. Among dozens of publication categories, we collect the publications in the category of journals, proceedings, letters, and magazines (`https://www.acm.org/`

publications/about-publications), which are composed of major publications of the digital library. The collected items are from January 2017 to December 2019, with information about the title of publications, the name and affiliation of the author(s), the publication time, the domain classification, and the uniform resource locator (URL) of the webpages. Table 4.2 shows an example of the scholarly data.

Table 4.2: An example of the scholarly data.

Item	Data
Title	BubbleNet: A Cyber Security Dashboard for Visualizing Patterns
Author	Sean McKenna ‖ Diane Staheli ‖ C Fulcher ‖ Miriah Meyer
Affiliations	University of Utah ‖ MIT Lincoln Laboratory ‖ MIT Lincoln Laboratory ‖ University of Utah
Published time	01 June 2016
Domain	Human-centered computing ‖ Interaction design ‖ Interaction design process and methods ‖ User-centered design
Domain ID	10003120 ‖ 10003120.10003123 ‖ 10003120.10003123.10010860 ‖ 10003120.10003123.10010860.10010859
URL	https://dl.acm.org/doi/10.5555/3071534.3071565

The collected dataset contains 137,818 publication items of publications, 305,815 authors, and 39,030 institutions 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. Specifically, many publication items have multiple authors. The name and affiliation of the authors are collected separately, but they follow the order as in the publications. In addition, the hierarchical structure is utilized to classify research domains for the publications (https://dl.acm.org/ccs). According to the ontological system, the computing field is divided into 13 domains, such as computing methodologies and human-centered computing. Within each domain, it is further divided into subdomains. Every domain or sub-domain has a predefined domain ID.

Before visualizing and analyzing the scholarly data, data cleaning and preprocessing have to be carried out. Data cleaning includes data restructuring and text encoding. The data is restructured into four types of entities, including paper, author, institution, and domain, and these entities are connected with extra tables (as Figure 4.2 shows). Text encoding converts non-English letters and mathematical symbols into the UTF-8 encoding system. Data preprocessing includes georeferencing, topological calculation, and aggregation. The georeferencing process converts the name or address of affiliations into longitudes and latitudes. We use the open-source geocoding service Nominatim (https://nominatim.org/) to obtain the coordinates. The topological calculation process projects the coordinates of the affiliations onto multiple administrative divisions, such as cities, regions, or counties. The administration boundary data is collected from the geodata open data portal `http://nnu.geodata.cn:8008/`. Finally, the number of publications, co-authorships, and domain information are aggregated onto administrative units.

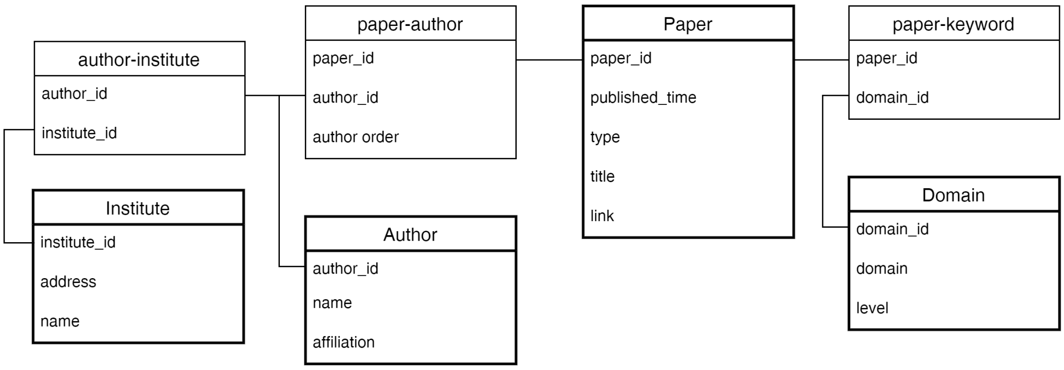


Figure 4.2: The data schema of the restructured scholarly data.

To explore the knowledge innovation patterns and to further support the decision-making in regional analytics, the network analytical methods on the network level are applied. We select scale, node, link, degree centrality, and gravity center as representative metrics to measure the academic networks. In addition, the geographic locations of the nodes are calculated, and the spatial distributions are analyzed. The scale measures the number of nodes in a network, indicating how many institutions or cities were involved in publications. The weights of nodes and links reflect the number of publications and collaborations. We further calculate the median and standard deviation of the weights to measure the statistical distributions of the publications and the collaborations at the institution and city levels. The degree centrality of network is calculated as the number of links a node has. We use it at the city level to measure how active each individual city is in reaching out for inter-city collaborations. The gravity center is the weighted center of the nodes, indicating the theoretical central location of the innovation network (Equations (4.4) and (4.5)). We calculate it at the city-level network to analyze the spatial distribution of the knowledge innovation in the region.

$$Longitude_{GC} = \frac{\sum_{i=0}^{n} Longitude_i Weight_i}{\sum_{i=0}^{n} Weight_i} \tag{4.4}$$

$$Latitude_{GC} = \frac{\sum_{i=0}^{n} Latitude_i Weight_i}{\sum_{i=0}^{n} Weight_i} \tag{4.5}$$

where GC is the gravity center, i is a node in a network.

4.4 Visual Design

Visual design in map-based dashboard aims to guide users to understand the scope of social environments and acquire insights. This section firstly introduces the role of the visual components in the sense-making process, then describes the typical visualization methods, interactions, and layout used in map-based dashboard.

Based on the analysis of the cognitive tasks introduced in Section 4.2, we relate the visual components and the sense-making process as shown in Figure 4.3. First, it is very important to introduce users to a specific topic with the titles, data source, and designers' information. Second, users read the related factors and the spatial coverage from the title of panels and the visualizations. Third, users learn the important values by reading such as the maps, KPI boards, and charts. Finally, users search details, compare the values, and relate to their previous knowledge to formulate assumptions and make predictions about social environments.

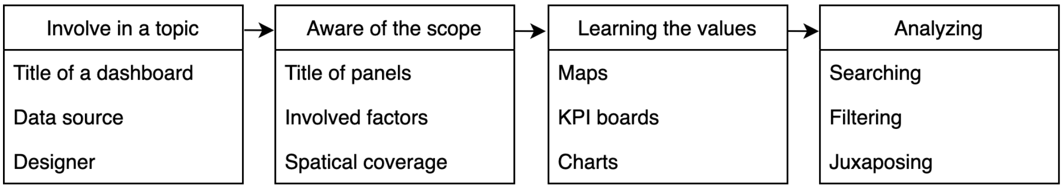
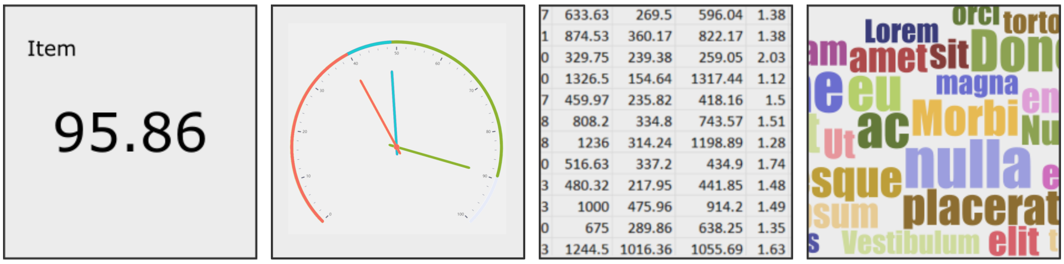


Figure 4.3: The sense-making processing of using a dashboard to understand the social environments.

More specifically, we introduce the visualizations that can be applied in map-based dashboards to support learn and analyze social environments. The visualizations are introduced in four groups, and they are supporting users in learning data values, ranges, relations, and spatial patterns.

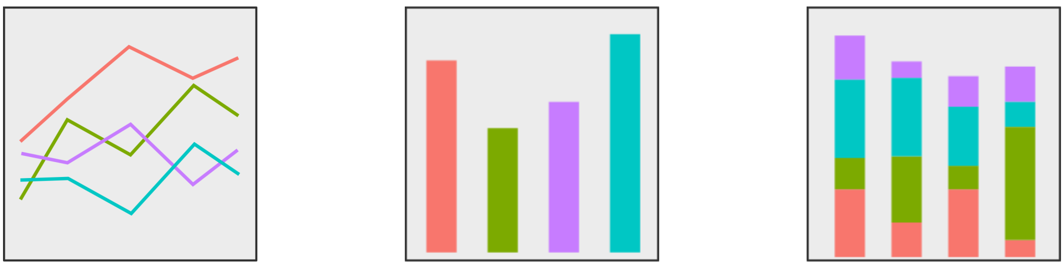
Data values. With regard to origin data and the selected important values are necessary to be shown to users directly. As shown in Figure 4.4, KPI boards and gauges are often applied to show the selected important values, such as the population of a city or the GDP of a region. Gauges additionally present the common ranges of the value to the users. The original numbers and texts are sometimes needed by the users, and the tables and word cloud can be used to represent the selected values.



(a) KPI board **(b)** Gauge **(c)** Table **(d)** Word cloud

Figure 4.4: Visualization methods presenting data values

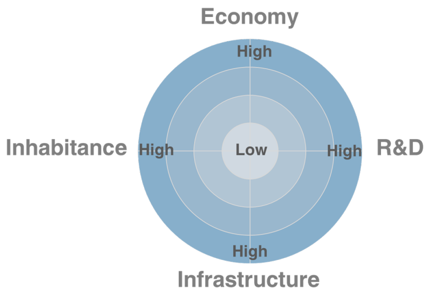
Ranges. Showing the value distribution and ranges can provide users an overview of social environments. We select line charts and bar charts to show these data characteristics, as illustrated in Figure 4.5. Line charts can be used to show and compare several selected factors in a certain time range. Bar charts can be used to highlight the top-ranked cities and the domain values. Staked bar charts can be used to juxtapose multiple domain values in multiple years of a city.



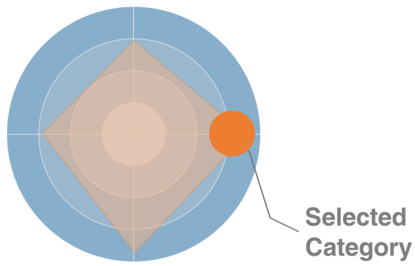
(a) Line chart **(b)** Bar chart **(c)** Stacked bar chart

Figure 4.5: Visualization methods presenting ranges

To present the multiple factors of social environments, we design a radar chart with multiple axes (as shown in Figure 4.6). Each of the axes represents one factor of a certain social environment, for example, economy, inhabitant, infrastructure, and R&D. As Figure 4.6a shows, the range of the axes is from low to high from the center to the border of the circle. Figure 4.6b shows an example of one selected factor marked in bright color. The mark changes as the users' selection of category changes. Supported by the radar chart, users can easily compare the values of various factors.



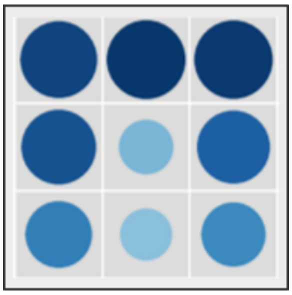
(a) A four-dimensional radar chart showing four factors.



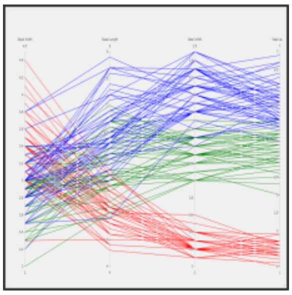
(b) An example showing one selected factor.

Figure 4.6: A designed radar chart representing multiple factors

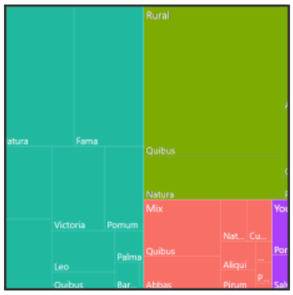
Relations. It requires users to view more complex data patterns such as hierarchical proportion, multi-variable correlations, temporal progression of multiple parameters, and linkage and correlation between multiple entities. Identifying these data patterns is cognitively effortful. We adopt treemap, parallel coordinates, horizon chart, arc diagram, and correlation matrix to present such patterns, respectively, as shown in Figure 4.7. Correlation matrix is used to show the relative correlation of multiple factors. The size of the bubble is proportional to the degree of correlation. Blue means positive correlation, and red means negative. Parallel coordinates show multiple factors of multiple cities. Each coordinate stands for a factor, and each line stands for a city. Treemap is used to show the proportion of the popular research domains. The bigger the rectangular size, the more popular the domain. Two levels of domains are shown in the treemap. The first level is shown in the same color, and the rectangular is further divided into smaller rectangular to show the second level. Arc diagram is applied to show research collaborations among counties. Each arc stands for the number of joint publications between two counties. The thinker the arc, the more connections between two counties. Horizon chart is applied to show the temporal changes of selected research domains. The horizontal axis stands for time. Each row stands for one research domain. The darker the color, the more popular the domain.



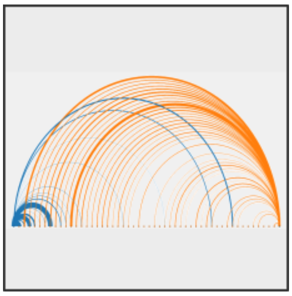
(a) Correlation matrix



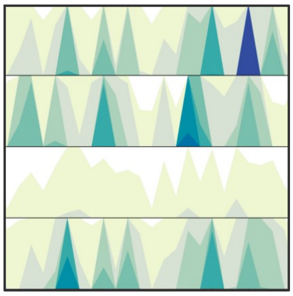
(b) Parallel coordinates



(c) Treemap



(d) Arc map

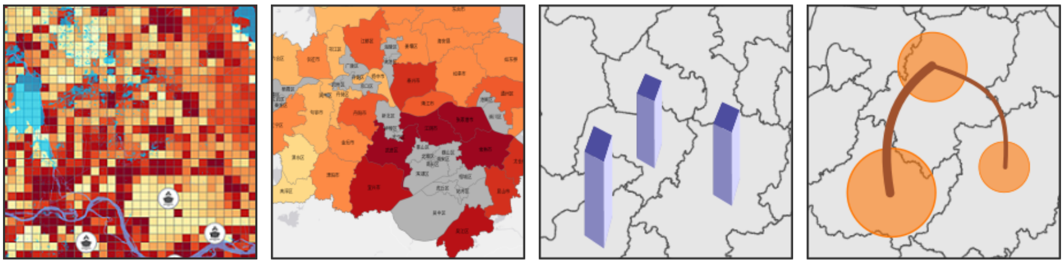


(e) Horizon chart

Figure 4.7: Visualization methods presenting relations

Spatial patterns. Various maps are designed to reveal spatial distribution of factors that involved in social environments. They could also guide users to identify spatial distributions and trigger them to acquire insights. We adopt heatmap, choropleth map, 3D symbol map, and flow map to visualize the distribution of factors on multiple aggregation levels, as shown in Figure 4.8. Heatmaps can show the spatial distribution of domain values on aggregated grids. Choropleth maps are good at displaying classified domain values over administrative areas. 3D symbol maps can show the magnitudes of factors using the height of symbols. Flow maps can show the linkage of multiple areas. The size of the circle shows the number of scientific publications, and the widths of the arc show the number of joint publications between cities. In addition, the aforementioned gauges and charts could be superimposed on maps to show the data patterns together with their spatial distributions.

Interactions provide users with the possibility of examining the values from multiple granularities (such as drill-down or roll-up) and setting the focus on their interests (such as zoom in or highlight). Although interactions may facilitate visual reasoning, they also largely increase the complexity of the interface design. It is therefore important to design efficient interactions while keeping the interactions concise for

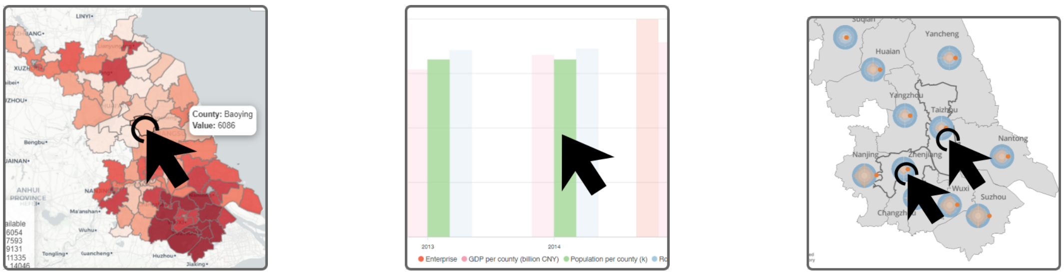


(a) Heatmap (b) Choropleth map (c) 3D symbol map (d) Flow map

Figure 4.8: Various maps presenting spatial distributions

the users. In this thesis, we grouped the interactions in four categories, including detail-on-demand, data manipulation, spatial interaction, and linking view.

