Overview

Information visualization

Chaomei Chen*

This overview introduces the key structure of the field of information visualization, a number of influential exemplars in the field, and challenging as well as promising directions of future developments. The focus is on explaining some of the most fundamental concepts, prominent approaches, and commonly held criteria. The overview also aims to point out theoretical and practical challenges that the community as a whole has been addressing. *WIREs Comp Stat* 2010 2 387–403 DOI: 10.1002/wics.89

Keywords: information visualization; visual analytics; interactive dynamic graphics/interactive; visual thinking; visual design; human perception

The term *information visualization* refers to computer generated interactive graphical representations of information. In this article, it also refers to the process of producing information visualization representations. The field of information visualization refers to the scientific community of researchers and practitioners who are contributing or have contributed to the field of study. This overview article aims to provide a brief introduction to information visualization. The overview is primarily intended for the audience who are not familiar with the field. The introduction will focus on core motivations and ambitions of information visualization, landmark and exemplar work in the field, emerging trends and promising directions of further growth. Specialized fields such as geovisualization, software visualization, and visual analytics are discussed briefly. For comprehensive coverage of these specialized fields, interested readers are referred to materials listed in the resources.

OVERVIEW

Information visualization is concerned with the design, development, and application of computer generated interactive graphical representations of information. This often implies that information visualization primarily deals with abstract, nonspatial data. Transforming such nonspatial data to intuitive and meaningful graphical representations is therefore of fundamental importance to the field. The transformation is also a creative process in which designers assign new meanings into graphical patterns. Like art, information visualization aims to communicate complex ideas to its audience and inspire its users for new connections. Like science, information visualization must present information and associated patterns rigorously, accurately, and faithfully.[1]

A common question is the relationship between information visualization and scientific visualization. A simple answer is that they are unique in terms of their corresponding research communities. They do overlap, but largely differ. Here are some questions that might further clarify the scope of information visualization.

First, is the original data numerical? Graphical depictions of quantitative information are often seen in the fields of data visualization, statistical graphics, and cartography. For example, is a plot of daily temperatures of a city for the last 2 years qualified as information visualization? The answer to this question may depend on another question: how easy or straightforward is it for someone to produce the plot? As Michael Friendly and Daniel J. Denis put it,[a] *unless you know its history, everything might seem novel*. By the same token, what is complex and novel today may become trivial in the future. A key point to differentiate information visualization from data visualization and scientific visualization is down to the presence or absence of data in quantitative forms and how easy one can transform them to quantitative forms. This is why researchers emphasize the ability to represent nonvisual data in information visualization.[2]

Second, if the data is not spatial or quantitative in nature, what does it take to transform it to something that is spatial and visual? This step involves visual design and the development of computer algorithms. It is this step that clearly distinguishes information visualization from its nearest neighbors such as quantitative data visualization. In a more formal terms, this step can be found in an earlier taxonomy

*Correspondence to: chaomei.chen@drexel.edu

College of Information Science and Technology, Drexel University, Philadelphia, PA 19104-2875, USA

DOI: 10.1002/wics.89

of information visualization,[3] which models the process of information visualization in terms of data transformation, visualization transformation, and visual mapping transformation. Data transformation turns raw data into mathematical forms. Visualization transformation establishes a visual–spatial model of the data. Visual mapping transformation determines the appearance of the visual–spatial model to the user. On the other hand, if the data is quantitative in nature, researchers and designers are in a better position to capitalize on this valuable given connection.

The connection between scientific and artistic aspects of information visualization is discussed in terms of *functional* information visualization and *aesthetic* information visualization.[4] The primary role of functional information visualization is to communicate a message to the user, whereas the goal of aesthetic information visualization is to present a subjective impression of a data set by eliciting a visceral or emotive response from the user.

Geometry, Structure, and Semantics

The absence of a predefined geometry associated with a given type of data is often seen as an important indicator that differentiates information visualization from scientific visualization. The distinction can be profound as well as subtle. In information visualization, researchers typically find themselves in need of defining the semantics of visual displays as an integral part of the design. In contrast, scientific visualization researchers may need to choose a reference system for the same reason, although the degree of complexity may vary from a well-defined theory of a scientific phenomenon to an initial exploration of a newly emerged phenomenon. Similarly to define the meaning of geometry, researchers need to characterize the meaning of structural patterns, for example, what a tightly coupled component of a network of coauthors represents.

The ultimate design question is whether salient features of geometric or structural patterns convey the intended message to the viewer. The challenge for information visualization is that the attachment of meaningful geometric or visual encoding is much more arbitrary than its counterparts in scientific visualization. Designers and viewers tend to share much more of a common ground when viewing a visualization of a storm or a tsunami than they do when viewing the evolution of a multidimensional and highly abstract topic. Science of science as an application domain of information visualization, for example, requires a higher order of abstraction than scientific visualization. This is the primary criterion as one draws the boundary of information visualization.

Creating the meaning of geometric and structural patterns is both a core challenge and a source of excitement in information visualization. From the design perspective, the best information visualization is the one that can effectively convey the meaning to its users. From a user's point of view, it should be either intuitive in that most people will have no problem to understand the intended meaning, or easy to learn so that one can easily figure out through interacting with the visualization. In reality, this is a long process before it becomes clear what works and with what constraints. More studies should follow the rigorous approaches demonstrated by earlier fundamental studies of elementary perceptual tasks.[5–7]

THE HOLY GRAIL

The holy grail of information visualization is for users to gain *insights*. In general, the notion of insight is broadly defined, including unexpected discoveries, a deepened understanding, a new way of thinking, eureka-like experiences, and other intellectual breakthroughs. In early years of information visualization, it is believed that the ability to view the entirety of a data set at a glance is important to discover interesting and otherwise hidden connections and other patterns. More recently, it is realized, with the rise of visual analytics, that the actionability of information visualization is essential and it emphasizes the process of searching for insights instead of the notion of insights *per se*.

Information visualization researchers have addressed issues concerning how to measure insights in the context of evaluative studies.[8,9] Unlike research in the data mining community on interestingness, which aims to develop metrics and algorithms to identify the interestingness in given data, relatively few concrete metrics have been derived in the field of information visualization.[10] Relevant concepts include saliency, uniqueness, recency, and burstness.

A tough and ultimate question for information visualization advocates to the general public and other scientific and technological fields is what information visualization can achieve that other ways can't or at least without paying a much higher price. In light of MySpace, FaceBook, Twitter, and more, social networking in cyberspace offers a new perspective to the issue. IBM's ManyEyes, for example, allows users to make their own visualizations and share their comments on visualizations made by others. The social dimension may offer a new route to increase the chance of getting insights, although research in this direction has merely began.[11] Social navigation, the ability to blaze trails in a cyber information space, as what Vennevar Bush talked about in As We May Think,[12] is likely to play more important roles in information visualization.

THE COMMUNITY

The world's first symposium on information visualization was the IEEE Symposium on Information Visualization (InfoVis) that took place in 1995. The IEEE InfoVis Symposium is widely regarded as the flagship conference of the community. Since 2008, InfoVis, the recently emerged Visual Analytics Science and Technology (VAST), and IEEE Visualization became closely coordinated events now known as VisWeek.

In Europe, the London-based International Conference on Information Visualization (IV) started in 1997 and it is still running. Another annual European symposium series was formerly called VisSym and renamed to EuroVis from 2006.

In Asia and Pacific countries, the regional information visualization symposium started in 2001 in Australia in the name of InVis.au. It became the later APVIS—Asia Pacific Symposium on Information Visualization later on until it took the current PacificVis in 2008.

The first book of a collection of important publications on information visualization was published in 1999. The first dedicated international journal, *Information Visualization*, was launched in 2002. More recently, InfoVis Symposium papers are published simultaneously as a conference paper and a journal paper.

Many data-driven or data-laden fields organize contests or competitions so that researchers and developers can work on the same data set with their own tools. Notable ones include American Statistical Association Data Expo, which is the source of the cars data set many have studied, KDD cup, VAST contest, and, of course, the InfoVis contests. Information visualization contests started in 2003. Topics of InfoVis contests in the past include the literature of the field itself, movie data, and sensor data. Lessons learned from 2003, 2004, and 2005 InfoVis contests are summarized by the organizers.[13] Robert Kosara, who organized more recent 2008 InfoVis contests, has shared some interesting reflections[b] on 'The Sad State of the InfoVis Contest'.

MILESTONES

A good place to see information visualization in a broader context is the list of milestones in the history of thematic cartography, statistical graphics, and data visualization maintained by Michael Friendly and Daniel J. Denis.[14,15] Their list starts with a Konya town map made 6200 BC, arguably the oldest known map of all. The list also includes a few milestones that the information visualization community is more familiar with, including Playfair's parallel time-series bar chart of prices of wheat, wages, and monarchs in 1700s and John Snow's dot map of cholera deaths in 1800s. Milestones after 1975 are intentionally sparse, but includes some of the milestones that have profound influences on the emergence of the field of information visualization, namely fisheye views in 1981,[16] aesthetics and information integrity in concepts such as data-link ratio,[17] Jacques Bertin's *Semiology of Graphics*,[5] parallel coordinate plots[18] for high-dimensional data analysis,[19] a computational extension of Bertin's semiology of graphics,[7] and Treemaps, a space-filling visualization of hierarchies in nested rectangular areas.[20]

In addition to the milestones on Friendly and Denis' list, the following ones are also remarkable to the evolution of the information visualization field. This is obviously an incomplete list.

- Force-directed graph drawing algorithms[21–23]; constraint-based graph layout[24];
- SemNet[25];
- Cone trees[26];
- Visualizing large networks[27,28];
- SageBook[29];
- Self-Organizing Maps[30];
- Landscape and galaxy views[2];
- Hyperbolic trees[31]; large networks in 3D hyperbolic trees[32];
- First books[33–36];
- Visual analytics research agenda[37].

Here we highlight some of the widely known works that have profoundly influenced the development of information visualization.

Telling a Story

Charles Minard's depiction of the Russian campaign of Napoleon's army is widely recognized as one of the most compelling exemplars of storytelling (Figure 1). It shows the size of the diminishing army on its way to Moscow and the even more staggering losses on

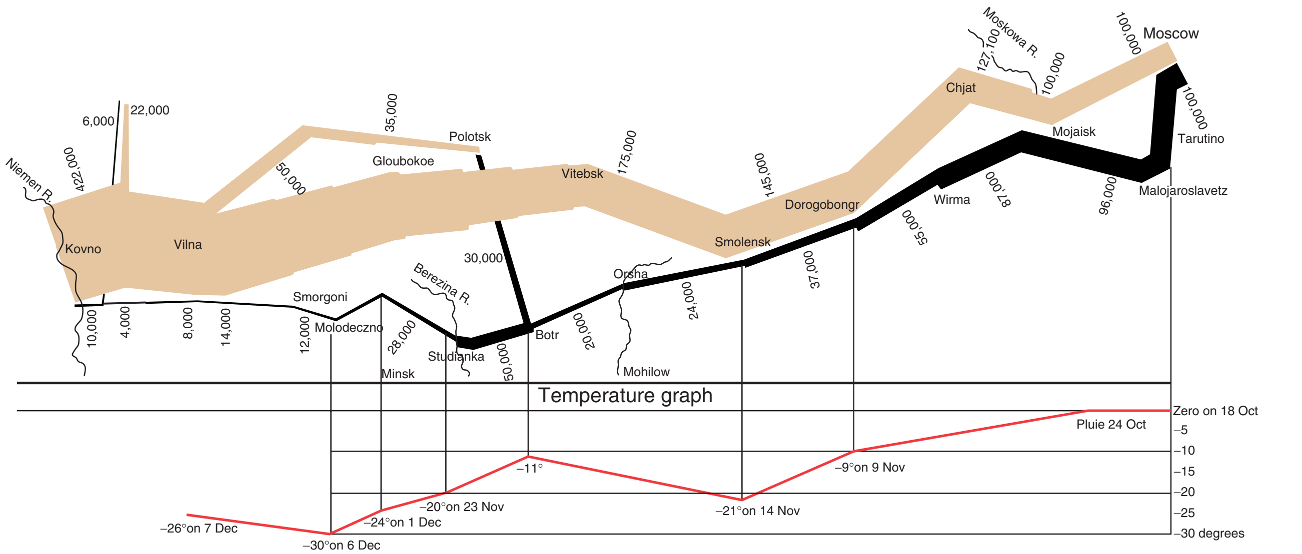

FIGURE 1 | Losses suffered by the Grande Armée during the Russian Campaign.[c]

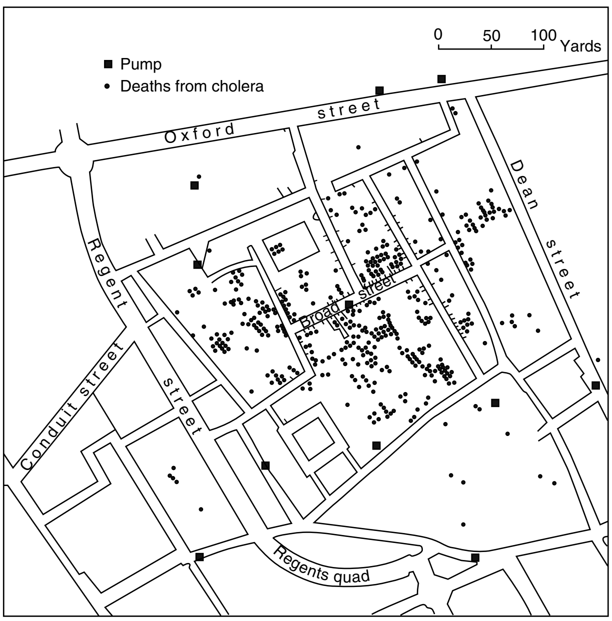

FIGURE 2 | John Snow's dot map of cholera deaths.[d]

its retreat. Many modern information visualization designers replicate the original design.

Making a Discovery

One of the most intriguing examples that have shed light on the power of visual thinking is due to John Snow's investigation of the deaths of a cholera epidemic in 1855s in London (Figure 2). Edward Tufte gives a vivid historical account in his Visual Explanations, pp. 27–37. Each death was shown as a dot on a simple street map of London. It was the concentration of dots that revealed the hidden connection between the deaths and the contaminated well.

Seeing the Big Picture

SmartMoney visualizes the ups and downs of stock prices in the financial market. It organizes the stock market in several high-level categories such as financial, energy, health care, transport, and technology. Each category is shown as a rectangular area, which is in turn divided further into smaller rectangular areas. Each of the smaller areas represents the status of a particular company, for example, Dell Computer +6.84% as shown in the map (Figure 3). This organization is known as a treemap, originally developed at the University of Maryland. SmartMoney is one of the few successful stories of how information visualization moves beyond research labs.

Seeing with Many Eyes

ManyEyes is one of a kind in information visualization. It is a new 'social kind of data analysis' in the own words of its designers at IBM's Visual Communication Laboratory. ManyEyes enables many people to have a taste of what is like to create your own information visualization that they would otherwise have no such chance at all. The public-oriented design significantly simplifies the entire process of information visualization. Furthermore, ManyEyes is indeed a community-building environment in which one can view visualizations made by other users, make

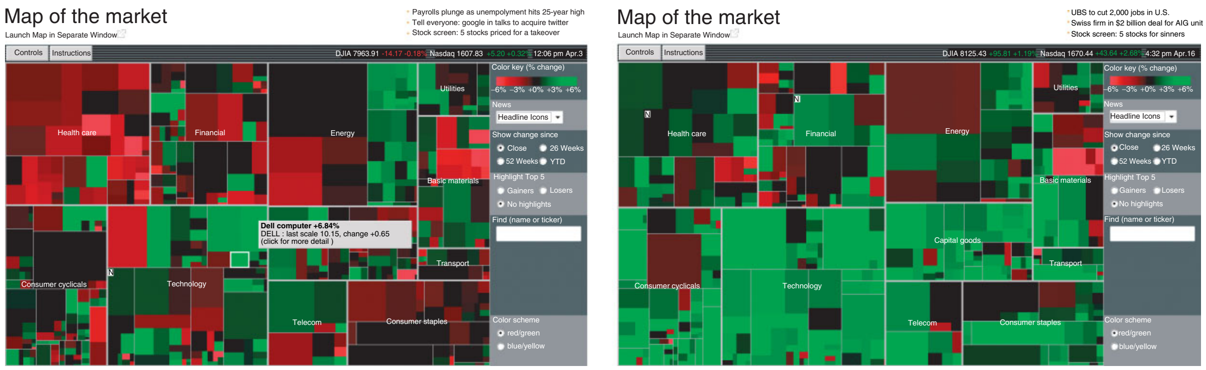

FIGURE 3 | SmartMoney on April 3, 2009 (left) and April 16, 2009 (right). http://www.smartmoney.com/map-of-the-market/.

comments, and make your own visualizations. These reasons alone would be enough to earn ManyEyes a unique position in the development of information visualization. ManyEyes and Wikipedia share some interesting characteristics—both tap in social construction and both demonstrate emergent properties of a self-organizing underlying system. Figure 4 shows the opening webpage of ManyEyes.