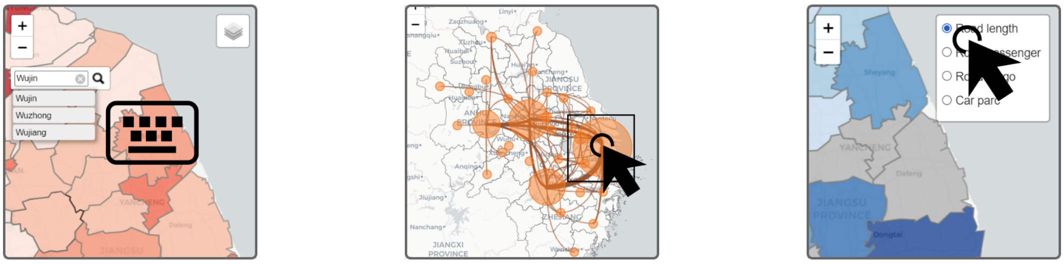
Detail-on-demand. This interaction is as well widely applied in maps, charts, and symbols. In using web-based interfaces, users might click anywhere where they expect more details. Pup-up windows and highlighting effects, as shown in Figure 4.9, are commonly used to support users to examine data values and focusing on a few places or factors.



(a) Detail on demand (b) Highlighting (c) Selecting

Figure 4.9: The visual design of interactions to support showing detailed information.

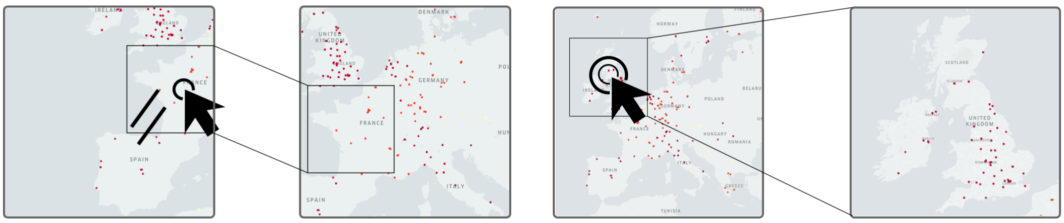
Data manipulation. Except for examining the data values in detail, users need to retrieve more information about a place, filter data according to certain criteria and switch to other related factors of the social environments. Figure 4.10 shows the visual design of these interactions. We adopt spatial searching and integrated it into maps to allow users to search for detailed information about a place. Filtering is applied to enable users to examine such as the spatial connections of one city. The switching function allows users to check and compare different factors of the cities.



(a) Searching **(b)** Filtering **(c)** Switching

Figure 4.10: The visual design of interactions to support data manipulation.

Spatial interaction. Panning and zooming are two interactions that allow users to learn more about spatial information. As Figure 4.11 shows, these interactions are widely used in mapping, so users can interactively set the range. In the social environment, zooming and panning are rather restricted to the study areas instead of allowing users to freely explore anywhere.



(a) Panning **(b)** Zooming

Figure 4.11: The basic interactions.

Linking view. It is a very helpful function in analytical map-based dashboards. It can synchronize the same subset of data and present it in different panels, such as in maps and other visualizations. By linking the visualizations, dashboards help users reduce their working memory affordance and increase the efficiency of visual analysis.

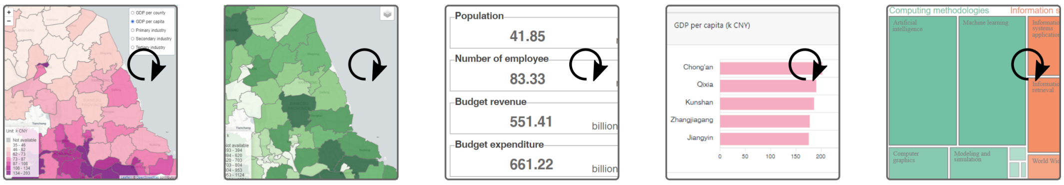


Figure 4.12: The linked views in dashboards across various visualizations.

Understanding social environments goes beyond the identification of the data patterns. The high-level insights, such as how the local research institutes collaborate with the international institutes and in which research domains, require users to relate several visualizations in their analysis. The designed arrangement of the visualizations can guide users to relate the factors better and think through the social environment. However, the layout design was not addressed enough in previous studies.

Designing interface of a map-based dashboard has to take various elements into consideration, including annotations, layout, and color. We studied the multi-panel dashboard design rules through an eye-tracking experiment [ZDM20]. Our previous study [ZDM20] shows that 1) clear division between the panels, 2) concise titles of the panels, 3) a logical arrangement of the panels, and 4) a unified color scheme and the font style are the keys for designing an easy-to-use dashboard. To be more specific, each panel in a map-based dashboard should have a single and clear purpose. For instance, showing the background information, showing spatial distribution, or supporting user input. The panels should be enclosed with clear borders so that users can easily distinguish various panels. For each panel, it should be named with a title that describes the expected insights, with which users can navigate easily to the needed information. The results show that the layout of the panels should follow the common reading habit. For instance, the panels are arranged from general to details, from a small scale to a large scale. The panels should also follow a united color and interaction style. Each panel should present a single perspective of the data or a single function. To better guide users to explore the data visually, the labeling of the panels should clearly express what types of information it conveys. The font plays a very important role in dashboard understanding. The font size should be big enough to read at a glance. Furthermore, the arrangement of panels should follow a logical order. The panels with similar content should be placed adjacently. When visualizing multi-granularity or multi-temporal data in several panels, the sequence of the panels should follow a common reading pattern, such as from large to small or from historical to present.

Designing the arrangement of the panels has to take users' reading approaches, the purpose of the dashboard, and the types of visualization into consideration. We propose the importance and the information abstraction as two references to conceptualize the disposition of panels in dashboards. In addition, the size of panels also influences the users' insight-gaining process: The bigger a panel is, the more attention of user will be led. Figure 4.13 shows an exemplary arrangement of a map-based dashboard, and it contains eight panels, namely a title panel (T), an information panel (I), two selection panels (S1 and S2), and four visualization panels (V1 – V4). The abstraction level of the information is organized from general at the top to detail at the bottom. To be more specific, the title (panel T) of map-based dashboard and the user selection zones (panel S1 and S2) contain general information and are located at the top, and the visualizations (panels V1–V4) contain more detailed information and are located at the bottom part. Also, the arrangement of panels V1–V4 follow the same logic. The spatial granularity follows the order from coarse to fine along with the reading flow from left to right, and the non-spatial granularity of factors follows the order from holistic to focus. The size of the panels is related to the importance of the represented information to users and the complexity of the visualization. For instance, spatial information is the most important and fundamental information in the social environment. Therefore the map panel can occupy a large area.

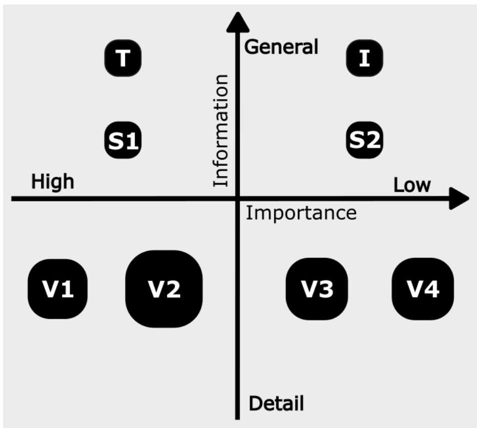


Figure 4.13: The conceptual design of dashboard layout in adherence to the reading flow of users.

4.5 Usability and User Experience

We design a systematic evaluation method to verify if a designed map-based dashboard can support users in acquiring spatial insights. The think-aloud method, eye-tracking, and interviews are used to identify the analytical-supporting ability, the visual elements, and users' attitude towards the design. Thus, the design issues in data analysis, visualization methods, interaction design, and layout design can be identified and improved according to users' feedback.

The user study consists four major phases - introduction, free exploration, task solving, and questions, as shown in Figure 4.14. During the *introduction*, the background of the experiment, the requests of the participants, and data privacy information are communicated to experiment participants. In addition, a standard instruction of map-based dashboards should be given to the experiment participants. In the *free exploration* step, the participants may freely try with the interfaces and ask usage-related questions. The *task-solving* step usually requires the participants to conduct pre-defined benchmark tasks. Finally, the participants are asked to give their feedback regarding the their experience in the *questions* step, such as their attitudes and suggestions.

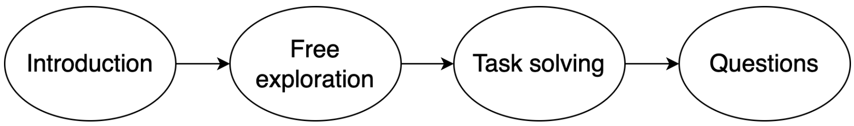


Figure 4.14: The general procedure of user studies to test map-based dashboards.

Insight acquisition. Since the map-based dashboards are designed to support users in acquiring spatial insights, it is important to know what insights can be acquired by the dashboard users. This can be analyzed by asking users to freely explore the dashboards and reporting their discoveries by the think-aloud method. The reported insight can be analyzed by categorizing them by cognitive effort. In the context of spatiotemporal analysis of social environments, the cognitive effort differs in *search area (single - multiple)*, *search time (single - multiple)*, *search attribute (single - multiple)*, and *query type (identification, comparison, summarization, deduction)*. The number and frequency of the reported insights among different user groups, such as age, gender, professions, and dashboard usage experiences, can be analyzed to analyze the ability of dashboards.

Visual Attraction. Knowing the visual attraction of the dashboards can help the dashboard designers to adjust the visual design accordingly. The visual attraction of the interface can be detected by analyzing dashboard viewers' eye movement. Furthermore, multiple visual elements could attract viewer' visual attention. They can be accessed by asking the viewers after they explored the dashboards. During the first few minutes of viewing the dashboards, the eye fixation of viewers could reflect the visual attraction of the visual elements of the dashboards. Therefore, we record the eye fixation of the viewers during their free exploration with eye-tracking devices. An example of gaze distribution over a map-based dashboard is shown in a heatmap in Figure 4.15. More specifically, the heatmaps can be applied to visualize the visual attention in time series so that the attention of users along time can be analyzed.

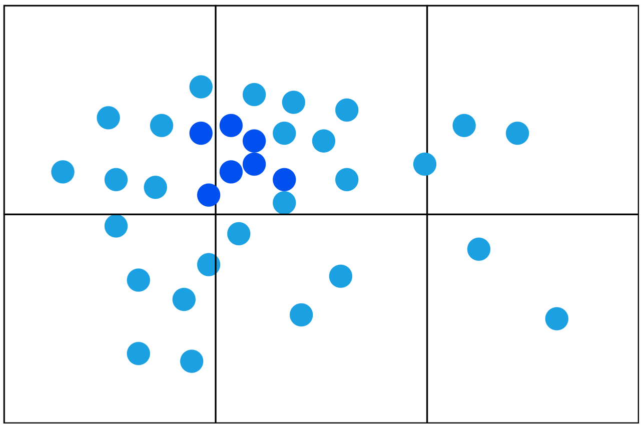


Figure 4.15: The conceptual heatmap of fixations on a map-based dashboard. The blue dots represent fixations.

Effectiveness and Efficiency. The effectiveness and efficiency of map-based dashboards can be reflected by the success rate and response time of conducting benchmark tasks. The benchmark tasks can be formulated according to cognitive efforts as shown in Table 4.3. The effectiveness and efficiency of solving the cognitively easy tasks can be used as a reference to check if the designed dashboard support the cognitively difficult tasks. Thus the dashboard designers can identify whether the visualization methods and data analytical methods should be adjusted to better fit to the design goals of certain visual analysis of spatiotemporal patterns.

Insight-gaining Strategy. Dashboard layout plays an important role in guiding dashboard users' reading strategy. Analyzing visual reading patterns of task-solving can help designers to improve the arrangement of the panels on dashboards. In our dashboards, each panel is designed with one focus, such as title, legend, spatial patterns, and certain panels are adjacently placed for comparison. Therefore, we mark them as areas of interests (AOIs) according to the functions of dashboard panels. Figure 4.16 shows the conceptual area of interests. The eye movement metrics of the AOIs, for instance the fixation sequences, dwell, and transitions, along the task-solving can be analyzed. According to [HNA+11], we listed the selected metrics of the eye movements in Table 4.4.

Table 4.3: The benchmark tasks of spatial insight acquisition a with increasing cognitive effort.

Cognitive Operation	Task Example
Identification	Find an attribute of a place.
Comparison	Find the attribute temporal trend of a place.
Identification and comparison	Identify a place with the highest attribute value.
Comparison and summary	Summarize the spatial distribution of an attribute.
Comparison and summary	Compare the spatial distribution of two attributes.
Comparison, summary, and deduction	Compare the spatial distribution of two attributes.

Table 4.4: The selected eye movement metrics.

Metric	Description
Sequence	The order of fixation within AOIs.
Dwell time	The sum of all fixations and saccades within an AOI.
Transition	The movements from one AOI to another.
Return	The transition to an AOI itself, also known as a revisit.

Figure 4.16: The conceptual area of interests (AOIs) according to the functions of each dashboard panel.

Users' feedback. Subjective opinions and suggestions can be obtained by asking users open questions in questionnaires or interviews and help to improve the interface from the users' point of view. To be specific, the general performance of the interface, specific visualizations, and specific interactions, users' likeness of the design, users' attitudes, and suggestions and comments can be collected as listed in in Table 4.5. By conducting the semi-structured interviews, users are guided to give feedback on various items.

Table 4.5: The questions to investigate the users' feedback on map-based dashboards.

Category	Questions
Usability	How do you rank the overall ease of this visualization interface? How do you rank the learnability of the visual interface? How do you rank the understandability of the visualizations? How do you rank the interactivity of the interface?
Design issues	Please list the design elements that helped the most during the task conduction. Please list the design items or elements that were not easy to understand or interact with. Do you think the information are sufficient enough to assist you in fulfilling all tasks? If not, what else would you like to add? Do you think the visualization methods are suitable to perform the benchmark tasks? If not, why?
Users' attitude	How do you rank the visual attractivity of the interface? How do you rank your trust into the interface? How do you rank the confidence of your findings based on the interface? How do you rank the entertainment of the interface?
Suggestions and comments	What would you suggest improving this visualization interface? If you have some additional comments, please write it down.

Users' feedback can be grouped into six categories, e.g., visualizations, layout, annotation, annotation, interaction, content, and aesthetics, as Table 4.6 shows. In this way, the feedback could be integrated into the iterative design of dashboards.

Table 4.6: The categories of users' feedback to the design issues.

Category	Items
Visualizations	Visualization method, color scheme, visual attributes
Layout	Arrangement of panels, margin
Annotation	Contextual information, font size, label, legend, axis
Interaction	Highlight, reorder, searching, selection, pop-up window, reset, help information
Content	Data, language
Aesthetics	Attraction, appearance
User experience	Learning time, mental workload

The purpose of visualization is insight, not pictures.
S. T. Kard, J. D. Mackinlay, & B. Scheiderman,
Readings in Information Visualization, using vision to think., 1999 (p. 6)

5

Dashboard Case Studies in Social Environment Analysis

Three case studies were conducted to evaluate the proposed design framework of map-based dashboard and illustrate the map-based dashboard design in practice. Three dashboards were designed to represent different social environments with different datasets:

Case study 1: InDash [ZDL+22] was designed to show the overview of the industrial innovation environment to the decision-makers in industry. Users could learn the overview of the environment and compare the important factors in multiple cities. The statistical data was collected from yearbooks. Moreover, the effectiveness in sight acquisition was evaluated with the think-aloud method. The *challenge* of this case study is how to integrate heterogeneous data on vairous topics and represent them logically on a single-screen dashboard.

Case study 2: EconDash [ZDM20] was designed show the differences of the socioeconomic environment in multiple municipalities to the entrepreneurs. Users were enabled to compare and relate the important factors in the socioeconomic environment. Similarly, the statistical data was used in this case study. In addition, eye-tracking method was applied to examine the visual attraction of the interface and externalize the viewing strategies of users. The *challenge* of this case study is how to trigger the viewers to visually compare the socio-economic environments among many municipalities.

Case study 3: KnowDash [ZDYM22] was designed to show the spatial network pattern of the knowledge innovation environment to the scholars. It could help users to learn the popular research domains and study the research collaborations at multiple levels. The big scholarly data was collected from an online research database and used to reveal the research networks. Furthermore, the users' attitude towards the interface was collected and analyzed. The *challenge* of this case study is how to efficiently represent the complex spatial network patterns and allow viewers to acquire detailed information on their demands.

More specifically, the three use cases are showing the analytical ability of map-based dashboard with different focuses and datasets. InDash and EconDash used the statistical data from yearbooks, while InDash shows the data at city level and focus on supporting learning and overviewing and EconDash shows data at county level and focus on comparing and relating. KnowDash used social network data and show it at three different levels to enable users to analyze spatial distribution of research domains and networks. InDash and EconDash took Province Jiangsu, China as the study area, which covers 107,200 km2 and 13 cities (containing 98 municipalities). KnowDash took Yangtze River Delta (YRD) region as the study area and it covers Shanghai, Jiangsu, Anhui, and Zhejiang. Figure 5.1 shows the location of the study areas.