Aesthetics and Functionality

The art and science are both integral parts of information visualization. Researchers and artists have attempted to derive criteria that may tell us when information visualization is art, when it is not, and when it is in between.[4] Gaviria distinguishes functional and aesthetic information visualization. The design of traffic lights is essentially functional, whereas a landscape painting is certainly more on the aesthetic side. It is often said that the purpose of information visualization is insight, not just pretty pictures. On the other hand, discussions of aesthetics are inevitable in the design of evaluative studies and searching for theoretical foundations of information visualization. Is TextArc, shown in Figure 5, art or science? What about Figures 6 and 7? Figure 5

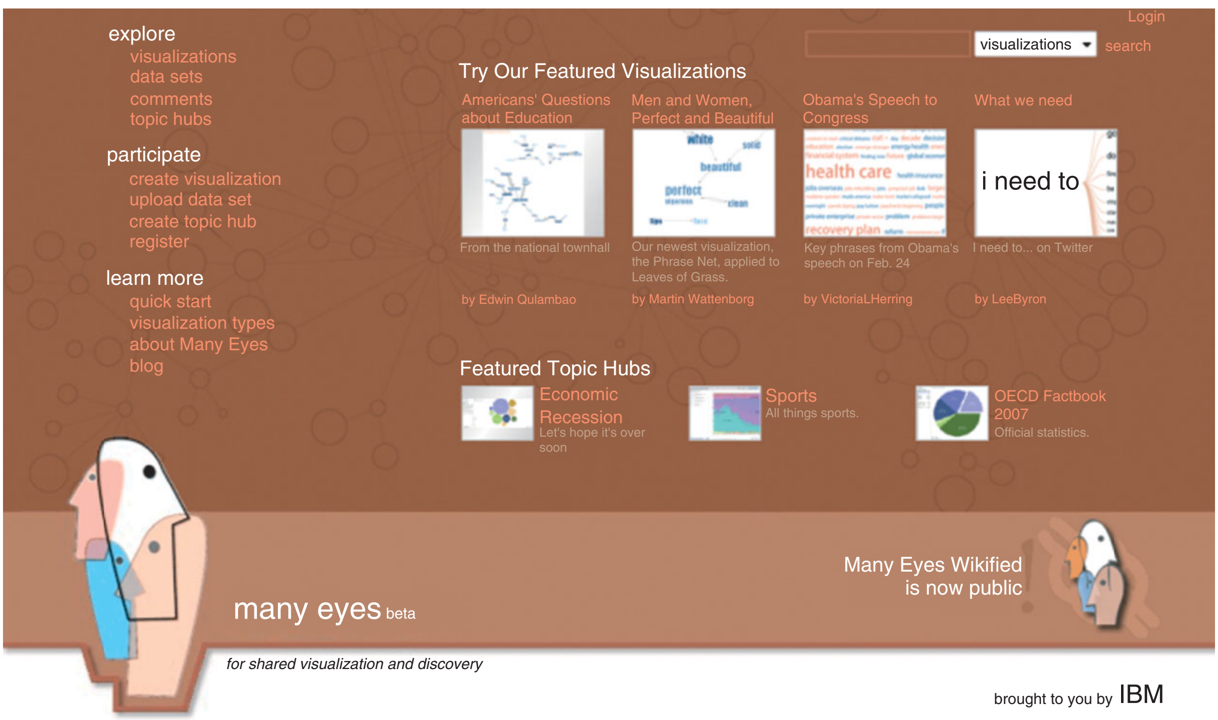

FIGURE 4 | ManyEyes. http://manyeyes.alphaworks.ibm.com/manyeyes/.

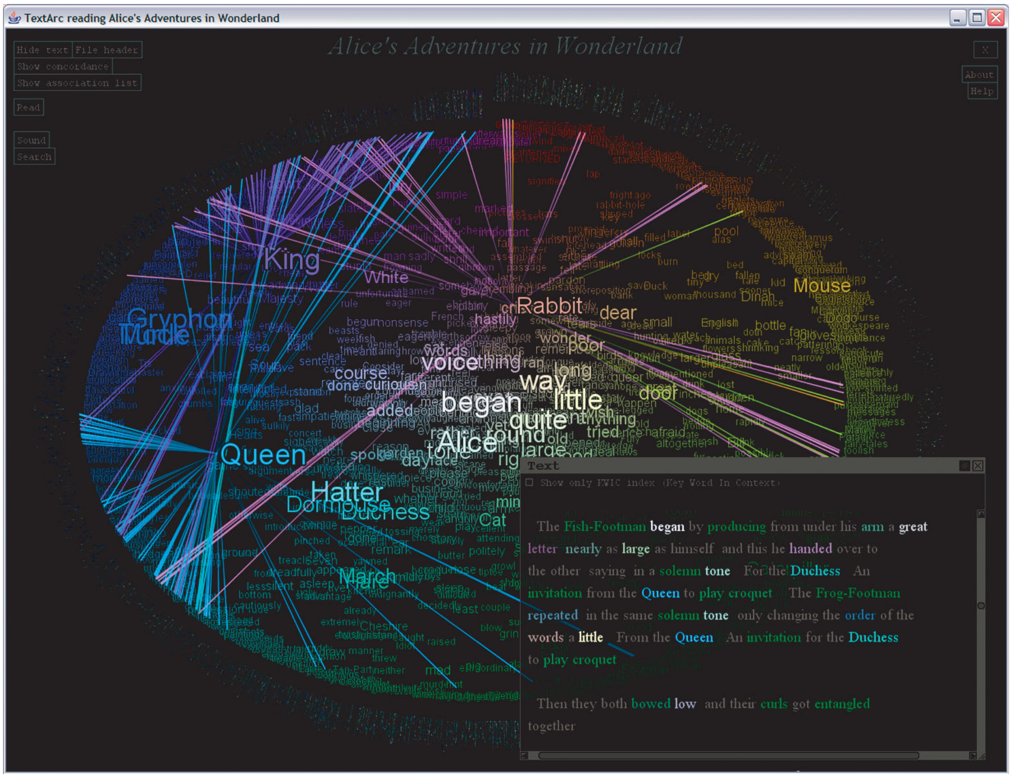

FIGURE 5 | Alice's adventures in Wonderland in TextArc. http://www.textarc.org/.

shows *Alice's Adventures in Wonderland* in TextArc. Although TextArc was designed as a tool, it has been in the Museum of Modern Art in New York. According to Bradford Paley, its creator, engineers see TextArc as a feat of engineering, artists as art, analysts as a tool, and designers as design. As it seems, such artefacts become a part of the world in their own right and each of us may relate to our own unique perspectives.

Figure 6 shows the three-dimensional virtual worlds of several scientific domains' citation landscapes.[38] Each sphere depicts a scientific publication. The height of the bar is proportional to the number of citations the corresponding publication

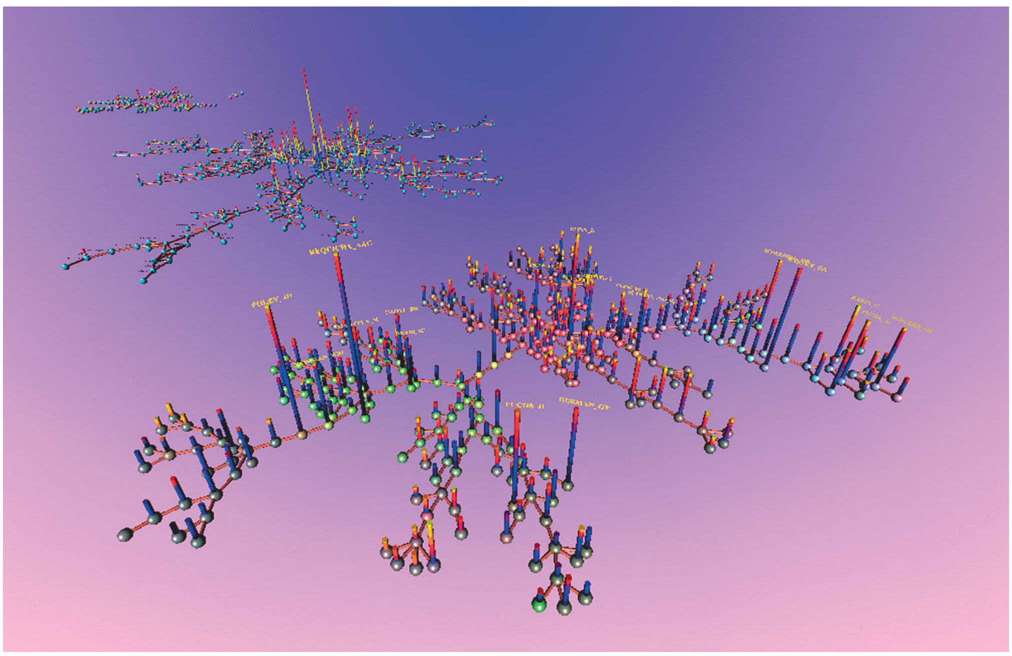

FIGURE 6 | Knowledge domain landscapes.[38]

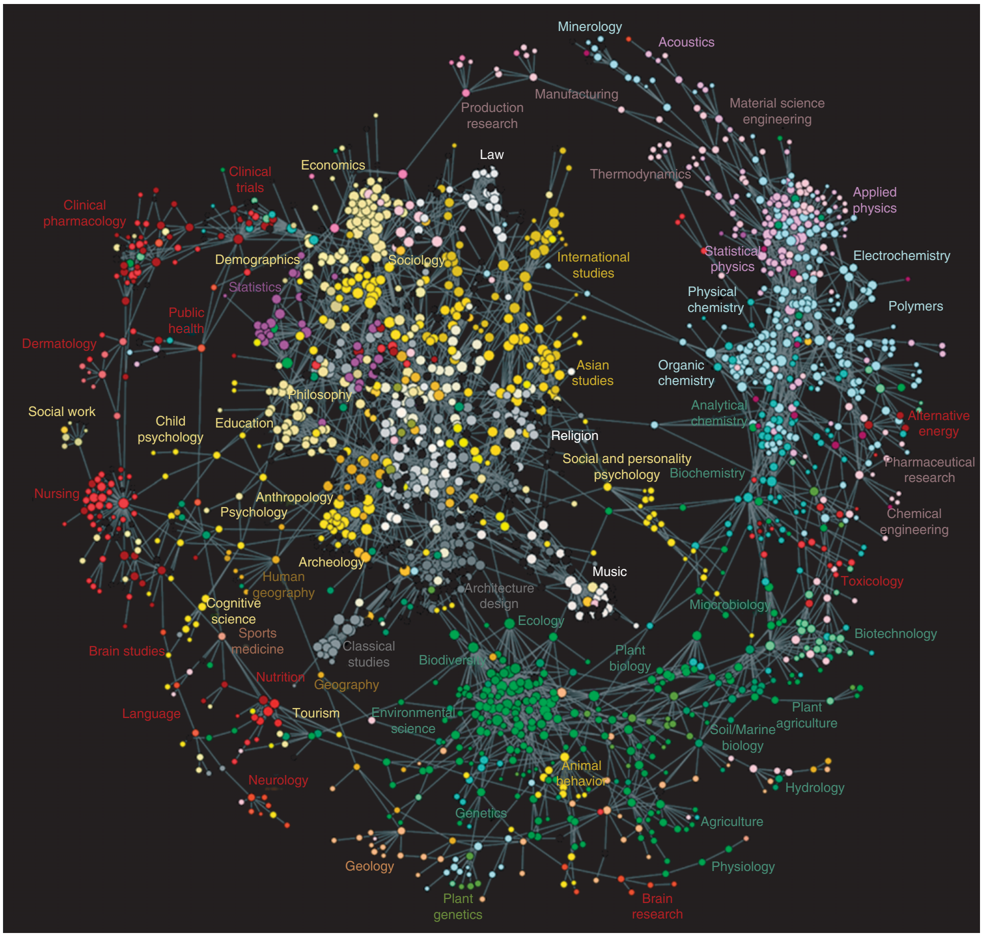

FIGURE 7 | A clickstream map of scientific journals.[39]

has received. The colors on the bar represent the time when citations were made. The most recent citations appear toward the top of the bar, whereas earlier citations are in the lower part of the bar.

Figure 7 shows a recent example of a network visualization of scientific journals derived from clickstreams. Each node in the graph is a journal. Two journals are connected if they are next to one another in clickstream data. The visualized graph is a simplified representation. A journal has to appear in at least 170 observations to qualify entering into the view. There are 2307 such journals. Then only the five strongest outbound relationships for each journal would be retained and subsequently, a single edge between any pair of journals is shown in the visualization. Finally, the largest connected component was selected for a fully interconnected visualization. The radius of each node is scaled to the natural logarithm of the journal's degree centrality in order to avoid a cluttered map.

Visual Analytics

A newly emerged field of visual analytics has its roots in information visualization as well as other contributing fields. Visual analytics aims to support analytical reasoning and decision-making activities through the use of information visualization, statistical analysis, data mining, and other techniques. One of the most notable documents in the relevant literature is the research agenda of visual analytics.[37]

Visual analytics is celebrating its fifth anniversary in 2009. It has its own symposium alongside the IEEE InfoVis symposium. It also runs annual contest. The community has substantial overlaps with the InfoVis community. Figure 8 is a screenshot of GeoTime, a visual analytic system for investigating events in both spatial and temporal dimensions.[40]

FOUNDATIONS

Gestalt Principles

Metaphors of an information space imply a definition of metric that measures the distance in the abstract space. The notion of an abstract space taps into what is known as *Gestalt Psychology*, which gives principles of our tendency to see patterns of individual items. The central idea of Gestalt principles is that, as far