Figure 5.1: The study areas of the three case studies.

5.1 InDash for Learning Industrial Innovation Environment

Industrial innovation is a driving force for regional economic growth and corporate competitiveness [FS97]. An industrial innovation environment (IIE) involves different aspects that support or impose new demands on innovative capabilities, such as economic dynamics, socio-cultural conditions, and environmental sustainability [Har04, Mor13]. Understanding and communicating the nature of IIE is an essential task for managers, public policy makers, corporate decision-makers, and the students of industry and business [DR95]. Understanding the industrial innovation environment requires learning and reasoning about the related factors, such as reading and comparing socioeconomic indexes [DLWV19], analyzing spatial distributions of relevant factors [MS14], and inferring spatial correlations [CV17, Lee05]. A previous study [Fis11] pointed out some challenges in environmental decision-making, such as how to obtain a holistic view of different perspectives while interpreting complex analytical results.

In this case study, we designed a map-based dashboard called InDash for the target users of entrepreneurs using open government data collected from yearbooks. 24 factors were selected to reflect the economy, inhabitant, infrastructure, and R&D. The purpose of InDash was to present the related factors in different categories, support users in learning the KPIs, and provide basic comparisons of the multiple factors. We designed a symbol map, a KPI board, a correlation matrix, and parallel coordinates to reveal four selected aspects of IIE. Furthermore, a user study with 30 participants was conducted to assess the ability of InDash in supporting users to gain knowledge of IIE.

5.1.1 InDash Interface Design

The graphic interface of InDash is shown in Figure 5.2, including the title and the background information in the title panel (T), the user interaction zones (S1 and S2), and the four data visualization panels (V1 – V4). More specifically, the different panels supported users to read and reason about the IIE from different perspectives and at different levels of detail. Panel T showed the title and the data source, which gave users the background information and makes them be aware of the context. Panel S1 and S2 showed all the available IIE categories and the names of the displayed cities, and they allowed users to select the categories and highlight different cities. Panel V1 – V4 showed details of the IIE, such as the factor values, their spatial distribution, and the correlations between factors. Panel V1 showed a map containing the designed charts serving as an anchor point for the users to understand the spatial information. In Panel V2, we designed a KPI board (Key Performance Indicators) to show the aggregated values of IIE factors. In Panel V3, we showed the correlation among the factors in a matrix heatmap, where the correlations between factor pairs were represented in a color scale from positive (blue) to negative (red). In Panel V4, we applied parallel coordinates to show multiple values of factors at the city level. In InDash, each panel showed relative independent content of IIE, and all the panels collectively serve users for their visual learning and reasoning. The web system of InDash is available at `http://129.187.45.33/InDash`.

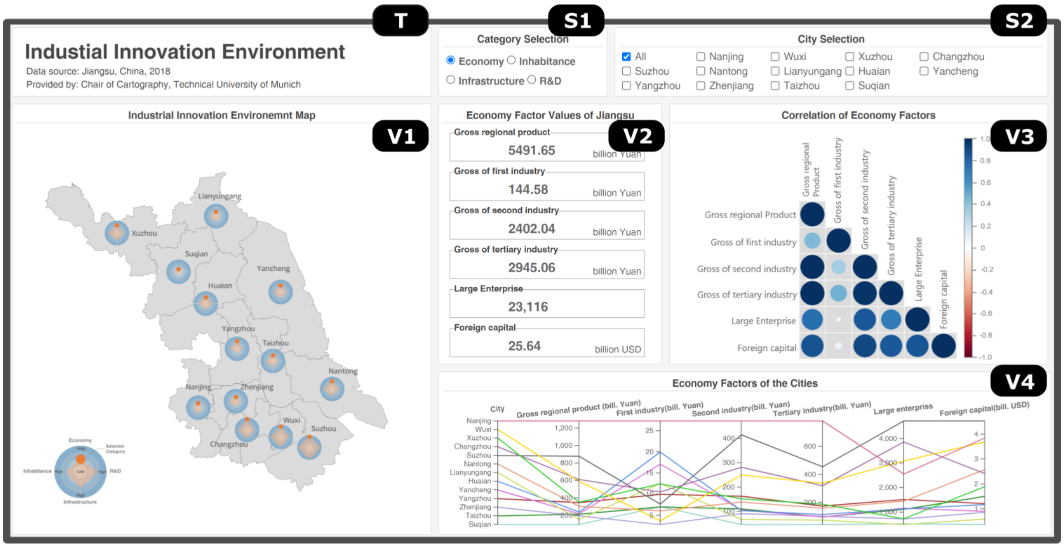


Figure 5.2: The visual interface of InDash with seven panels.

Domain value analysis. Early studies [Har04] suggested that regional IIE patterns should be understood to support effective policy-making. Using InDash could support the learning of the local IIE easily. For example, from the radar charts on the map in Figure 5.2, we could easily identify that the IIE in the south is, in general better than in the north with regard to multiple aspects. Economically stronger cities, such as Nanjing, Suzhou, and Xuzhou, had more balanced and higher values of the indexes in the IIE, whereas economically weaker cities, such as Suqian, Zhengjiang, and Taizhou, had higher values in inhabitant and infrastructure than in economy and R&D. Besides the spatial distribution, we could also easily interpret the statistics from Panel V2 (as shown in Figure 5.3). The secondary and tertiary sectors of the economy counted for about 97% of the regional gross, while the foreign investment took only about 3%. The gross regional product per capita was about 131,210 Chinese Yuan. The number of employees were about doubled the population. The bus rate per 1,000 population was 1.05, which means high-capacity public transportation. The number of granted patents per million population was 9,780, which is very high compared to the values of other regions.

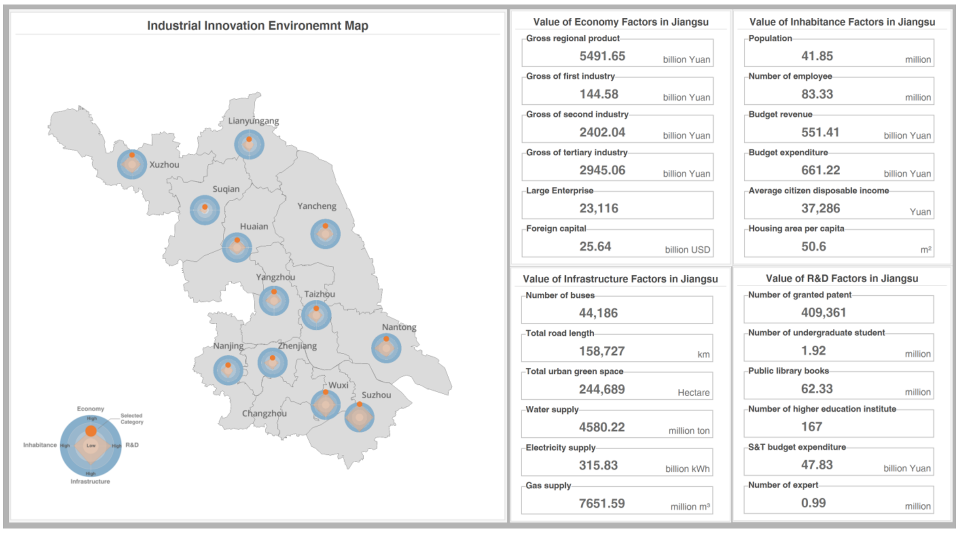


Figure 5.3: InDash panels with an overview map and statistical values of six factors in each of the four categories.

Comparing multiple factors among cities. A recent study [WHY20] investigated the economic inequality among the southern and northern cities in Jiangsu. Their findings could be visually perceived and validated using InDash. The overview of the regional patterns could be perceived from the map in InDash V1, as shown in Figure 5.2. The values of factor could be further visually analyzed with the parallel coordinates in InDash V4, as shown in Figure 5.4. We took the southern cities Nanjing and Suzhou and the northern cities Huaian and Suqian as analysis examples. Nanjing and Suzhou led in various aspects of the whole region. Although Nanjing had a higher GDP value, Suzhou was stronger than Nanjing in its industry with more large-sized enterprises, attracted more foreign capital, higher disposable income per capita, and more budget for R&D. Huaian and Suqian lagged in the economic growth in Jiangsu Province. The major contribution of their gross came from the secondary industry and tertiary industry. Compared to Nanjing and Suzhou, their average housing area and road length were equivalent, while other industrial innovation factors had low values.

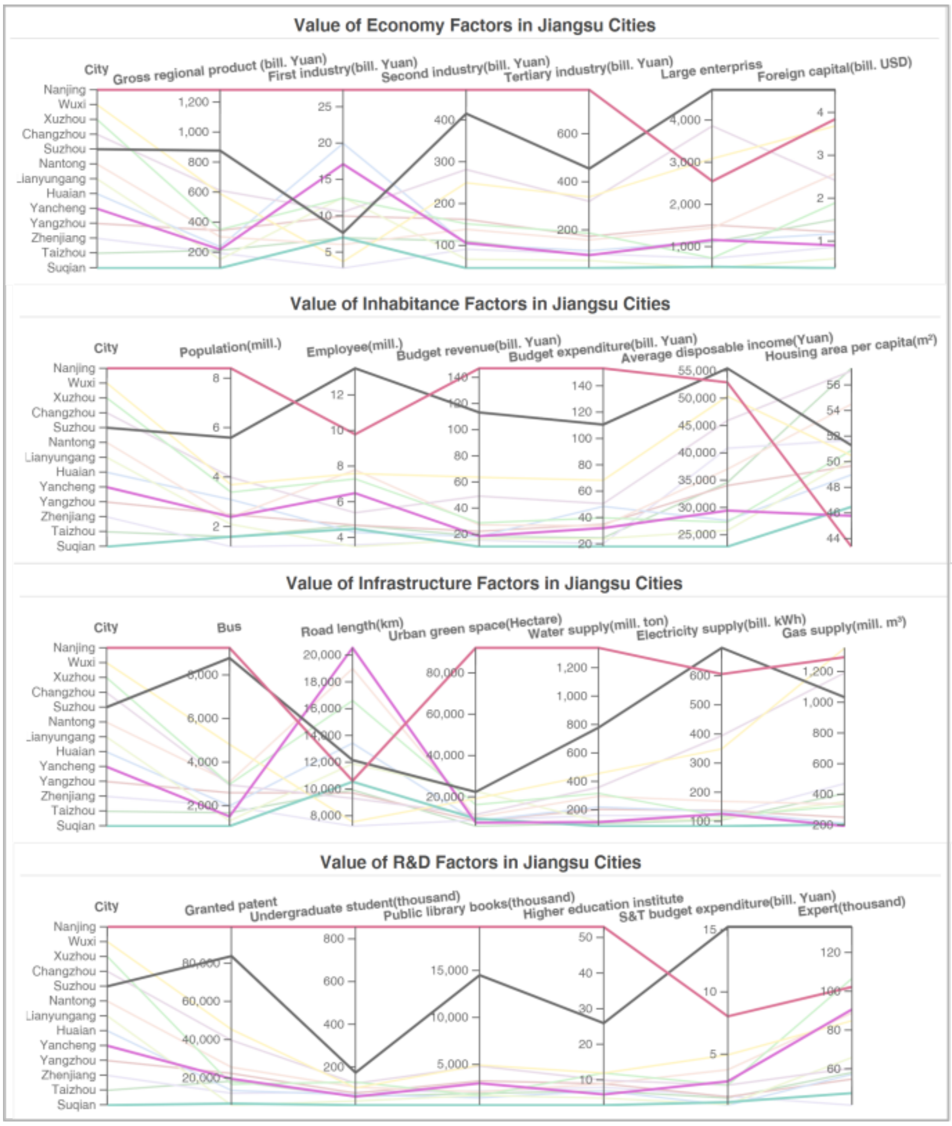


Figure 5.4: Comparison of the selected cities in the parallel coordinate presentation of six factors in each of four categories.

5.1.2 Evaluation of Insight Acquisition with Think-aloud Method

To evaluate the effectiveness of InDash in supporting IIE learning and reasoning, we conducted a think-aloud experiment with 30 participants with various profiles. The experiment collected the knowledge about participants when they freely explored InDash without intensive training. The aim of the evaluation was to understand what users can learn and discovery, and examine whether the proposed graphic interface can guide users to meet the design requirements.

The experiment was conducted in January and February 2021 and remotely online performed by the participants. The 30 participants were composed of 16 females and 14 males, 27 of them were between 18 and 30 years old, and three were between 30 to 50 years old. They were mostly students with various study backgrounds. Three of them were staff members in different companies, and one participant was a decision-maker. Three of them had never used a dashboard, 22 participants had limited usage experience, and five of them were familiar with dashboards. Among the participants, seven of them lived in the study area, 16 of them had visited the area, and seven of them had never been there.

The evaluation was proceeded one by one with the participants. Each participant was firstly given a three-minute introduction consisting of the background information of this evaluation and a short description of InDash. Then the participants were asked to explore the InDash displayed on a laptop freely and encouraged to speak aloud about their findings from InDash. The time for the free exploration was not limited, and the participants could end this session at any time.

With the aim of understanding what insights users can acquire, we extracted the insights from the free exploration session and further divided them into groups based on their characteristics. Five insight groups are identified, including fact, min/max, comparison, relation, and distribution, as shown in Table 5.1. Compared to the previous study [YEB18] about insight coding, we highlighted the spatial-related insights as distribution.

Table 5.1: The description and examples of defined insight groups.

Insight group	Description	Example
Domain value	The records that can be identified directly.	The population of Jiangsu was 41.85 million.
Min/Max	The extreme values of a set of data.	The number of granted patents in Nanjing was the highest.
Comparison	The description of two or more values, places, or topics.	The social environment factors of Suzhou were more balanced than Suqian.
Correlation	The relation across attributes	The GDP was strongly and positively correlated with the gross domestic product of secondary industry.
Spatial distribution	Location-related patterns	Xuzhou was the strongest city in north Jiangsu.

We compared the number of reported insights across three characteristics of the participants: different genders, experiences of the study area, and experiences of dashboard usage. Due to the small sample size of the group without dashboard usage experience, we analyzed the results using the median test. The median number of acquired insights and the frequency of different groups are shown in Table 5.2 and the distribution of the statistics are shown in boxplots in Figure 5.5. The results show that the median of total number of insights is 20, and the median duration is 21.4 minutes. It means that in general the participants were interested in reading InDash and acquired insights without intensive effort. To be specific, more insights about comparison and correlation were reported than the insights about facts and min/max, whereas only a few insights about distribution were reported. The results do not show a significant differences between different genders and people with different experiences of dashboard usage. However, the more background knowledge the participants had about the study area, the larger number of insights and the more complicated insights the participants can acquire. The insights related to spatial distribution is relatively hard for users to acquire, and therefore the number of such insights is not much reported by the users. They require the users to compare or correlate different values in multiple areas and in multiple environments.

Table 5.2: The median number of duration, frequency, and count of acquired insights from the participants.

Category	Group	N.	Duration	Frequency	Median count of insights					
					Fact	Extremum	Comparison	Correlation	Distribution	Total
Overall	Overall	30	21.4	1.0	1	5	7	6	0	20
Gender	Female	16	21.9	0.9	1	5	7	7	0	19
	Male	14	19.8	1.2	1	6	9	5	0	22
Test	Live	7	22.6	0.9	1	6	9	7	1	23
Area	Visited	16	21.4	1.0	1	6	8	6	0	20
Experience	None	7	17.5	0.9	3	2	3	5	0	15
Dashboard	Familiar	5	24.8	1.1	1	6	3	10	0	19
Usage	Used	22	20.6	1.0	1	5	7	6	1	21
Experience	None	3	21.7	0.8	0	3	9	5	0	18

N stands for number of participants.