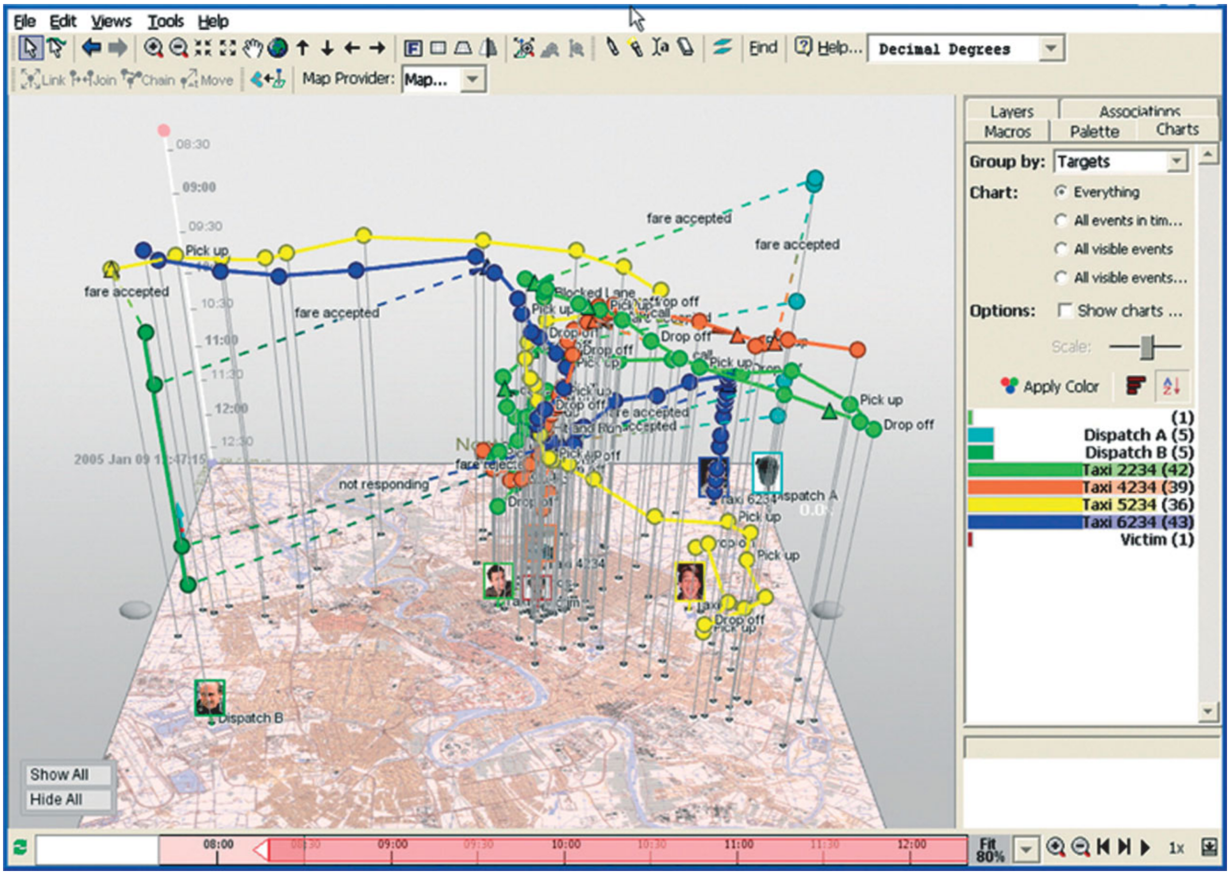

FIGURE 8 | A screenshot of GeoTime.

as our perception and cognition are concerned, the whole is more than the sum of parts. Most commonly known Gestalt principles are proximity, similarity, continuity, closure, figure and ground, and symmetry. A comprehensive explanation of these principles and a rich set of examples can be found in *How Maps Work*[41] in the context of cartography, that is, the design of geographically based thematic maps.

The proximity principle says that we tend to see groupings of individual items in a visual arrangement based on the proximity between these items. Items that are relatively close to one another tend to give us a sense of similarity. In other words, we see individual items in groups of some underlying similarity. This principle has been adapted by the information visualization community from the early stage. Algorithms that can arrange information items in this fashion tap into the proximity principle. Some interesting examples include Bead[42] and Information Mural.[43]

The similarity principle from Gestalt psychology says that visual attributes such as the shape, color, and texture are cues for us to group items, for example, all the circles in one group and all the triangles in another. The proximity and similarity principles can be used simultaneously to reinforce each other.

The Mantra of Visual Information Seeking

The most widely known visual information seeking mantra is given by Ben Shneiderman, University of Maryland: Overview first, zoom and filter, then details-on-demand.[44] This mantra insightfully summarizes the essential elements of interacting with graphically presented information.

Designers of visual overviews commonly capitalize on metaphors that can give users a sense of intuitiveness and familiarity. Naturally, metaphors of an information space are particularly popular, especially in 1990s, ranging from two-dimensional maps, three-dimensional landscape views and contours, to star fields and galaxies of information. An important function of an overview is to depict interrelationships among units of information. Units of textual information include words, sentences, documents, and collections of documents such as websites. Units of visual information include scenes, episodes, and libraries of videos.

Information space metaphors naturally invite navigational operations such as zoom, pan, tilt, and rotate. One of the earlier claims and design goals of information visualization is that good information visualization should present information to users intuitively. Many filtering operations have been adapted to enable users interact with dynamic information visualization, including brushing, linking, dynamic queries, and coordinated views.

In 1996, Shneiderman offered a taxonomy for visual information seeking.[44] The taxonomy divides general visual information seeking into seven data types and seven tasks. This taxonomy is one of the earliest and most influential contributions to the information visualization field.

Seven Data Types

- one-dimensional data;
- two-dimensional data;
- three-dimensional data;
- temporal data;
- multidimensional data;
- tree data;
- network data.

Seven Tasks

- overview;
- zoom;
- filter;
- details-on-demand;
- relate;
- history;
- extract.

The data type by task taxonomy has influenced a generation of information visualization researchers. Other notable efforts include the data state reference model.[3]

The Pursuit of Insights

Reflections on insight-centric evaluation are motivated by the increasing concern of how to establish the effectiveness of interacting with information visualization interfaces. On the one hand, it is almost a community-wide consensus that insight is the ultimate goal of information visualization. On the other hand, the definition of insight in the information visualization literature *per se* has been vague and ambiguous. The nature of insight has been extensively studied in the context of scientific discovery in cognitive science, psychology, and history of science. Few connections have been established so far between the study of insight in other disciplines and the field of information visualization. An intriguing introduction to some of the recent understanding of the brain activities that lead to insights can be found in a *New Yorker* article *The Eureka Hunt*.[45] *The Nature of Insight* is a comprehensive collection of studies of insight.[46]

In a recently developed explanatory and computational theory of scientific discovery, the nature of insight is characterized by a brokerage mechanism and a burst function of recognition.[47] The brokerage mechanism echoes what is described in the Eureka Hunt in that one arrives at insights by linking previously unconnected thoughts. The theory is computational and it is possible to formulate the search for insights as a problem of searching for the potential linkage between even the most unthinkable relations. Initial studies of transformative discoveries such as Nobel Prize winning discoveries are particularly promising. This approach is particularly relevant to visual analytics and insight-based evaluative studies because they can characterize insightful patterns in terms of structural and temporal properties.

Within the information visualization community, notable efforts on characterizing and measuring insights include exploratory approaches as opposed to benchmark-based experimental studies,[9] lessons learned from the first 3 years of InfoVis contests,[48] and more recent reflection in the context of visual analytics.[8] An interesting framework of evaluating interactive visualizations is proposed recently in Ref 49. The framework is built on top of a generic conceptual model in human–computer interaction, namely Don Norman's Seven Stages of Action.[50] According to the Seven Stages of Action, two stages of interacting with computer interfaces are particularly problematic: execution and evaluation. The gulf of execution and the gulf of evaluation are used to refer to these problematic stages. The gulf of execution, for example, should be narrowed so that users can accomplish their tasks smoothly and seamlessly. The gulf of evaluation should be narrowed so that users can judge their progress accurately.

Much of the discussions in information visualization on insights primarily address practical and methodological issues concerning how evaluative studies should be designed to capture the effectiveness of an information visualization design in terms of insights. The types of insights that are relevant to information visualization and evaluative studies have theoretical and practical implications. We found two meta-analysis studies of information visualization.[51,52] Given the growing calls for theoretical foundations in the field, this is expected to be a significant topic of research.

Theoretical Frameworks

The general consensus, as reported by a recent workshop and a few other public presentations, was that information visualization currently lacks adequate theoretical foundations.[53] As a result, many approaches are *ad hoc* in nature. A week-long seminar took place at Dagstuhl, Germany in mid-2007, for example, addressed four potential directions for developing new theories. The lack of theories becomes particularly prominent in information visualization courses and when designing empirical and evaluative studies.

The search for theoretical foundations increasingly introduces and adopts theories and conceptual frameworks from other fields and disciplines. For example, distributed cognition in human–computer interaction is seen as a potential candidate for a theoretical framework for information visualization.[54] Norman's Seven Stages of Action, also in human–computer interaction, provides a new insight into interacting with information visualizations, specifically on the gulf of execution and the gulf of evaluation.[49]

Many information visualizations lack a quantitative measure that could indicate the overall quality, uncertainly, novelty, and other evaluative metrics. The focus on gulfs of execution and evaluation, for example, has the potential to make progress in this direction.

TECHNICAL ADVANCES

Some of the recent developments in information visualization are worth noting. At the technical level, scalability issues remain to be a long-lasting challenge.[55] Some of the algorithms developed for clustering large-scale data sets in machine learning are particularly appealing, such as Refs 56,57 and one can expect these algorithms will soon find their ways to information visualizations.

Edge Bundling

Edge bundling is an emerging strategy to solve a common problem in visualizing a densely connected graph due to cluttered images caused by overlapping edges. Avoiding edge crossings has been long recognized as one of the constraints that could improve the clarity of resultant visualizations. Recently, an interesting strategy has emerged—that is the use of edge bundling techniques in a variety of graph visualizations to increase the clarity of visualized patterns. Bundling reduces visual clutter. Visualizations with bundled edges make it easier for viewers to see underlying patterns than non-bundled versions,[58] for example, as shown in Figure 9. Edge bundling is a generic technique in nature because it can be applied virtually to all node-and-link diagrams regardless the underlying layout algorithms. In this sense, it is similar to other generic display techniques such as fisheye views. A geometry-based edge bundling example appears recently, showing prominent patterns of migration in the USA[59] (Figure 10).

Constraint-Based Graph Drawing

Another trend originated from the graph drawing community is constraint-based graph drawing. Tim Dwyer et al. are the leading researchers in this research area.[60] Many graph visualization applications can benefit from the new development in this direction because of the generic and valuable role in establishing visual hierarchies and grouping (Figure 11).

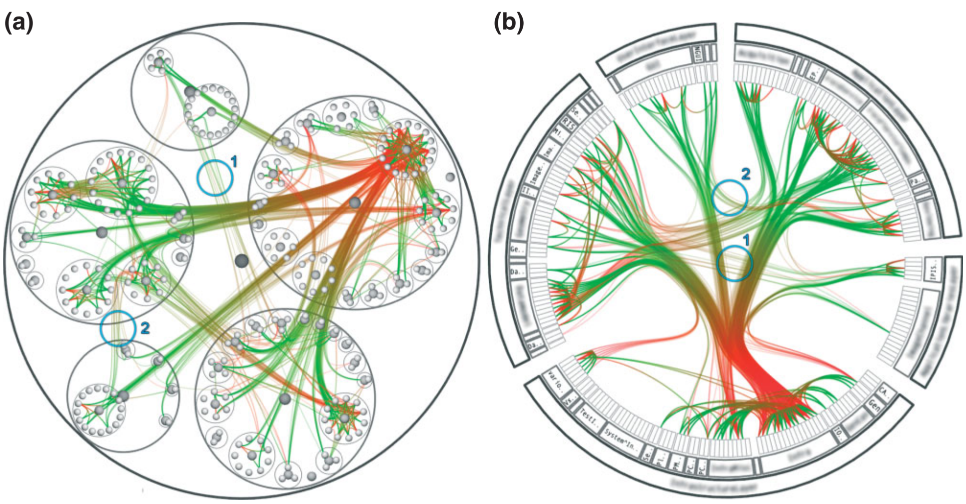

FIGURE 9 | Bundled edges in graph visualization.[58]

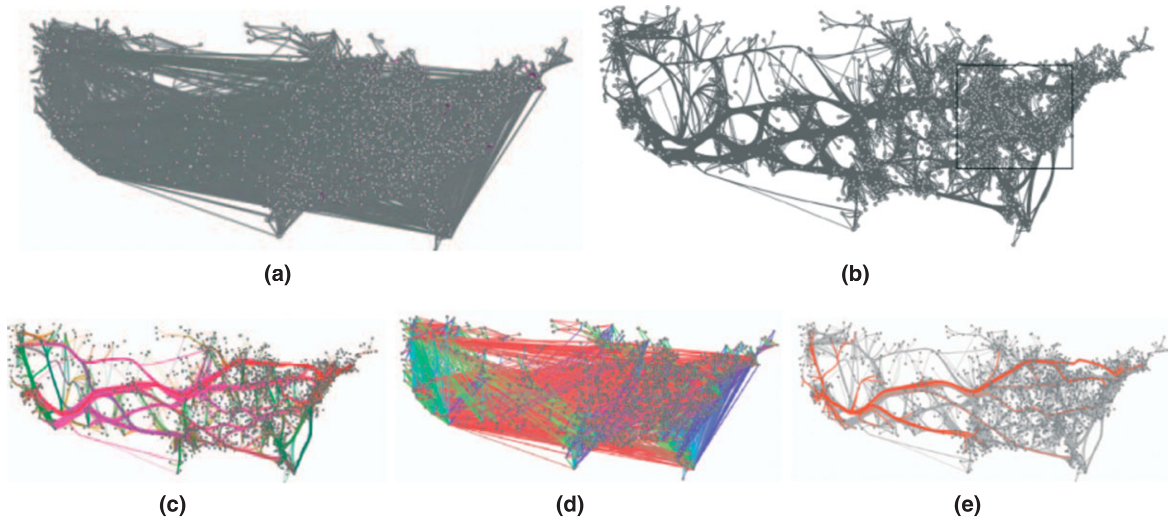

FIGURE 10 | Geometry-based edge bundling.[59]

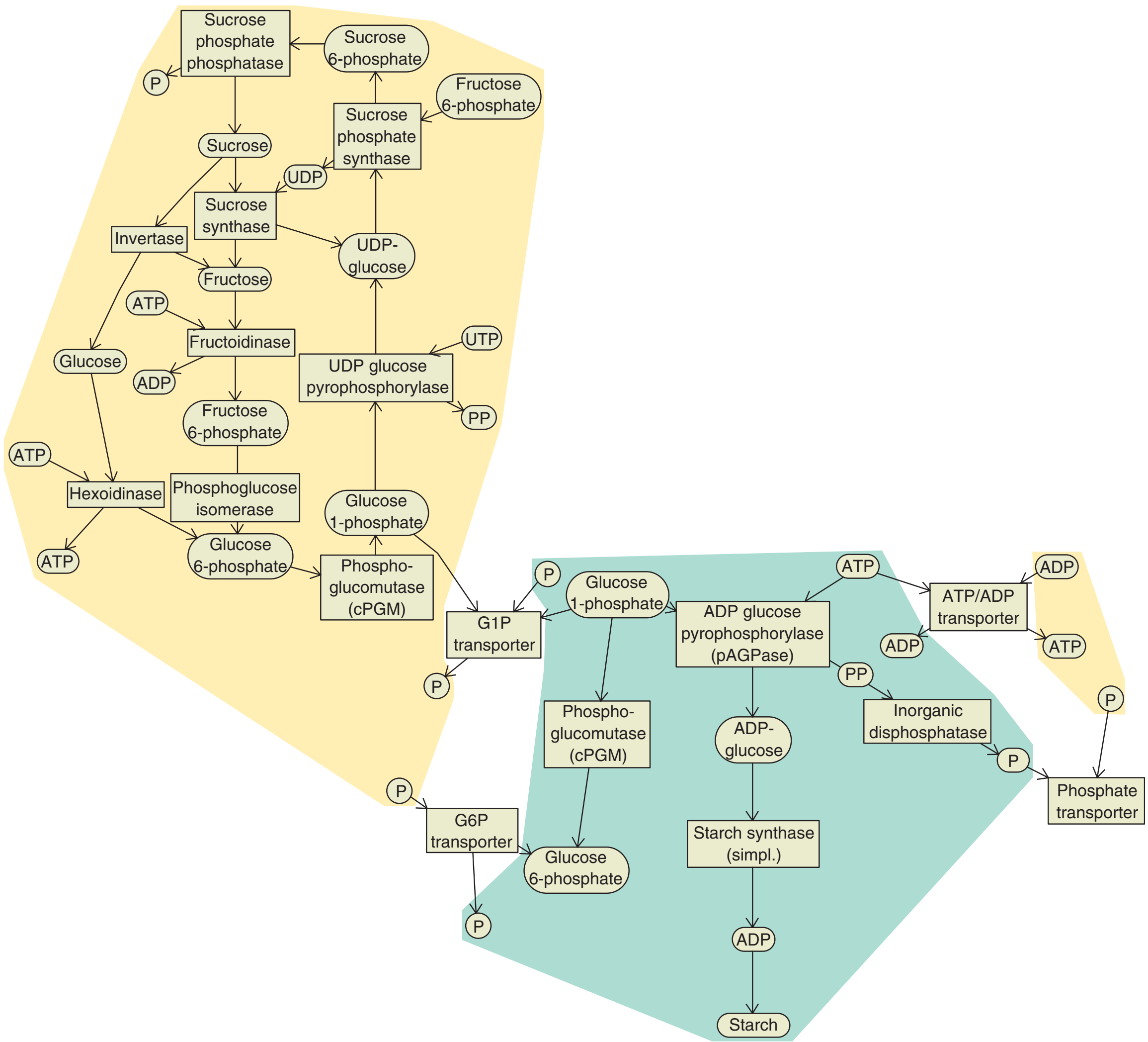

FIGURE 11 | Layout with constraints.[60]