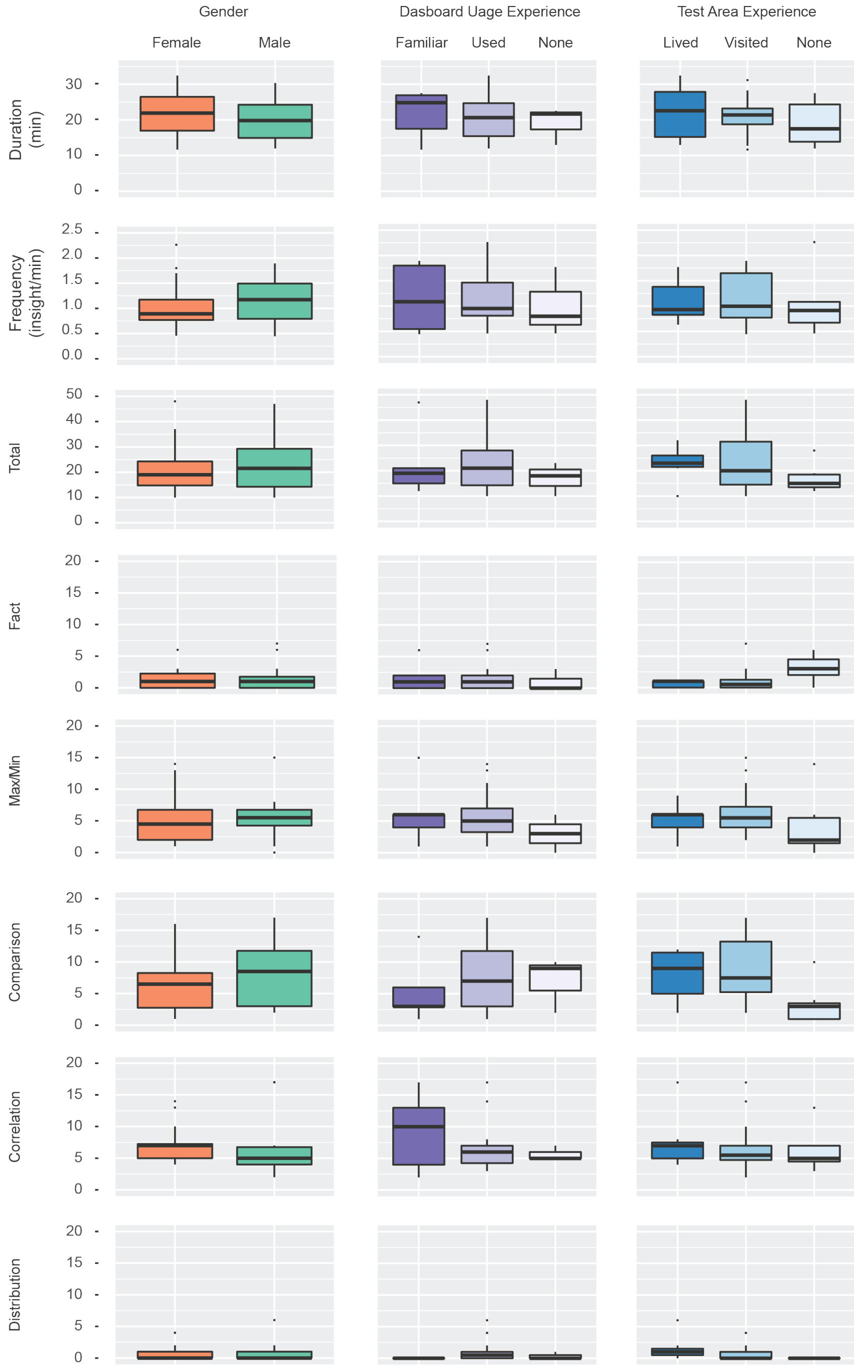


Figure 5.5: The boxplots show the duration of the exploration and frequency and count of the acquired insights across gender, dashboard usage experience, and test area knowledge.

5.2 EconDash for Comparing and Correlating Socioeconomic Environment

Understanding and communicating the nature of the socioeconomic environment is an essential task for corporate decision-makers and other stakeholders [DR95]. It requires learning and reasoning about the related factors, such as reading and comparing socioeconomic indexes [DLWV19], analyzing spatial distributions of relevant factors [MS14], and inferring spatial correlations [CV17, Lee05]. Literature related to the socioeconomic environment shows that a socioeconomic environment involves different aspects that support or impose new demands on innovative capabilities, such as economic dynamics, socio-cultural conditions, and environmental sustainability [Har04, Mor13].

In this section, we demonstrate the map-based dashboard EconDash. It is designed to support the comparing and correlating the socioeconomic environment using open government data collected from yearbooks. Four synchronized choropleth maps and multiple bar charts are designed to show the spatiotemporal distribution of the factors. We evaluated the visual attractiveness of the panels on EconDash and the viewing strategies of participants in discovering spatial knowledge. We collected the eye movement data of 39 participants while they were freely exploring EconDash and performing pre-defined knowledge discovery tasks.

5.2.1 EconDash Interface Design

EconDash consists of eight panels, as Figure 5.6 shows, the title panel (T), the toolbar panel (i), the spatial panels (M1 – M4), the temporal panel (B1), and the ranking panel (B2). The eight panels display the socioeconomic environment from multiple perspectives. Panel T shows the title that aims to make users aware of the topic. Panel I allows users to reset the dashboard and shows explanations for the various factors. The panels M1 – M4 present the spatial distribution of multiple factors on maps. Each map shows a socioeconomic category, i.e., enterprise, GDP, population, and logistics. There are several layers within each map, and each layer shows one factor. They provide a spatial overview of the multiple factors, show detailed values in pop-up windows, juxtapose multiple maps, and support search and highlight functions. Panel B1 shows the historical data of a selected county. Panel B2 shows the top-ranked counties in each active map layer. The visualizations are linked and updated according to users selection, and serve together to support users' visual analysis of value reading, comparison, correlating, and overviewing. EconDash is available at `http://129.187.45.33/EconAnalyticalDash/`.

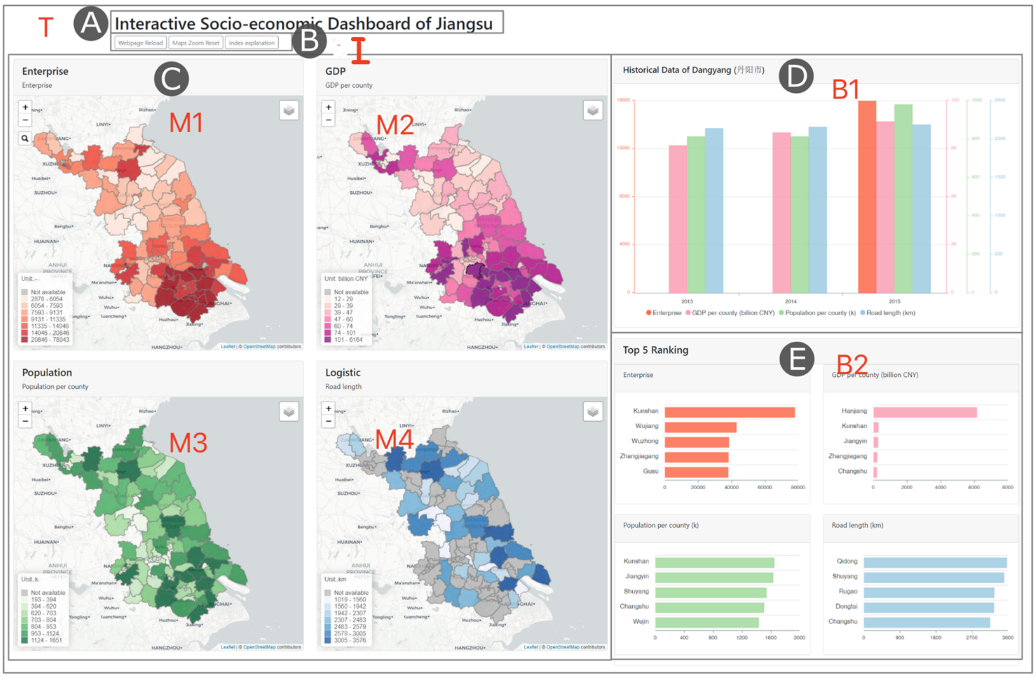
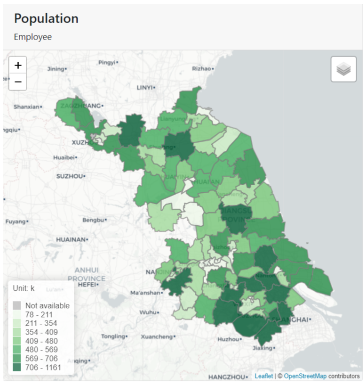
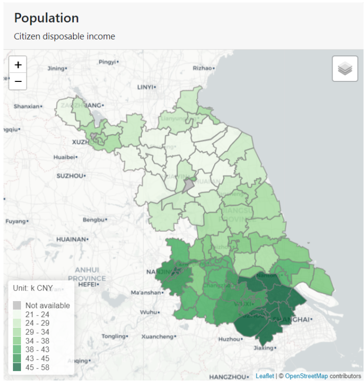


Figure 5.6: The panels of the dashboard design. The panels are labeled as: (T) the title panel, (I) the toolbar panel, (M1 – M4) the spatial panel, (B1) the temporal panel, (B2) the ranking panel.

Spatial correlation. Correlation between factors was analyzed in many previous studies. For example, Salim et al. [SRH+20] analyzed the correlation between the factors of employment and income. Similarly, EconDash could support the visual comparison of the number of employees and citizen disposable income in Jiangsu, as Figure 5.7 shows. We could see that the number of employees varies from county to county, while the citizen disposable income shows a spatial pattern that is much strong in southern part of Jiangsu. The spatial distribution of the number of employees and citizen disposable income did not show a strong correlation.



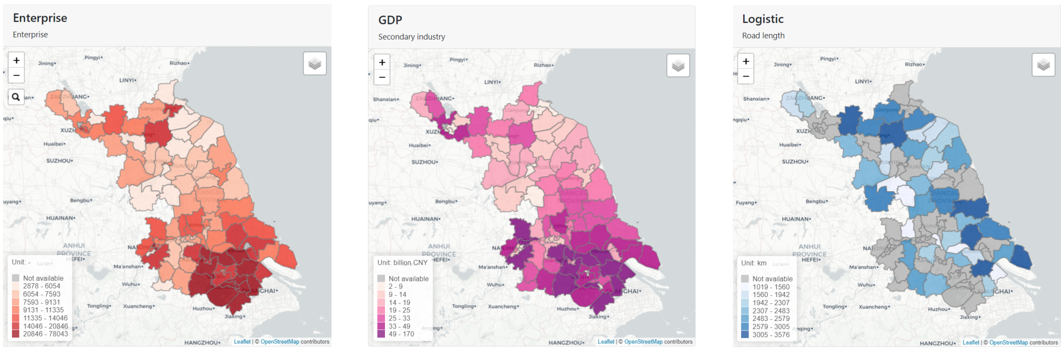
(a) Number of employees



(b) Citizen disposable income

Figure 5.7: The spatial distribution of the number of employees and citizen disposable income.

Spatial prediction. The correlation between the factors and general concepts such as industrialization could be deduced. For example, Figure 5.8 shows the spatial distribution of enterprises, secondary GDP, and road length. We could see that the spatial distributions of enterprises and secondary GDP in Jiangsu share a similar pattern, with higher values in the southern part than the northern part of Jiangsu and higher values in the northwest than the northeast. The values of road length were largely unavailable, but we could still perceive a pattern that does not reveal higher values in the southern part. We could preliminarily predict that the industrialization level in the south part is higher.



(a) Number of enterprise **(b)** Secondary GDP **(c)** Road length

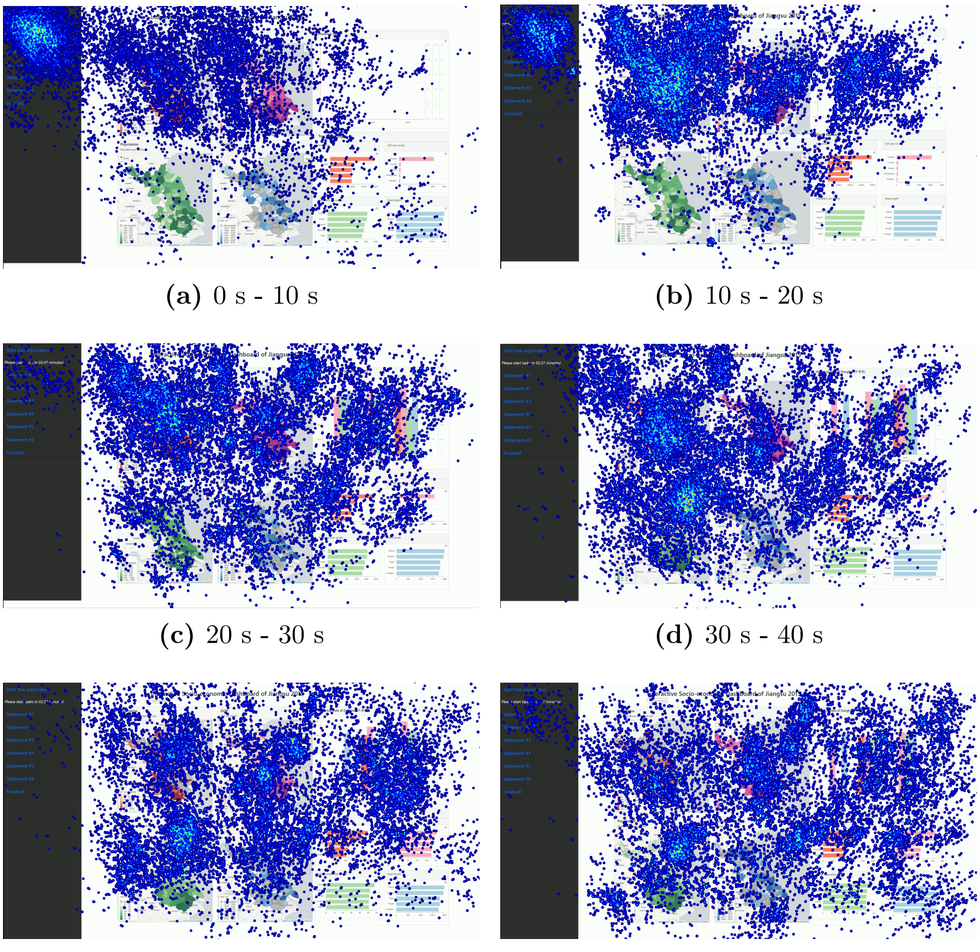
Figure 5.8: The spatial distribution of industrialized level with the number of enterprises, secondary GDP, and road length.

5.2.2 Evaluation of Layout Design with Eye-tracking Method

The experiment was conducted in November and December 2019 in an eye-tracking lab. We recruited 40 volunteered participants with the means of short introductions in university lecture rooms, posters on campus, and online advertisements. One of the participants quit because of an eyesight problem. The remaining 39 participants had normal or corrected-normal eye sights and completed the experiment smoothly. After the experiment, we found that the eye-tracking ratios of seven participants were less than 70% and could not be considered. Thus, the analysis was based on the recorded eye movement data from the remaining 32 participants. Among the 32 participants, there were 17 females and 15 males. Their average age was 25.9, with a standard deviation of 2.36. They had diverse educational backgrounds: one participant with a high school or equivalent degree, 17 participants with bachelors' degree, and 14 participants with masters' degree. They had various usage experiences with interactive dashboards: 12 participants had used dashboards more than five times, 5 participants had used less than five times, 8 participants had heard about it but not used, 7 participants had never heard about it. None of the participants had ever lived in the study area.

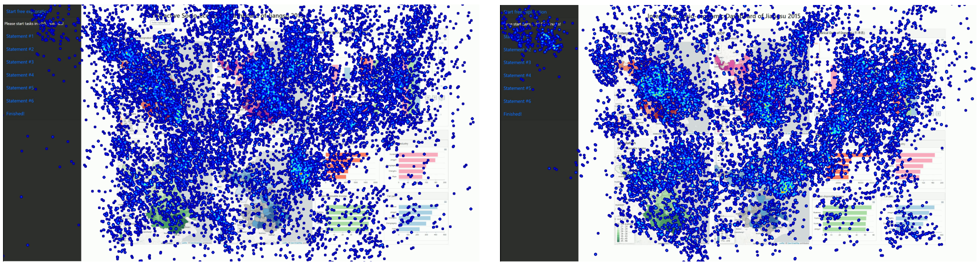
Participants were asked to explore the interface for three minutes freely. Their eye fixations during the exploration were collected, showing the fixation distribution in the first 80 seconds at the free exploration stage on heat maps (Figure 5.9). Considering the amount of our sample, we chose 10 seconds as the interval to include enough fixations in forming clusters in each interval. The heatmaps exhibit different patterns of the fixation distribution in 0 - 20 seconds (Figure 5.9a and 5.9b) and 20 - 80 seconds (Figure 5.9c - 5.9h). In the first 10 seconds (Figure 5.9a), we can see that the fixations were mostly on the title, task area, and top-left area. Between 10 to 20 seconds (Figure 5.9b), the fixations scattered as well over the map on the top-middle and other maps. From 20 seconds onwards (Figure 5.9c - 5.9h), the attention of the participants was located more evenly on each dashboard panel. In general, the panels located in the center won more attention than other panels at the free exploration stage. For example, the maps in the top-left area were focused at the beginning of the exploration. Besides, the dynamic visualizations attracted much attention. For example, the bar chart drew much attention, as the chart in the panel changes when there was a mouse hover or click event. Additionally, much attention went to the task panel in the beginning because the participants needed to click and read the task items. The author inferred that bright colors also play an important role in attracting users' attention. Last but not least, anomalous patterns, such as incomplete data and outliers, also attracted the participants' attention. In the 0 -

30 seconds, more fixations were at the *Panel Logistic* than *Panel Population*, where a large gray area indicated of unavailable data.