Logarithmic Transformations

One of the problems identified earlier on by the information visualization community is the tension between showing more local details while maintaining users' contextual awareness. The problem is known as the focus + context problem. Many widely known techniques in information visualization were indeed developed originally to deal with such problems, notably including fisheye views[16] and hyperbolic views.[31] Along a similar vein, Figure 12 shows a logarithmic view.[61] Logarithmic transformations are commonly used by astronomers when they need to deal with multiple vast scales. The major advantage of a logarithmic view is its computational scalability. Like fisheye views, a logarithmic view also enlarges some areas of display at the expenses of other areas. Figure 13 illustrates the use of logarithmic transformations of the sky. Astronomical objects distributed across a wide span of multiple scales are depicted in the same single sky map. See video of mapping the universe with Sloan Digital Sky Survey.[e]

Other enabling and supporting techniques include fast point-feature labeling algorithms,[62] and fast network scaling algorithms that improve semantically desirable but computationally expensive algorithms such as Pathfinder network scaling.[63–66]

EMERGING TRENDS AND FUTURE DIRECTIONS

Mixed-Initiative Interaction

Integrating perceptual guidelines from human vision with an AI-based mixed-initiative search strategy is a promising but challenging direction for information visualization.[67] Mixed-initiative interaction is motivated by the observation that even experienced designers cannot be expected to know everything about how to construct effective visualizations due to the diverse range of situated requirements. Furthermore, designers often repeatedly utilize the same basic design strategy. Consequently, the resulting visualization may not be the best possible design option. It is often more effective to be able to explore the same set of data from different perspectives through different visualization designs. Therefore, the goal of mixed-initiative interaction is to augment designers with an easy access to the existing body of knowledge of proven and effective visualization design options in a given scenario. The underlying principle is very similar to the concept of design pattern and design language in the field of human–computer interaction.

FIGURE 12 | Logarithmic view centered at the Capitol in Washington.[61] Points northwest of the capitol are mapped to a vertical line in the middle of the image. Points southeast are mapped to the very left and the very right.

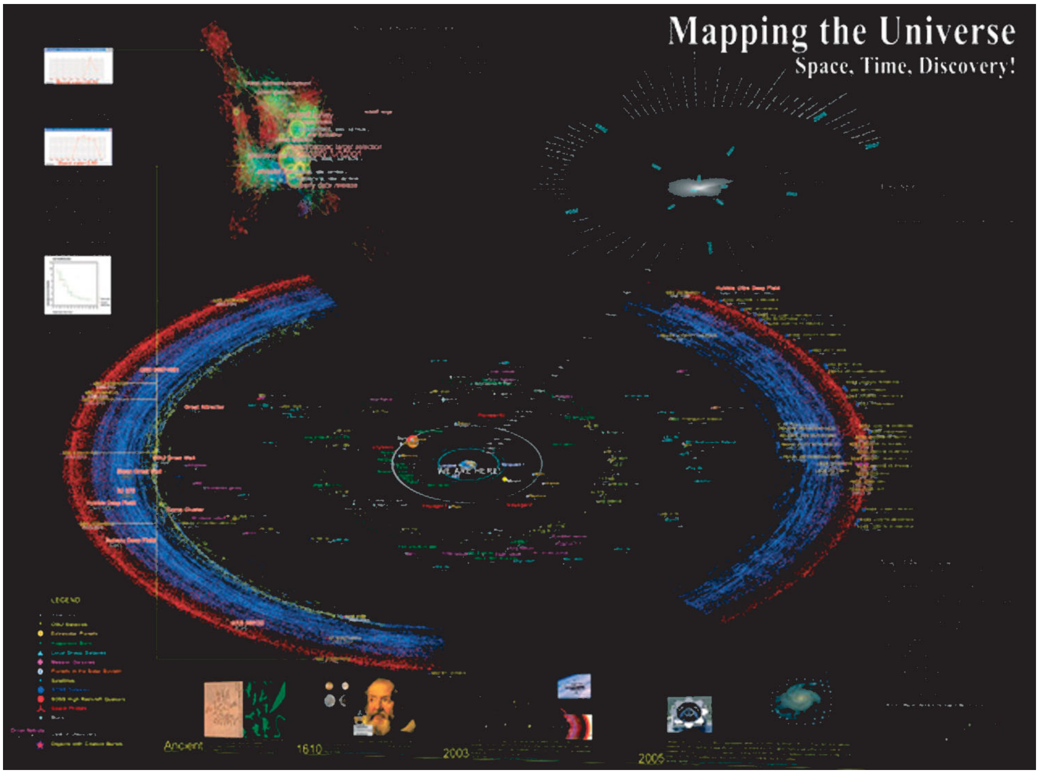

FIGURE 13 | A circular logarithmically transformed map of the universe.[f] The circular structure in this visualization is a 2D projection of the Universe based on the right ascension and the distance between the Earth and an astronomical object, such as stars, galaxies, or quasars.

Collaborative and Social Visualization

Tom Erickson is a pioneer in social computing. His work on social translucence is about designing social infrastructures that make collective activity visible.[g] The key message is making social clues visible and persistent helps online groups to govern their activities. More recently, ManyEyes has opened up a whole new area of information visualization. Implications and dynamics of social navigation and exploration through information visualization are expected to raise many theoretical and practical questions about the nature of insight and how one may achieve various goals with and without sharing a potentially diverse range of views offered by a growing social network. Relevant readings in this direction include Refs 68–71 on social clarity.

Future Directions

An increasing number of activities and writings have examined the contemporary information visualization field and looked ahead for motivating problems and enlightening challenges that would lead the field to a high level of development.

In 2004 IEEE Visualization, a panel specifically focused on can we determine the top unresolved problems of visualization?[72] Each panelist subsequently published their own lists of unsolved problems. For example, the top 10 unsolved problems[73] identify the need for new methodologies for empirical evaluations and more attention to a better understanding of elementary perceptual–cognitive tasks. The list of problems also includes a better understanding of the role of prior knowledge of viewers in maintaining an effective dialog between information visualization and its users. Many of the top 10 problems remain to be challenges, such as the intrinsic quality measure problem, the scalability problem, and the causality, visual inference, and prediction problem. Some of the problems identified in the more recent visual analytics are also relevant, for example, integrating information-theoretic views, developing metrics for saliency and novelty, and bridging between macroscopic and microscopic views.[10]

Illuminating the Path[37] is an ambitious research agenda for visual analytics. It addresses many issues relevant to information visualization.

A major reflection by leading researchers in both information visualization and scientific visualization is summarized in the 2006 NIH/NSF/VRC Final Report:

the full report[74] and an abridged summary.[75] The report recommends an increase of funding levels from government agencies and industries.

The 2007 Dagstuhl Workshop identified collaborative information visualization and theory building as major directions for future development.[53] In the Visualization Summit 2007, groups of researchers projected where information visualization would be in 10 years.[76] Their report also represents collective thinking of researchers on the future of the field.

The field of information visualization is interdisciplinary in nature. The field will benefit from interconnections with a wide variety of other fields, for example, developing and adopting some of the most promising algorithms for information visualization purposes. For example, some potentially significant directions include scalable, high-performance algorithms developed by machine learning and complex network analysis for analyzing and visualizing large-scale, multidimensional data, such as scalable community finding algorithms[77] and fast EM clustering algorithms,[57] and algorithms of greater interpretability for dimensionality reduction and automated summarization, such as nonnegative matrix factorization and tensor factorization.[78,79]

CONCLUSION

We conclude the overview with some recommendations based on emerging trends and the most promising directions for future research. All recommendations are organized in terms of priority areas. Most of them are interdisciplinary in nature.