(a) 0 s - 10 s
(b) 10 s - 20 s
(c) 20 s - 30 s
(d) 30 s - 40 s
(e) 40 s - 50 s
(f) 50 s - 60 s

Figure 5.9: The heatmaps of all the participants' fixations during at free exploration stage. The blue dots represent fixations. (Cont.)



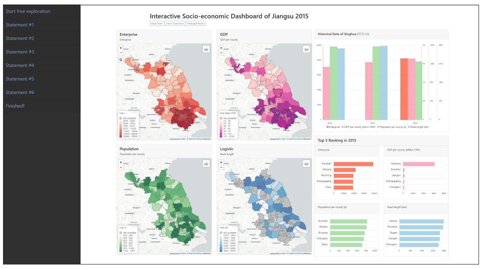
(g) 60 s - 70 s **(h)** 70 s - 80 s

Figure 5.9: The heatmaps of all the participants' fixations during at free exploration stage. The blue dots represent fixations.

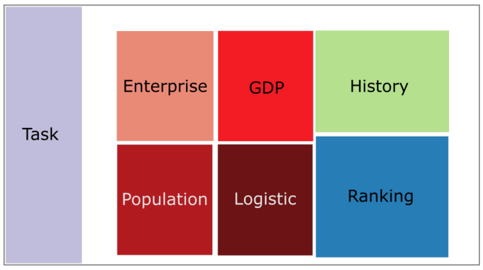
The fixation sequence of AOIs revealed the viewing strategies of the participants in insight acquisition. We asked the participants to find the insights as shown in Table 5.3 and visualized the fixation sequence in Figure 5.10. In general, the success rate of correlation analysis was higher than correlation deduction, the time used for correlation analysis was shorter than correlation deduction, and the participants used more panels in correlation deduction than correlation analysis. The reason might be that correlation deduction is harder than correlation analysis and it requires more background knowledge. Moreover, the adjacent panels were more likely to be read in sequence.

Table 5.3: Six exemplary statements and their possible answers.

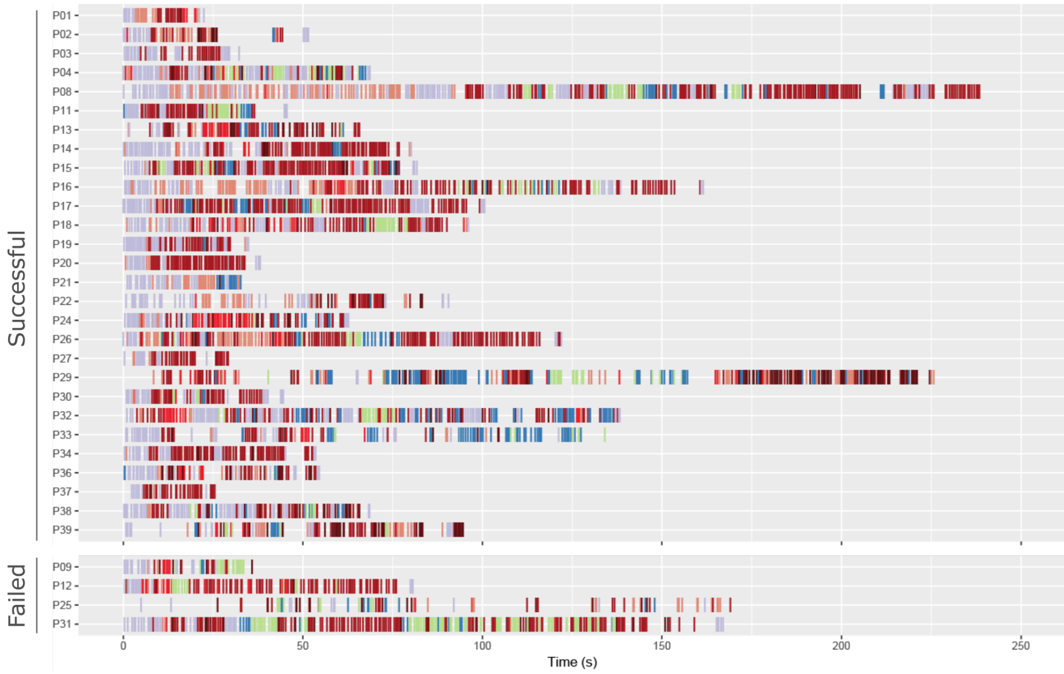
Task number	Statement	Answer
Task 1	In 2015, the Tertiary Industry value of City A was 85 billion dollar.	Wrong
Task 2	The number of enterprises in City B increases from 2013 to 2015.	Unknown
Task 3	In 2015, among all the counties in State A, County A had the largest number of enterprises.	Correct
Task 4	The south part of State A is economically stronger than the north part.	Correct
Task 5	In State A, the more employees in a county, the higher the citizens' disposable income is.	Wrong
Task 6	In State A, the longer the total length of the road of a county, the higher the industrialized level is.	Wrong



(a) EconDash interface



(b) Legend of the AOIs



(c) Visual sequence during spatial correlation

Figure 5.10: The sequence charts of eye movement during the insight acquisition. (Cont.)

The participants were asked to list the design elements that helped them during the task-solving procedure. We grouped the feedback into four groups: panel, layout, interaction, and others. Table 5.4 summarizes the positive feedback with the associated frequency in each group. In the panel group, the most helpful panel is the spatial panel (nine mentions). In the layout group, the color scheme is very helpful in supporting the participants in organizing the information. In the interaction group, the search function is the most frequently mentioned item of helpfulness.

(d) Visual sequence during spatial prediction

Figure 5.10: The sequence charts of eye movement during the insight acquisition.

Table 5.4: Grouped positive items mentioned by the participants in the interview.

Group	Item	Frequency
Views	The spatial panel is helpful	9
	The temporal panel is helpful	6
	The ranking panel is helpful	2
Layout	The color scheme helped in organizing information	13
	The grouped layers helped in factor finding	4
	The juxtaposition benefits comparison	3
	The structured design gives a good overview	2
Interaction	The search function is useful in finding places	15
	The interactions of the temporal panel help them find data quickly	6
	The pop-up windows are helpful	5
	The layer switching function is efficient	2
Others	Natural to use	1

Similarly, we grouped the negative design items named by the participants in the interview. Table 5.5 shows the items in detail. Compared to the positive items, the negative ones are more related to specific issues. The most frequently mentioned items are: the top margin of the bars in the temporal panel is sometimes too narrow, the shifting of the maps is disturbing, and the font size is too small. It is noticeable that three participants though the temporal panel was too informative, and one participant pointed out the temporal panel did not follow the four category structure as other panels. Two participants also raised the problem that they did not know where to look for the needed information.

Table 5.5: Grouped negative items mentioned by the participants in the interview.

Group	Item	Frequency
Penal	The top margin of the bars in the temporal panel is sometimes too narrow	5
	The maps shift when the mouse moves close to their boundaries	5
	The temporal panel is too informative	3
	The legend intervals are confusing	3
	The axes in the temporal panel change their ranges	1
	The axes in the temporal panel are not necessary	1
	The temporal panel should be split into four charts as other panels	1
	Mark the important places in the spatial panel	1
Layout	Hard to compare two layers in one map	3
	The color scheme is not good for color-blinded people	3
	Only one map in the spatial panel is preferred	2
	The color hue should be increased in the temporal panel and ranking panel	2
	The color of the unavailable data should be lighter	1
Interaction	The search bar should be on each map / outside the spatial panel	5
	The ranking panel should be clickable	4
	The selected place should be highlighted on all the maps	2
	The map legends should be clickable	2
Others	The font size is too small	5
	The unavailable data increases the difficulty	3
	No idea where to look on the dashboard	2
	The dashboard is too informative	2
	The listing of the top five municipalities is not interested	1
	A learning time is required	1
	Lack of the economic background information	1

5.3 KnowDash for Analyzing Knowledge Innovation Environment

Knowledge innovation environment is one of the fundamental factors for regional development. Supporting knowledge innovation and collaboration becomes increasingly important in modern regional planning. Science and technology parks are good practices in encouraging knowledge innovations. 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 [HSK18]. This in turn attracts ideas, knowledge, innovation, population, and financial investment to the regions [CYGL14]. Therefore, fostering knowledge innovation, collaboration, and academic spin-offs are key strategies in regional planning and management [FS12, FY04]. Understanding knowledge innovation patterns and regional development is essential for policy and decision-making.

In this section, we demonstrate the map-based dashboard KnowDash. It is designed to support the exploration of spatial connections and gaps of various research domains. The data is collected from the public scholar database ACM Digital Library [1]. We designed linked treemap and flow maps to reveal the spatial connections and gaps. We evaluated the user experience of KnowDash with seven experts in cartography.

5.3.1 KnowDash Interface Design

We developed KnowDash to support users' understanding and analysis of the knowledge innovation patterns in terms of the spatial distribution of the institutions, the collaborations, and the popular domains. KnowDash is composed of eight panels as shown in Figure 5.11, including the title panel (T), the information panel (I), an area selector panel (S), the research domains visualized in treemaps (V1), and the spatial distribution of the academic network in the YRD region (V2), in the Chinese domestic area (V3), and the international coverage (V4). To be more specific, Panel V1 contains a treemap showing the popular topics in two levels. Each rectangle represents one topic, and the rectangles in the same color belong to a group. The size of the rectangle indicates the number of scientific publications on a topic. Panel V2, V3, and V4 represent the publications and the collaborations in joint publications in flow maps. The users follow the natural reading order from top to bottom (knowing the general background to the details) and left to right (from the innova-

[1] https://dl.acm.org/

tion domains to their spatial distributions). The interactive tool can be visited at `http://129.187.45.33/KnowledgeDash/`.

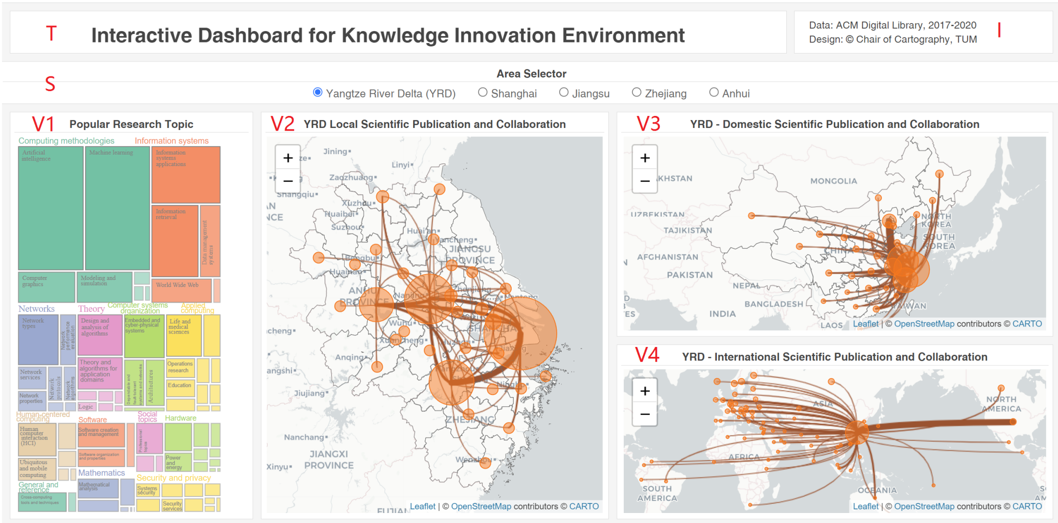
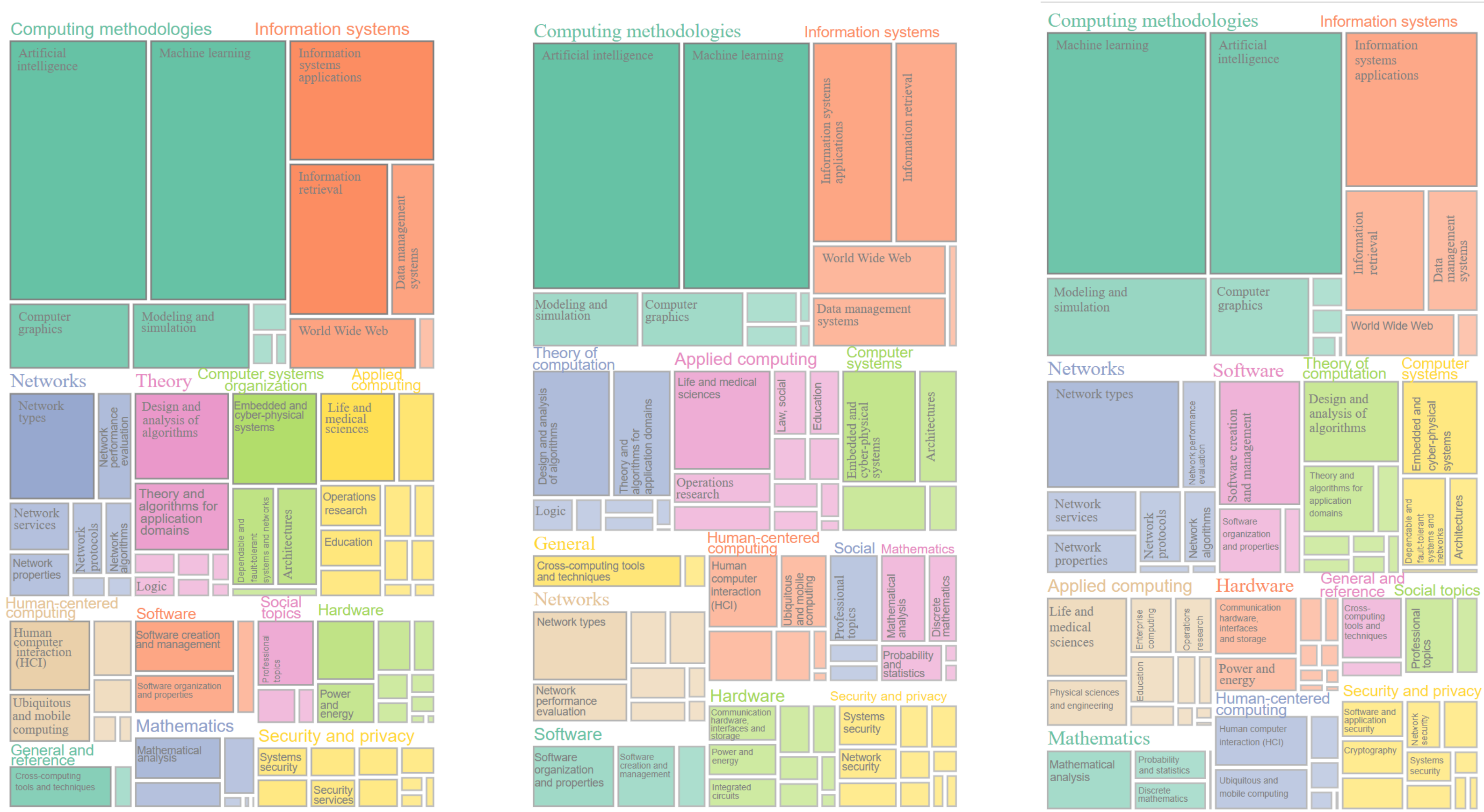


Figure 5.11: The visual interface of KnowDash. The panels are labeled as: (T) the title panel, (I) the information panel, (S) the selection panel, and (V1 – V3) the visualization panel.

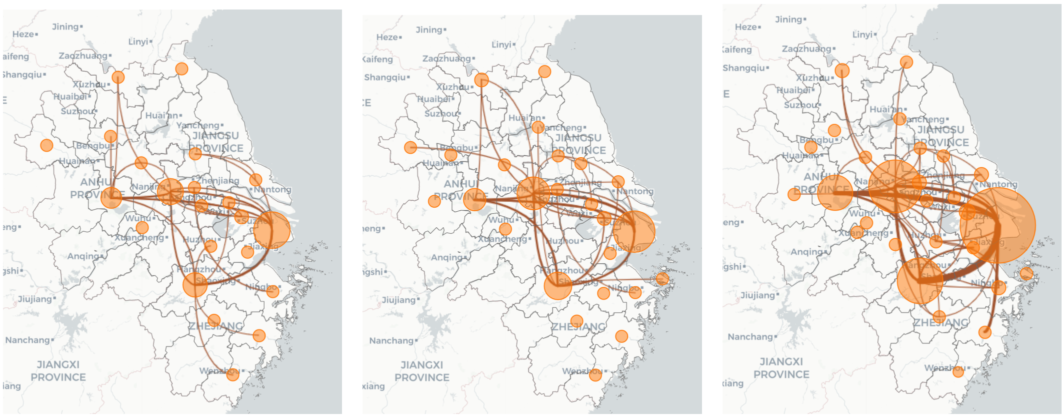
Analysis of popular topics. The regional competitiveness of knowledge innovation can be partially reflected in research domains of computing and information technology. We used treemaps to visualize the popular research topics in YRD, Shanghai, and Jiangsu as an example, as shown in Figure 5.12. The two most popular first-level topics were the same in three regions, and the two first-level topics took almost half share of all the topics. The rest first-level topics were diverse in those three regions. Within the most popular topic Computing methodologies, the two most popular sub-topics were Artificial Intelligence and Machine Learning in all three regions.



(a) YRD (b) Shanghai (c) Jiangsu

Figure 5.12: The treemaps showing the most popular research domains in YRD, Shanghai, and Jiangsu.

Analysis of spatial relations. We also analyzed the characteristics of the academic network based on prefectural cities. The number of publications was aggregated at the city level, and the number of co-authorship connections among the cities was calculated. We designed a flow map to illustrate the number of publications in each city and the collaborations among the cities, as shown in Figure 5.13. We can see that the networks were well connected in the center while surrounded by some isolated nodes. Besides, not all cities were involved in the network. The most active institutions were located in a few cities: Shanghai, Hangzhou, Nanjing, and Hefei. The scale of the network did not show an obvious growth, but the number of publications and the connections in the hotspot cities had a large growth. It suggested that not many new research institutions appeared between 2017 and 2018, but the academic activities in the existing institutions were largely increasing.



(a) 2017 **(b)** 2018 **(c)** 2019

Figure 5.13: The flow maps showing the number of publications and the co-authorship connections among cities in YRD.

5.3.2 Assessment of Users' Attitude with Interviews

To understand the user experience of KnowDash, we investigated the users' attitude towards dashboard visualization. The experiment was conducted in May 2022 and performed online. We recruited seven cartographic experts to conduct online questionnaires. Among the participants, there were five females and two males. Their ages were between 25-59 years old. All of them were very familiar with dashboards and were known of scientific publication data. During the experiment, the users were firstly asked to find out the knowledge as shown in Section 5.3.1 and then answered the questions reflecting their attitudes.

In the experiment, the participants were asked to rate KnowData in four items as shown in Figure 5.14, and the five scales in rating are represented in five colors from purple to orange. From the results, we can see that most responses were positive. From the view of understanding the knowledge innovation environment, they thought KnowDash is understandable and trustworthy. From an entertaining point of view, the participants held diverse opinions. Except two participants mentioned the interface is trendy, most of them thought it is neutral, which might help users to focus on acquiring insights.

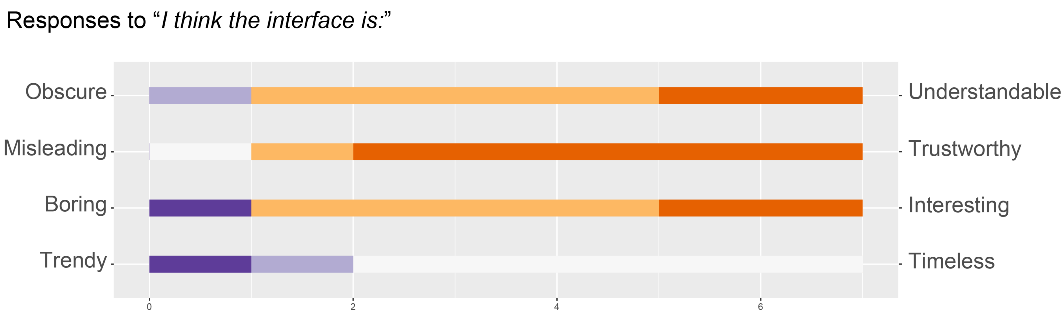


Figure 5.14: The results of the participants' attitudes towards KnowDash.

To understand the reasons for their attitudes, we asked the participants to write down the possible reasons for their feelings. The answers might indicate the factors that influence understandability, trust, entertainment, and visual attractiveness. Table 5.6 shows the extracted keywords from the participants' answers and the font size indicates the mentioned frequency. There are several keywords mentioned by different participants. To be specific, a *concise layout* is very important to understandability, and it contributes to trust. *Interaction* and *Visualization* play complex roles in the interface. *Simple interaction* has a positive influence on understandability and trust but a negative influence on entertainment and visual attractiveness. *Complex visualization* can reduce trust while increasing visual attractiveness. Likewise, *easy accessibility*, *clear data source*, and *clear data processing methods* contribute to understandability and trust. *Background knowledge* increases understandability, leads to insights, and contributes to entertainment.

5.4 Summary

The three dashboards illustrate the three different designs for different types of social environments and distinct analysis tasks. The dashboard design processes take the nature of the environments, the data patterns, and the users' needs into consideration.

Table 5.6: The actors of users' attitudes

	Understandability	Trust	Entertainment	Visual attractiveness
Increase	Easy accesibility Background knowledge Concise layout Simple interaction	Concise layout Simple interaction Indicate data source Clear data processing methods	Various visualizations Lead to insights Linked-view Colorful design	Color design Complex visualization
Decrease	Small font size Symbol overlap Lack explanation	Symbol overlap Complex visualization	Simple interaction	Simple functions

Although the usability was reflected in previous sections, we summarized some

From the three case studies, some design suggestions for map-based dashboards are summarized considering different factors:

Visual design. Unlike reading texts linearly from one side to another, dashboard readers might follow different reading directions. Therefore, the visualizations on dashboards should be arranged in a logic order, following the information abstraction level from general to detail. For example, the functional tools like selection and switch should be located in one zone, rather than scattered in different panels; the interactions in different panels should follow a unified design style like click or zoom-in.

Annotation. In the analytical context, it is very important to help users to firstly aware the background and context of the displayed information before their reasoning process. For designers, the titles and other texts on the dashboards should be carefully considered from a dashboard viewer's perspective. These text can help viewers to establish an ambient knowledge before they reason with the displayed information on dashboards. For example, the titles of the dashboard panels show the insights types (e.g., Popular Research Topics), instead of the name of the data table or visualization methods (e.g., Research Topics, Treemap).

Audience. The relevance between displayed information and the audience influence very much of the effectiveness and efficiency of acquiring insights from dashboard reading. Therefore, dashboards target to specific groups or diverse groups should represent the intrinsic and extrinsic data patterns, respectively.

> For instance, KnowDash targets to scholars and it shows the intrinsic patterns of the scholarly data, such as spatial distribution, hierarchical structures, and research topics; InDash expects diverse groups to read and it shows the extrinsic patterns, such as comparison between indexes and comparison between cities.

Although this study comprehensively demonstrated the benefits of map-based dashboards, there are several challenges with regard to the applications and technologies.

The evaluation of map-based dashboards remains challenging. Map-based dashboards go beyond data presentation and they mainly support visual reasoning to discover new knowledge. This means it is important to evaluate the performance and user experience of map-based dashboards in knowledge acquisition. However, the effectiveness and efficiency of knowledge acquisition via map-based dashboards are largely influenced by the background knowledge of individuals. A potential solution would be to observe the usage in practice and analyze it over the long term. In addition, a large sample size of the experiment participants is preferred to examine the design, but it is difficult to reach many participants for the in-lab experiment.

There are several technical challenges in designing map-based dashboards. The first challenge is to access data sources of good quality. Although the number of open data portals is increasing, the data for specific social environments is fragmentally stored in various agencies, following different standards. Data collection and cleaning take a much longer time than data modeling in the process. The second challenge is concerned with adaptation to different screens to show the map-based dashboards. To be more specific, the layout is important for users to read map-based dashboards, but various screens with different sizes and resolutions would sometimes distort the layouts. Designing responsive dashboards for various screens is a non-trivial effort. Finally, real-time data updating is challenging. Analytical map-based dashboards usually integrate complex modeling and computing tasks, especially big data modeling, and therefore real-time calculating and updating are technically hard for webpages and servers.

6

Conclusion and Outlook

6.1 Summary of thesis

The growing needs of the public in understanding the complex spatial patterns of social environments requires an effective and efficient visual language. This thesis is dedicated to proposing a design framework for map-based dashboard that can support learning the scope of a social environment and guiding users to acquire knowledge of social environment. The framework takes advantage of visual analytical methods and integrates human cognition to improve the visual design.

This study has contributions to both social environment understanding and visual analytics. More specifically, we summarize the following contributions.

- We enrich the visual language in understanding of social environments by identifying map-based dashboard to represent big data, proposing a design framework, and demonstrating three case studies. We highlighted the spatial information by the intensive use of maps. We carefully design the layout and visualizations to guide the users in analyzing the social environments in an "overview + detail" manner.

- Traditional visual analytics focus on using visual representations to explore the hidden patterns in data. We extend the visual analytical method by integrating users in the process. By considering users' experiences, the visual representations can be better adjusted to reach a boarder audience, including experts and novices.

Targeting at supporting the understanding of social environments and highlight the spatial information, the thesis design map-based dashboard to answer the proposed research questions:

Research question 1: Why is map-based dashboard selected to support the understanding of complex social environments?

Chapter 2 stated the growing need of using visual languages to support various stakeholders in social environment understanding. Previous studies introduced various visual vocabularies that can involve various users from various backgrounds, trigger human visual reasoning, and reduce information load. Map-based dashboards can synthesize multi-agency data, provide overview and details through interaction, and support multi-granularity and multi-factor analysis. With the intensive use of maps, map-based dashboards can highlight the spatial component in social environments and improve the spatial awareness of the citizens in social matters.

Research question 2: What are the important elements in map-based dashboards that facilitate users in acquiring knowledge from large amount of heterogeneous data?

Chapter 3 reviewed many studies on dashboards with various usages. The previous studies focus on proposing interactive map-based dashboard that can support users to explore given datasets and gain insights. To better involve users in the social environment analysis, we identified that visual layout in the interface design and user feedback are very important but have not been sufficiently addressed.

Research question 3: How can a map-based dashboard be designed that considers users' demands, perceptions, and data characteristics?

We proposed a framework in Chapter 4 with five components: 1) *Design goals* include information needs and users' experience. 2) *Users' cognitive tasks* contain such as topic understanding and data pattern discovering. 3) *Data* requires to collect suitable data, identify their geo-location, and transfer them into understandable variates. 4) *Interface* requires designing visualizations, interaction, layout, and description. 5) *Users' feedback* involves collecting user feedback on usability and user experience, that can be used to improve the dashboard design and examine the *design goals.*

Research question 4: How can the proposed design framework be implemented and evaluated for selected case studies?

The design framework of map-based dashboard was evaluated in Chapter 5 with three case studies regarding on the industrial innovation environment, socioeconomic environment, and knowledge innovation environment. These three case studies used open government data and big scholarly data, demonstrated three interfaces to support different cognitive tasks, and conducted user studies on usability, visual design, and user experience.

In summary, map-based dashboards are among the first attempts to bridge the gap between the social environments reflected by large volume geodata and the increasing stakeholders of social environments, although it is still very hard, if not impossible, to propose a generic design template to match different social environments.

6.2 Research Transferability

This research demonstrated the potential of map-based dashboards for the visual analysis of three complex social environments. Further complex social environments, such as public health monitoring, geometadata management, and hydrological environment analysis, could also be visually analyzed and communicated using analytical map-based dashboards.

Public health monitoring. Public health has many influence factors that are spatially correlated. As our experimental design in [ZZD21], map-based dashboards can be designed for monitoring public health situations to increase social awareness. It offers a compound interface for users to learn the environment from multiple factors and update the information frequently. For instance, the key performance indicators, spatial clusters, and temporal trends can be visualized on a map-based dashboard. The application would allow users to explore the data patterns and analyze the correlations between factors and also help frequent users to build a reading habit.

Geometadata management. Visual interface can support users in finding their needed datasets in geodata portals. We designed a preliminary map-based dashboard [ZGDM22] to provide an overview of the attributes of geometadata and help users to understand the available datasets, filter and search the needed datasets by location, time, or other attributes. For instance, the spatial coverage, time span, data quality, and data size can be visualized on a map-based dashboard. In addition, with the interactions embedded in the map-based dashboards, users may effectively find data in a large geodata portal.

Hydrological environment analysis. The hydrological environment involves multiple parameters of chemical portions at different places in the river systems.

Learning the changes of multiple measurements from the monitoring stations in different time periods can help scholars to analyze the water quality. Due to the complicated process and geographic-related knowledge, map-based dashboards can be designed to present the discrepancies between predicted values and true values intuitively. For instance, parameter dynamics of the measurement locations can be shown on a map-based dashboard. In addition, interactions can be designed to set focus on different parameters and filter different time periods.

6.3 Outlook

Except for the abovementioned achievements, we identified several directions for future research on map-based dashboard.

Design style and user perception. Map-based dashboard designers are challenged by balancing the complex information and interface legibility. The position, adjecency, and size of dashboard panels influence users' acquisition of insights. A quantitative research on the information perception and the graphic density, layer of depth, layout style, and information complexity should be carried.

Participatory interface. Map-based dashboards could allow users to contribute their own data, share their discovered knowledge, and collaboratively participate in decision-making. For instance, users are allowed to create and modify the data according to their knowledge. The dashboards could also allow users to contribute their own understanding of social environments.

Extended reality. Extended reality technology provides the potential to add more information to the surroundings and more human-computer interactions. Designing map-based dashboard with extended reality can provide additional information according to the users' location. It allows flexible display sizes and use body gestures to interact with the data-driven visualizations.

Interactive computing. Map-based dashboard could be further combined with more advanced computing methods, such as deep learning methods, to perform more computing intensive spatial analysis. By exposing more analysis methods parameters and intermediate results, the visual interface can help people to gain more reasoning results.

Human values. Except for effectiveness and efficiency, an in-depth consideration of human values, such as trustworthiness and engagement, should be considered and evaluated in the entire dashboard design process. In another word, the human values should be examined throughout the design goals, users' cognitive tasks, data, interface, and users' feedback process.

Tracking of the long-term influence. Due to the time limitation, the long-term influence of using dashboards is not investigated in this thesis. A number of questions, e.g., whether dashboards can cultivate visual thinking and what its influence is, whether the use of dashboards may lead to KPI-oriented thinking, remain open for future research.

Bibliography

[AA16] Ahmet Atıl Aşıcı and Sevil Acar. Does income growth relocate ecological footprint? *Ecological Indicators*, 61:707–714, 2016.

[AAG03] Natalia Andrienko, Gennady Andrienko, and Peter Gatalsky. Exploratory spatio-temporal visualization: an analytical review. *Journal of Visual Languages & Computing*, 14(6):503 – 541, 2003. Visual Data Mining.

[ACG14] Danielle Albers, Michael Correll, and Michael Gleicher. Task-driven evaluation of aggregation in time series visualization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, page 551–560, New York, NY, USA, 2014. Association for Computing Machinery.

[AES05] R. Amar, J. Eagan, and J. Stasko. Low-level components of analytic activity in information visualization. In *IEEE Symposium on Information Visualization, 2005. INFOVIS 2005.*, pages 111–117, 2005.

[APM20] Francesca Abastante, Stefano Pensa, and Elena Masala. *The Process of Sharing Information in a Sustainable Development Perspective: A Web Visual Tool*, pages 339–350. Springer International Publishing, Cham, 2020.

[APP11] Daniel Archambault, Helen Purchase, and Bruno Pinaud. Animation, small multiples, and the effect of mental map preservation in dynamic graphs. *IEEE Transactions on Visualization and Computer Graphics*, 17(4):539–552, 2011.

[Bal76] W. G. V. Balchin. Graphicacy. *The American Cartographer*, 3(1):33–38, 1976.

[BC01] Elizabeth Barnett and Michele Casper. A definition of "social environment". *American journal of public health*, 91(3):465–470, 2001.

[BGW01] Tom Brinck, Darren Gergle, and Scott D Wood. *Usability for the web: Designing web sites that work.* Morgan Kaufmann, San Francisco, CA, 2001.

[BHBPW13] Claudia R. Binder, Jochen Hinkel, Pieter W. G. Bots, and Claudia Pahl-Wostl. Comparison of frameworks for analyzing social-ecological systems. *Ecology and Society*, 18(4):1–26, 2013.

[BHW+21] Mona Bartling, Clemens R. Havas, Stefan Wegenkittl, Tumasch Reichenbacher, and Bernd Resch. Modeling patterns in map use contexts and mobile map design usability. *ISPRS International Journal of Geo-Information*, 10(8):527, 2021.

[BJ18] Edyta P. Bogucka and Mathias Jahnke. Feasibility of the space–time cube in temporal cultural landscape visualization. *ISPRS International Journal of Geo-Information*, 7(6):209, 2018.

[BJB09] Jeff Baker, Donald Jones, and Jim Burkman. Using visual representations of data to enhance sensemaking in data exploration tasks. *Journal of the Association for Information Systems*, 10(7):2, 2009.

[BSG+16] Marcus A Badgeley, Khader Shameer, Benjamin S Glicksberg, Max S Tomlinson, Matthew A Levin, Patrick J McCormick, Andrew Kasarskis, David L Reich, and Joel T Dudley. Ehdviz: clinical dashboard development using open-source technologies. *BMJ open*, 6(3):e010579, 2016.

[BWAL20] Michael Burch, Gunter Wallner, Sergiu Lazar Angelescu, and Peter Lakatos. Visual analysis of fifa world cup data. In *2020 24th International Conference Information Visualisation (IV)*, pages 114–119, 2020.

[CH20] Paweł Cybulski and Tymoteusz Horbiński. User experience in using graphical user interfaces of web maps. *ISPRS International Journal of Geo-Information*, 9(7):412, 2020.

[CLCY20] Shuai Chen, Sihang Li, Siming Chen, and Xiaoru Yuan. R-map: A map metaphor for visualizing information reposting process in social media. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):1204–1214, 2020.

[CLZ+17] Eun Kyoung Choe, Bongshin Lee, Haining Zhu, Nathalie Henry Riche, and Dominikus Baur. Understanding self-reflection: How people reflect on personal data through visual data exploration. In *Proceedings of the 11th EAI International Conference on Pervasive Computing Technolo-*

gies for Healthcare, PervasiveHealth '17, page 173–182, New York, NY, USA, 2017. Association for Computing Machinery.

[CMS99] Stuart K. Card, Jock D. Mackinlay, and Ben Shneiderman, editors. *Readings in Information Visualization: Using Vision to Think*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1999.

[CV17] Chia-Lin Chen and Roger Vickerman. Can transport infrastructure change regions' economic fortunes? some evidence from europe and china. *Regional Studies*, 51(1):144–160, 2017.

[CYGL14] Francisco Javier Carrillo, Tan Yigitcanlar, Blanca García, and Antti Lönnqvist. *Knowledge and the city: Concepts, applications and trends of knowledge-based urban development*. Routledge, New York, 2014.

[DE98] Alan Dix and Geoffrey Ellis. Starting simple: Adding value to static visualisation through simple interaction. In *Proceedings of the Working Conference on Advanced Visual Interfaces*, AVI '98, page 124–134, New York, NY, USA, 1998. Association for Computing Machinery.

[DK20] Ieva Dobraja and Menno-Jan Kraak. Principles of dashboard adaptability to get insights into origin-destination data. *Journal of Location Based Services*, 14(1):28–48, 2020.

[DLWV19] Soumitra Dutta, Bruno Lanvin, and Sacha Wunsch-Vincent. *Global Innovation Index 2019*. Cornell University, 2019.

[DR95] Mark Dodgson and Roy Rothwell. *The handbook of industrial innovation*. Edward Elgar Publishing, Massachusetts, USA, 1995.

[DS09] Kathleen Dolan and Kira Sanbonmatsu. Gender stereotypes and attitudes toward gender balance in government. *American Politics Research*, 37(3):409–428, 2009.

[DWC+20] Z. Deng, D. Weng, J. Chen, R. Liu, Z. Wang, J. Bao, Y. Zheng, and Y. Wu. Airvis: Visual analytics of air pollution propagation. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):800–810, Jan 2020.

[fDPC21] European Centre for Disease Prevention and Control. Covid-19 situation dashboard, https://qap.ecdc.europa.eu/public/extensions/covid-19/covid-19.html, 2021. (accessed on 14 January 2021).

[Few06] Stephen Few. *Information dashboard design: The effective visual communication of data*, volume 2. O'reilly Sebastopol, CA, 2006.

[FHWR20] Anna Fredriksson Häägg, Charlotte Weil, and Niklas Rönnberg. On the usefulness of map-based dashboards for decision making (in press). 2020.

[Fis11] Rob D. Fish. Environmental decision making and an ecosystems approach: Some challenges from the perspective of social science. *Progress in Physical Geography: Earth and Environment*, 35(5):671–680, 2011.

[FPV+13] Nivan Ferreira, Jorge Poco, Huy T Vo, Juliana Freire, and Cláudio T Silva. Visual exploration of big spatio-temporal urban data: A study of new york city taxi trips. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2149–2158, 2013.

[FS97] Christopher Freeman and Luc Soete. *The economics of industrial innovation*. Psychology Press, New York, 1997.

[FS12] Ali Fikirkoca and Ozcan Saritas. Foresight for science parks: the case of ankara university. *Technology Analysis & Strategic Management*, 24(10):1071–1085, 2012.

[FY04] J Friedmann and Li Yong. Planning for the global cities: The endogenous development model. In *Urban Planning Forum*, volume 4, 2004.

[GCML06] Diansheng Guo, Jin Chen, Alan M. MacEachren, and Ke Liao. A visualization system for space-time and multivariate patterns (vis-stamp). *IEEE Transactions on Visualization and Computer Graphics*, 12(6):1461–1474, nov 2006.

[GMH+06] Amy L. Griffin, Alan M. MacEachren, Frank Hardisty, Erik Steiner, and Bonan Li. A comparison of animated maps with static small-multiple maps for visually identifying space-time clusters. *Annals of the Association of American Geographers*, 96(4):740–753, 2006.

[GOR+07] Arnulf Grübler, Brian O'Neill, Keywan Riahi, Vadim Chirkov, Anne Goujon, Peter Kolp, Isolde Prommer, Sergei Scherbov, and Erika Slentoe. Regional, national, and spatially explicit scenarios of demographic and economic change based on sres. *Technological forecasting and social change*, 74(7):980–1029, 2007.

[GOR20] Izabela Golebiowska, Tomasz Opach, and Jan Ketil Rød. Breaking the eyes: How do users get started with a coordinated and multiple view geovisualization tool? *The Cartographic Journal*, 57(3):235–248, 2020.

[GRXY04] Stuart R Gaffin, Cynthia Rosenzweig, Xiaoshi Xing, and Greg Yetman. Downscaling and geo-spatial gridding of socio-economic projections from the ipcc special report on emissions scenarios (sres). *Global Environmental Change*, 14(2):105–123, 2004.

[Har04] Vesa Harmaakorpi. *Building a competitive regional innovation environment: The regional development platform method as a tool for regional innovation policy.* Helsinki University of Technology, Helsinki, Finland, 2004.

[HCP+20] Dongming Han, Wei Chen, Rusheng Pan, Yijing Liu, Jiehui Zhou, Ying Xu, Tianye Zhang, Changjie Fan, Jianrong Tao, et al. Graphfederator: Federated visual analysis for multi-party graphs. *arXiv preprint arXiv:2008.11989*, 2020.

[HES+20] Farook Hamzeh, Ali Ezzeddine, Lynn Shehab, Salam Khalife, Ghali El-Samad, and Samir Emdanat. Early warning dashboard for advanced construction planning metrics. In *Construction Research Congress 2020: Project Management and Controls, Materials, and Contracts*, pages 67–75. American Society of Civil Engineers Reston, VA, 2020.

[HJS+18] Lukáš Herman, Vojtěch Juřík, Zdeněk Stachoň, Daniel Vrbík, Jan Russnák, and Tomáš Řezník. Evaluation of user performance in interactive and static 3d maps. *ISPRS International Journal of Geo-Information*, 7(11):415, 2018.

[HJZ+14] Yaohuan Huang, Dong Jiang, Dafang Zhuang, Yunqiang Zhu, and Jingying Fu. An improved approach for modeling spatial distribution of water use profit—a case study in tuhai majia basin, china. *Ecological indicators*, 36:94–99, 2014.

[HNA+11] Kenneth Holmqvist, Marcus Nyström, Richard Andersson, Richard Dewhurst, Jarodzka Halszka, and Joost van de Weijer. *Eye Tracking : A Comprehensive Guide to Methods and Measures.* Oxford University Press, United Kingdom, 2011.

[HP12] Rex Hartson and Pardha Pyla. *The UX Book: Process and Guidelines for Ensuring a Quality User Experience*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1st edition, 2012.

[HSE+22] Ali Hamdi, Khaled Shaban, Abdelkarim Erradi, Amr Mohamed, Shakila Khan Rumi, and Flora D Salim. Spatiotemporal data mining: a survey on challenges and open problems. *Artificial Intelligence Review*, 55:1441–1488, 2022.

[HSK18] Iago Cotrim Henriques, Vinicius Amorim Sobreiro, and Herbert Kimura. Science and technology park: Future challenges. *Technology in Society*, 53:144–160, 2018.

[HVdB11] Noor Huijboom and Tijs Van den Broek. Open data: an international comparison of strategies. *European journal of ePractice*, 12(1):4–16, 2011.

[HZC+20] Zhaosong Huang, Ye Zhao, Wei Chen, Shengjie Gao, Kejie Yu, Weixia Xu, Mingjie Tang, Minfeng Zhu, and Mingliang Xu. A natural-language-based visual query approach of uncertain human trajectories. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):1256–1266, 2020.

[IHBD18] Indratmo, Lee Howorko, Joyce Maria Boedianto, and Ben Daniel. The efficacy of stacked bar charts in supporting single-attribute and overall-attribute comparisons. *Visual Informatics*, 2(3):155–165, 2018.

[JCS+19] Zhuochen Jin, Nan Cao, Yang Shi, Hanghang Tong, and Yingcai Wu. Ecolens: Visual analysis of urban region dynamics using traffic data. *ArXiv*, abs/1908.00181, 2019.

[JDLL19] Changfeng Jing, Mingyi Du, Songnian Li, and Siyuan Liu. Geospatial dashboards for monitoring smart city performance. *Sustainability*, 11(20):5648, 2019.

[JZHA22] S. Jamonnak, Y. Zhao, X. Huang, and M. Amiruzzaman. Geo-context aware study of vision-based autonomous driving models and spatial video data. *IEEE Transactions on Visualization & Computer Graphics*, 28(01):1019–1029, jan 2022.

[KAF+08] Daniel Keim, Gennady Andrienko, Jean-Daniel Fekete, Carsten Görg, Jörn Kohlhammer, and Guy Melançon. *Visual Analytics: Definition,* *Process, and Challenges*, pages 154–175. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.

[KE03] Michael J Kane and Randall W Engle. Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to stroop interference. *Journal of experimental psychology: General*, 132(1):47–70, 2003.

[Kei02] D.A. Keim. Information visualization and visual data mining. *IEEE Transactions on Visualization and Computer Graphics*, 8(1):1–8, 2002.

[KGF+22] Junfeng Kang, Xingyu Guo, Lei Fang, Xiangrong Wang, and Zhengqiu Fan. Integration of internet search data to predict tourism trends using spatial-temporal xgboost composite model. *International Journal of Geographical Information Science*, 36(2):236–252, 2022.

[KI13] Daniel F. Keefe and Tobias Isenberg. Reimagining the scientific visualization interaction paradigm. *Computer*, 46(5):51–57, 2013.

[KK22] Caglar Koylu and Alice Kasakoff. Measuring and mapping long-term changes in migration flows using population-scale family tree data. *Cartography and Geographic Information Science*, 49(2):154–170, 2022.

[KLM15] Rob Kitchin, Tracey P. Lauriault, and Gavin McArdle. Knowing and governing cities through urban indicators, city benchmarking and real-time dashboards. *Regional Studies, Regional Science*, 2(1):6–28, 2015.

[LAS+20] Roger A. Leite, Alessio Arleo, Johannes Sorger, Theresia Gschwandtner, and Silvia Miksch. Hermes: Guidance-enriched visual analytics for economic network exploration. *Visual Informatics*, 4(4):11–22, 2020.

[LB18] Thomas B Long and Vincent Blok. Integrating the management of socio-ethical factors into industry innovation: towards a concept of open innovation 2.0. *International food and agribusiness management review*, 21(4):463–486, 2018.

[LBI+12] Heidi Lam, Enrico Bertini, Petra Isenberg, Catherine Plaisant, and Sheelagh Carpendale. Empirical studies in information visualization: Seven scenarios. *IEEE Transactions on Visualization and Computer Graphics*, 18(9):1520–1536, 2012.

[LBLS15] Bireswar Laha, Doug A. Bowman, David H. Laidlaw, and John J. Socha. A classification of user tasks in visual analysis of volume data. In *2015 IEEE Scientific Visualization Conference (SciVis)*, pages 1–8, 2015.

[LCL21] Kim Lowell, Brian Calder, and Anthony Lyons. Measuring shallow-water bathymetric signal strength in lidar point attribute data using machine learning. *International Journal of Geographical Information Science*, 35(8):1592–1610, 2021.

[LCZ+19] J. Li, S. Chen, K. Zhang, G. Andrienko, and N. Andrienko. Cope: Interactive exploration of co-occurrence patterns in spatial time series. *IEEE Transactions on Visualization and Computer Graphics*, 25(8):2554–2567, Aug 2019.

[LD05] Tom Lowne and Carmel Diezmann. Fourth grade students' performance on graphical languages in mathematics. In H Chick and J Vincent, editors, *Proceedings of the 29th Conference of the International Group for the Psychology of Mathematics Education (PME 29)*, pages 265–272. University of Melbourne, Australia, 2005.

[Lee05] Chien-Chiang Lee. Energy consumption and gdp in developing countries: A cointegrated panel analysis. *Energy Economics*, 27(3):415 – 427, 2005.

[Li19] Mingyang Li. Big data based outdoor sports monitor and analysis system design of university. In *2019 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS)*, pages 245–248, 2019.

[LKGA18] Nancy Li, Julia Kramer, Pierce Gordon, and Alice Agogino. Co-author network analysis of human-centered design for development. *Design Science*, 4:e10, 2018.

[LLW+20] Quan Li, Huanbin Lin, Xiguang Wei, Yangkun Huang, Lixin Fan, Jian Du, Xiaojuan Ma, and Tianjian Chen. *MaraVis: Representation and Coordinated Intervention of Medical Encounters in Urban Marathon*, page 1–12. Association for Computing Machinery, New York, NY, USA, 2020.

[LVG+21] Dongyu Liu, Kalyan Veeramachaneni, Alexander Geiger, Victor OK Li, and Huamin Qu. Aqeyes: Visual analytics for anomaly detection and examination of air quality data. *arXiv preprint arXiv:2103.12910*, 2021.

[LWCZ20] Liang Liu, Song Wang, Ting Cai, and Weixin Zhao. Vabd: Visual analysis of spatio-temporal trajectory based on urban bus data. In *2020 IEEE 6th International Conference on Computer and Communications (ICCC)*, pages 2018–2022, 2020.

[LXR19] Dongyu Liu, Panpan Xu, and Liu Ren. Tpflow: Progressive partition and multidimensional pattern extraction for large-scale spatio-temporal data analysis. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):1–11, 2019.

[LYL+19] Xiaonan Luo, Yuan Yuan, Zhihao Li, Minfeng Zhu, Ying Xu, Liang Chang, Xiyan Sun, and Zi'ang Ding. Fbva: A flow-based visual analytics approach for citywide crowd mobility. *IEEE Transactions on Computational Social Systems*, 6(2):277–288, 2019.

[LZL+21] Liyan Liu, Hongxin Zhang, Jiaxin Liu, Shangching Liu, Wei Chen, and Jiaju Man. Visual exploration of urban functional zones based on augmented nonnegative tensor factorization. *Journal of Visualization*, 24(2):331–347, 2021.

[MGC+18] Martijn Millecamp, Francisco Gutiérrez, Sven Charleer, Katrien Verbert, and Tinne De Laet. A qualitative evaluation of a learning dashboard to support advisor-student dialogues. In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*, LAK '18, page 56–60, New York, NY, USA, 2018. Association for Computing Machinery.

[MHSW19] Eva Mayr, Nicole Hynek, Saminu Salisu, and Florian Windhager. Trust in Information Visualization. In Robert Kosara, Kai Lawonn, Lars Linsen, and Noeska Smit, editors, *EuroVis Workshop on Trustworthy Visualization (TrustVis)*. The Eurographics Association, 2019.

[MK97] Alan M. Maceachren and Menno-Jan Kraak. Exploratory cartographic visualization: Advancing the agenda. *Computers & Geosciences*, 23(4):335–343, 1997. Exploratory Cartograpic Visualisation.

[ML21] Hartmut Müller and Marije Louwsma. The role of spatio-temporal information to govern the covid-19 pandemic: A european perspective. *ISPRS International Journal of Geo-Information*, 10(3):166, 2021.

[MOC21] Ab Mosca, Alvitta Ottley, and Remco Chang. Does interaction improve bayesian reasoning with visualization? In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21, New York, NY, USA, 2021. Association for Computing Machinery.

[Moo15] Moodle. Moodle documentation, Apr 2015. (https://docs.moodle.org/28, accessed on 14 May 2022).

[Mor13] Kevin Morgan. Regional innovation strategies: the challenge of collaboration and governance. In *conference Local Economic Growth: Recession, Resilience and Recovery, Cambridge UK*, pages 11–12, 2013.

[MS14] Tomoya Mori and Tony E. Smith. A probabilistic modeling approach to the detection of industrial agglomerations. *Journal of Economic Geography*, 14(3):547–588, 2014.

[MSFM16] S. McKenna, D. Staheli, C. Fulcher, and M. Meyer. Bubblenet: A cyber security dashboard for visualizing patterns. *Computer Graphics Forum*, 35(3):281–290, 2016.

[MWR+20] Bill Moss, Crystal Watson, Lainie Rutkow, Brian Garibaldi, Beth Blauer, and Lauren Gardner. Covid-19 map, https://coronavirus.jhu.edu/map.html, 2020. (accessed on 16 May 2022).

[NB19] Kawa Nazemi and Dirk Burkhardt. Visual analytical dashboards for comparative analytical tasks – a case study on mobility and transportation. *Procedia Computer Science*, 149:138 – 150, 2019. ICTE in Transportation and Logistics 2018 (ICTE 2018).

[Nor06] C. North. Toward measuring visualization insight. *IEEE Computer Graphics and Applications*, 26(3):6–9, 2006.

[OGF14] Tomasz Opach, Izabela Gołębiowska, and Sara Irina Fabrikant. How do people view multi-component animated maps? *The Cartographic Journal*, 51(4):330–342, 2014.

[OLP+16] Katie O'Connell, Yeji Lee, Firaz Peer, Shawn M Staudaher, Alex Godwin, Mackenzie Madden, and Ellen Zegura. Making public safety data accessible in the westside atlanta data dashboard. *arXiv preprint arXiv:1609.09756*, 2016.

[PBG+14] Robert E. Patterson, Leslie M. Blaha, Georges G. Grinstein, Kristen K. Liggett, David E. Kaveney, Kathleen C. Sheldon, Paul R. Havig, and Jason A. Moore. A human cognition framework for information visualization. *Computers & Graphics*, 42:42–58, 2014.

[PFG08] Catherine Plaisant, Jean-Daniel Fekete, and Georges Grinstein. Promoting insight-based evaluation of visualizations: From contest to benchmark repository. *IEEE Transactions on Visualization and Computer Graphics*, 14(1):120–134, 2008.

[PHŘ+19] Stanislav Popelka, Lukáš Herman, Tomas Řezník, Michaela Pařilová, Karel Jedlička, Jiří Bouchal, Michal Kepka, and Karel Charvát. User evaluation of map-based visual analytic tools. *ISPRS International Journal of Geo-Information*, 8(8):363, 2019.

[PMSR18] Scott Pezanowski, Alan M MacEachren, Alexander Savelyev, and Anthony C Robinson. Senseplace3: a geovisual framework to analyze place–time–attribute information in social media. *Cartography and Geographic Information Science*, 45(5):420–437, 2018.

[QLR+20] Dezhan Qu, Xiaoli Lin, Ke Ren, Quanle Liu, and Huijie Zhang. Air-explorer: visual exploration of air quality data based on time-series querying. *Journal of Visualization*, 23(6):1129–1145, 2020.

[QZX16] Shengsheng Qian, Tianzhu Zhang, and Changsheng Xu. Multi-modal multi-view topic-opinion mining for social event analysis. In *Proceedings of the 24th ACM International Conference on Multimedia*, MM '16, page 2–11, New York, NY, USA, 2016. Association for Computing Machinery.

[RGH+17] Erna Ruijer, Stephan Grimmelikhuijsen, Michael Hogan, Sem Enzerink, Adegboyega Ojo, and Albert Meijer. Connecting societal issues, users and data. scenario-based design of open data platforms. *Government Information Quarterly*, 34(3):470–480, 2017.

[Rot13] Robert E Roth. Interactive maps: What we know and what we need to know. *Journal of Spatial Information Science*, 2013(6):59–115, 2013.

[RRM15] Robert E. Roth, Kevin S. Ross, and Alan M. MacEachren. User-centered design for interactive maps: A case study in crime analysis. *ISPRS International Journal of Geo-Information*, 4(1):262–301, 2015.

[SB11] Christian Schmidt and Peter Buxmann. Outcomes and success factors of enterprise it architecture management: empirical insight from the international financial services industry. *European Journal of Information Systems*, 20(2):168–185, 2011.

[SBv+21] Wonyoung So, Edyta P. Bogucka, Sanja Šćepanović, Sagar Joglekar, Ke Zhou, and Daniele Quercia. Humane visual ai: Telling the stories behind a medical condition. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):678–688, 2021.

[SBY+21] Oybek Shavkatbekovich Sabirov, Bakhriddin Tavasharovich Berdiyarov, Alisher Shavkatbekovich Yusupov, Akram Tolliboyevich Absalamov, and Adham Ilkhomjon Ugli Berdibekov. Improving ways to increase the attitude of the investment environment. *REVISTA GEINTEC-GESTAO INOVACAO E TECNOLOGIAS*, 11(2):1961–1975, 2021.

[SCB+19] A. Sarikaya, M. Correll, L. Bartram, M. Tory, and D. Fisher. What do we talk about when we talk about dashboards? *IEEE Transactions on Visualization and Computer Graphics*, 25(1):682–692, Jan 2019.

[SCS21] Dipto Sarkar, Colin A. Chapman, and Raja Sengupta. Mapping research networks supported by the national geographic society through spatial social networks. *International Journal of Geographical Information Science*, 35(12):2442–2462, 2021.

[SHH+21] D. Seebacher, J. Hausler, M. Hundt, M. Stein, H. Muller, U. Engelke, and D. A. Keim. Visual analysis of spatio-temporal event predictions: Investigating the spread dynamics of invasive species. *IEEE Transactions on Big Data*, 7(03):497–509, jul 2021.

[SHL+21] Lei Shi, Congcong Huang, Meijun Liu, Jia Yan, Tao Jiang, Zhihao Tan, Yifan Hu, Wei Chen, and Xiatian Zhang. Urbanmotion: Visual analysis of metropolitan-scale sparse trajectories. *IEEE Transactions on Visualization and Computer Graphics*, 27(10):3881–3899, 2021.

[Shn03] Ben Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. In BENJAMIN B. BEDERSON and BEN SHNEIDERMAN, editors, *The Craft of Information Visualization*, Interactive Technologies, pages 364–371. Morgan Kaufmann, San Francisco, 2003.

[SMP+19] Daniel Seebacher, Matthias Miller, Tom Polk, Johannes Fuchs, and Daniel A. Keim. Visual analytics of volunteered geographic information: Detection and investigation of urban heat islands. *IEEE Computer Graphics and Applications*, 39(5):83–95, 2019.

[SND05] P. Saraiya, C. North, and K. Duca. An insight-based methodology for evaluating bioinformatics visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 11(4):443–456, 2005.

[SRH+20] Agus SALIM, Andi RUSTAM, Haeruddin HAERUDDIN, Asriati ASRIATI, and Aditya Halim Perdana Kusuma Putra. Economic strategy: Correlation between macro and microeconomics on income inequality in indonesia. *The Journal of Asian Finance, Economics, and Business*, 7(8):681–693, 2020.

[SSS+14] Dominik Sacha, Andreas Stoffel, Florian Stoffel, Bum Chul Kwon, Geoffrey Ellis, and Daniel A. Keim. Knowledge generation model for visual analytics. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1604–1613, 2014.

[SWH14] Arvind Satyanarayan, Kanit Wongsuphasawat, and Jeffrey Heer. Declarative interaction design for data visualization. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*, UIST '14, page 669–678, New York, NY, USA, 2014. Association for Computing Machinery.

[Tid01] Joe Tidd. Innovation management in context: environment, organization and performance. *International Journal of Management Reviews*, 3(3):169–183, 2001.

[Tor70] Hägerstrand Torsten. What about people in regional science? *Regional Science Association*, 24(1):6–21, 1970.

[Tuf16] Edward Tufte. *The Visual Display of Quantitative Information (1983):*, pages 219–230. De Gruyter (A), 2016.

[UTS10] David Urbano, Nuria Toledano, and Domingo Ribeiro Soriano. Analyzing social entrepreneurship from an institutional perspective: Evidence from spain. *Journal of Social Entrepreneurship*, 1(1):54–69, 2010.

[vdEvW13] Stef van den Elzen and Jarke J. van Wijk. Small multiples, large singles: A new approach for visual data exploration. *Computer Graphics Forum*, 32(3pt2):191–200, 2013.

[vW13] Jarke J. van Wijk. Evaluation: A challenge for visual analytics. *Computer*, 46(7):56–60, 2013.

[War19] Colin Ware. *Information visualization: perception for design.* Morgan Kaufmann, 2019.

[WCR13] Geoffrey F. Woodman, Nancy B. Carlisle, and Robert M. G. Reinhart. Where do we store the memory representations that guide attention? *Journal of Vision*, 13(3):1–1, 02 2013.

[WHY20] Shijin Wang, Guihong Hua, and Lizhu Yang. Coordinated development of economic growth and ecological efficiency in jiangsu, china. *Environmental Science and Pollution Research*, 27(29):36664–36676, 2020.

[Wil12] Leland Wilkinson. *The Grammar of Graphics*, pages 274–292. Springer Berlin Heidelberg, Berlin, Heidelberg, 2012.

[WL90] S. Wehrend and C. Lewis. A problem-oriented classification of visualization techniques. In *Proceedings of the First IEEE Conference on Visualization: Visualization ‘90*, pages 139–143, 1990.

[WO07] Richard Webster and Margaret A Oliver. *Geostatistics for environmental scientists.* John Wiley & Sons, 2007.

[WSPC20] Patrick Würstle, Thunyathep Santhanavanich, Rushikesh Padsala, and Volker Coors. The conception of an urban energy dashboard using 3d city models. In *Proceedings of the Eleventh ACM International Conference on Future Energy Systems*, e-Energy ’20, page 523–527, New York, NY, USA, 2020. Association for Computing Machinery.

[WZD+21] Di Weng, Chengbo Zheng, Zikun Deng, Mingze Ma, Jie Bao, Yu Zheng, Mingliang Xu, and Yingcai Wu. Towards better bus networks: A visual analytics approach. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):817–827, 2021.

[YEB18] M. A. Yalçın, N. Elmqvist, and B. B. Bederson. Keshif: Rapid and expressive tabular data exploration for novices. *IEEE Transactions on Visualization and Computer Graphics*, 24(8):2339–2352, Aug 2018.

[YKSJ07] Ji Soo Yi, Youn ah Kang, John Stasko, and J.A. Jacko. Toward a deeper understanding of the role of interaction in information visualization. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1224–1231, 2007.

[YKSJ08] Ji Soo Yi, Youn-ah Kang, John T. Stasko, and Julie A. Jacko. Understanding and characterizing insights: How do people gain insights using information visualization? In *Proceedings of the 2008 Workshop on BEyond Time and Errors: Novel EvaLuation Methods for Information Visualization*, BELIV '08, New York, NY, USA, 2008. Association for Computing Machinery.

[YS20] B. Yu and C. T. Silva. Flowsense: A natural language interface for visual data exploration within a dataflow system. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):1–11, Jan 2020.

[ZDL+22] Chenyu Zuo, Linfang Ding, Xiaoyu Liu, Hui Zhang, and Liqiu Meng. Map-based dashboard design with open government data for learning and analysis of industrial innovation environment. volume 0, pages 1–17. Taylor & Francis, 2022.

[ZDM20] Chenyu Zuo, Linfang Ding, and Liqiu Meng. A feasibility study of map-based dashboard for spatiotemporal knowledge acquisition and analysis. *ISPRS International Journal of Geo-Information*, 9(11):636, 2020.

[ZDYM22] Chenyu Zuo, Linfang Ding, Zhuo Yang, and Liqiu Meng. Multiscale geovisual analysis of knowledge innovation patterns using big scholarly data. *Annals of GIS*, 28(2):197–212, 2022.

[ZF98] Michelle X. Zhou and Steven K. Feiner. Visual task characterization for automated visual discourse synthesis. In *Proceedings of the SIGCHI*

Conference on Human Factors in Computing Systems, CHI '98, page 392–399, USA, 1998. ACM Press/Addison-Wesley Publishing Co.

[ZGDM22] Chenyu Zuo, Mengyao Gao, Linfang Ding, and Liqiu Meng. Space-time cube for visual queries over metadata of heterogeneous geodata. *KN-Journal of Cartography and Geographic Information*, 72(1):29–39, 2022.

[ZLJ+19] Ruoxin Zhu, Diao Lin, Michael Jendryke, Chenyu Zuo, Linfang Ding, and Liqiu Meng. Geo-tagged social media data-based analytical approach for perceiving impacts of social events. *ISPRS International Journal of Geo-Information*, 8(1):15, 2019.

[ZMTK17] Vladimir Zdraveski, Kostadin Mishev, Dimitar Trajanov, and Ljupco Kocarev. Iso-standardized smart city platform architecture and dashboard. *IEEE Pervasive Computing*, 16(2):35–43, 2017.

[ZQP+15] Wangsheng Zhang, Guande Qi, Gang Pan, Hua Lu, Shijian Li, and Zhaohui Wu. City-scale social event detection and evaluation with taxi traces. *ACM Trans. Intell. Syst. Technol.*, 6(3), may 2015.

[ZSB+12] Leishi Zhang, Andreas Stoffel, Michael Behrisch, Sebastian Mittelstadt, Tobias Schreck, René Pompl, Stefan Weber, Holger Last, and Daniel Keim. Visual analytics for the big data era — a comparative review of state-of-the-art commercial systems. In *2012 IEEE Conference on Visual Analytics Science and Technology (VAST)*, pages 173–182, 2012.

[ZSC+21] Ting Zhang, Shi Shen, Changxiu Cheng, Kai Su, and Xiangxue Zhang. A topic model based framework for identifying the distribution of demand for relief supplies using social media data. *International Journal of Geographical Information Science*, 35(11):2216–2237, 2021.

[ZWL+21] Ruoxin Zhu, Yujing Wang, Diao Lin, Michael Jendryke, Mingxia Xie, Jianzhong Guo, and Liqiu Meng. Exploring the rich-club characteristic in internal migration: Evidence from chinese chunyun migration. *Cities*, 114:103198, 2021.

[ZZD21] Hui Zhang, Chenyu Zuo, and Linfang Ding. Visualizing spatiotemporal epidemic clusters on a map-based dashboard: A case study of early covid-19 cases in singapore. *Proceedings of the ICA*, 4:117, 2021.

Related Publications

Journals

- C. Zuo; L. Ding; X. Liu; H. Zhang; L. Meng: Map-based dashboard design with open government data for learning and analysis of industrial innovation environment. International Journal of Cartography 9 (1), 2023, 97-113, doi:10.1080/23729333.2022.2049106

- C. Zuo; L. Ding; Z. Yang; L. Meng: Multiscale geovisual analysis of knowledge innovation patterns using big scholarly data. Annals of GIS 28 (2), 2022, 197-212, doi:10.1080/19475683.2022.2027012

- C. Zuo; M. Gao; L. Ding; L. Meng: Space-Time Cube for Visual Queries over Metadata of Heterogeneous Geodata. KN - Journal of Cartography and Geographic Information 72 (1), 2022, 29-39, doi:10.1007/s42489-022-00096-5

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

Conferences

- C. Zuo; J. Grübel: Using Map-based Dashboard to Improve Geo-knowledge Communication: a use case on Digital Twins, Abstr. Int. Cartogr. Assoc., 6, 295, 2023, doi.org/10.5194/ica-abs-6-295-2023

- A. Keler; S. Ludwig; C. Zuo: Using Interactive Space-Time Cube Visualisation for Pattern Mining in Bicycle Trajectories and Traffic-Related Parameters, 31st Geographical Information Science Research UK (GISRUK 2023), Glasgow, United Kingdom Geneva: CERN, April 19-21, 2023, doi:10.5281/zenodo.7839558

- C. Zuo; L. Ding; L. Meng: The reading strategies of a map-based dashboard – acquisition and assessment. Abstracts of the ICA 3, 2021, 1-2, doi:10.5194/ica-abs-3-331-2021

- C. Zuo; Y. Zhuo; L. Ding; L. Meng: Visual Analysis of the Knowledge Innovation Environment Using Big Scholarly Data. 17th International Conference on Computational Urban Planning and Urban Management, 2021

- H. Zhang; C. Zuo; L. Ding: Visualizing Spatiotemporal Epidemic Clusters on a Map-based Dashboard: A case study of early COVID-19 cases in Singapore. Proceedings of the ICA 4, 2021, 1-5, doi:10.5194/ica-proc-4-117-2021

- C. Zuo; B. Liu; L. Ding; E. Bogucka; L. Meng: Usability Test of Map-based Interactive Dashboards Using Eye Movement Data. The 15th International Conference on Location Based Services, Vienna University of Technology, 2019, doi:10.34726/lbs2019

- C. Zuo; L. Ding; E. Bogucka; L. Meng: Map-based Dashboards versus Storytelling Maps. The 15th International Conference on Location Based Services, Vienna University of Technology, 2019, doi:doi:10.34726/lbs2019

- C. Zuo; L. Ding; L. Meng: Visual Analytics for Regional Economic Environment Factors Based on a Dashboard Design. Proceedings of the ICA 2, 2019, 1-8, doi:10.5194/ica-proc-2-158-2019