Theoretical Foundations

- Pursuing the nature of insight should be broadened to incorporate studies of creativity, discovery, and problem solving in other fields and disciplines.[46,80]
- Theoretical conceptualizations should adapt design and communication frameworks.
- Theory building efforts should integrate information theory and other theories that are capable of defining metrics of information,[81] uncertainty,[82] and interestingness.[83]
- More research should focus on social dimensions of information visualization enabled in social information foraging and social networking.

Metrics

- Information metrics on uncertainty, interestingness, saliency, and rarity.[10]
- Diagnostic and evaluative metrics of information visualization.

Algorithms

- Develop and adopt scalable, high-performance algorithms for analyzing and visualizing large-scale, multidimensional data, such as scalable community finding algorithms[76] and fast EM clustering algorithms.[57]
- Incorporate algorithms of greater interpretability for dimensionality reduction and automated summarization, such as nonnegative matrix factorization and tensor factorization.[77,78]

Design

- Establish design languages and design patterns.
- Develop conceptual and operational taxonomies and enable mixed-initiative interaction.
- Focus on gulfs of execution and gulfs of evaluation.

Visual thinking, focusing on the big picture, pursuing deep insights, and many other characteristics of information visualization make it a compelling candidate across a wide variety of learning and information processing tasks in a diverse range of application domains. Information visualization needs to maintain an open-minded community and reaches out to other disciplines for motivating challenges as well as adaptable techniques. Information visualization is facing not only a challenging and exciting future, but also an increasing expectation and responsibility for an insightful and enlightening world.

NOTES

[a] http://www.math.yorku.ca/SCS/Gallery/milestone/sec1. html.
[b] http://eagereyes.org/blog/2008/sad-state-of-infovis-contest.html.
[c] http://www.napoleonic-literature.com/1812/1812.htm.
[d] http://xxi.ac-reims.fr/fig-st-die/actes/actes_2000/thouez/t13.gif.
[e] http://video.google.com/videoplay?docid=-8252705102362324792.
[f] http://scimaps.org/dev/big_thumb.php?map_id=166.
[g] http://www.visi.com/~snowfall/Soc_Infrastructures_CACMfmt.pdf.

REFERENCES

1. Baker MP, Bushell C. After the storm: considerations for information visualization. *IEEE Comput Graph Appl* 1995, 15:12–15.
2. Wise JA, Thomas JJ, Pennock K, Lantrip D, Pottier M, Schur A, Crow V. Visualizing the non-visual: Spatial analysis and interaction with information from text documents. *Proceedings of IEEE Symposium on Information Visualization '95*, Atlanta, Georgia, USA, 1995, 51–58.
3. Chi EH. *A taxonomy of visualization techniques using the data state reference model*. Proceedings of InfoVis, 2000, Salt Lake City, UT, October 9–10, 2000, (pp. 69–75) .
4. Gaviria AR. When is information visualization art? determining the critical criteria. *Leonardo* 2008, 41:479–482.
5. Bertin J. *Semiology of Graphics: Diagrams, Networks, Maps*. Wisconsin: University of Wisconsin Press; 1983.
6. Cleveland WS, McGill R. Graphical perception: theory, experimentation, and application to the development of graphical methods. *J Am Stat Assoc* 1984, 79: 531–553.
7. Mackinlay JD. Automating the design of graphical presentations of relational information. *ACM Trans Graph* 1986, 5:110–141.
8. Chang R, Ziemkiewicz C, Green TM, Ribarsky W. Defining insight for visual analytics. *IEEE Comput Graph Appl* 2009, 29:14–17.
9. North C. Towards measuring visualization insight. *IEEE Comput Graph Appl* 2006, 26:6–9.
10. Chen C. An information-theoretic view of visual analytics. *IEEE Comput Graph Appl* 2008, 28:18–23.
11. Heer J, Viegas FB, Wattenberg M. Voyagers and voyeurs: supporting asynchronous collaborative visualization. *Commun ACM* 2009, 52:87–97.
12. Bush V. As we may think. *Atlantic Monthly* 1945, 176:101–108.
13. Plaisant C, Fekete J, Grinstein G. Promoting insight based evaluation of visualizations: from contest to benchmark repository. *IEEE Trans Vis Comput Graph* 2008, 14:120–134.
14. Friendly M. A brief history of data visualization. In: Chen C, Härdle W, Unwin A, eds. *Handbook of Computational Statistics: Data Visualization*, vol. 3. Heidelberg: Springer; 2007, 1–34.
15. Friendly M, Denis DJ. (2001, Oct 16, 2008). Milestones in the history of thematic cartography, statistical graphics, and data visualization: An illustrated chronology of innovations Retrieved March 30, 2009, from http://www.math.yorku.ca/SCS/Gallery/milestone/milestone.pdf
16. Furnas GW. Generalized fisheye views. In *Proceedings of CHI '86*, 1986, 16–23.
17. Tufte ER. *The Visual Display of Quantitative Information*. Cheshire, CT: Graphics Press; 1983.
18. Inselberg A. The Plane with Parallel Coordinates. *Visual Computer* 1985, 1:69–91.
19. Wegman E. Hyperdimensional Data Analysis Using Parallel Coordinates. *J Am Stat Assoc* 1990, 85: 664–675.
20. Johnson B, Shneiderman B. *Tree-maps: A space filling approach to the visualization of hierarchical information structures*. In Proceedings of IEEE Visualization 1991, 91, 284–291.
21. Eades P. A heuristic for graph drawing. *Congressus Numerantium* 1984, 42:149–160.
22. Fruchterman TMJ, Reingold EM. Graph drawing by force-directed placement. *Software- Practice and Experience* 1991, 21:1129–1164.
23. Kamada T, Kawai S. A general framework for visualizing abstract objects and relations. *ACM Trans Graph* 1991, 10:1–39.
24. Dwyer T, Marriott K, Schreiber F, Stuckey PJ, Woodward M, Wybrow M. Exploration of networks using overview+detail with constraint-based cooperative layout. *IEEE Trans Vis Comput Graph* 2008, 14:1293–1300.
25. Fairchild K, Poltrock S, Furnas G. SemNet: Three-dimensional graphic representations of large knowledge bases. In: Guidon R, ed. *Cognitive Science and its Applications for Human-Computer Interaction* Hillsdale, NJ: Lawrence Erlbaum Associates; 1988, 201–233.
26. Robertson GG, Mackinlay JD, Card SK. *Cone trees: Animated 3D visualizations of hierarchical information*. In Proceedings of CHI '91, New Orleans, LA, 1991, 189–194.
27. Eick SG, Wills GJ. *Navigating large networks with hierarchies*. In Proceedings of the IEEE Conference on Visualization, Los Alamitos, CA, 1993.
28. Wills GJ. NicheWorks: Interactive visualization of very large graphs Retrieved 4th August, 1998, from http://www.bell-labs.com/user/gwills/NICHEguide/nichepaper.html
29. Chuah MC, Roth SF, Kolojejchick J, Mattis J, Juarez O. *SageBook: Searching data-graphics by content*. In Proceedings of CHI '95, Denver, CO, 1995, 338–345.
30. Kohonen T. *Self-Organizing Maps*: Springer; 1995.

31. Lamping J, Rao R, Pirolli P. A focus+context technique based on hyperbolic geometry for visualizing large hierarchies. In *Proceedings of ACM CHI'95 Conference on Human Factors in Computing Systems*, 1995, 401–408.
32. Munzner T. H3: Laying out large directed graphs in 3D hyperbolic space. In: Dill J, Gershon N, eds. *Proceedings of the 1997 IEEE Symposium on Information Visualization*. Phoenix: IEEE CS Press 1997, 2–10.
33. Card S, Mackinlay J, Shneiderman B, eds. *Readings in Information Visualization: Using Vision to Think*. San Francisco, CA: Morgan Kaufmann; 1999.
34. Chen C. *Information Visualisation and Virtual Environments*. London: Springer; 1999.
35. Spence B. *Information Visualization*. New York: Addison-Wesley; 2000.
36. Ware C. *Information Visualization: Perception for Design*. San Francisco: Morgan Kaufmann Publishers; 2000.
37. Thomas J, Cook K. *Illuminating the Path, the Research and Development Agenda for Visual Analytics*. Phoenix: IEEE CS Press; 2005.
38. Chen C, Paul RJ. Visualizing a knowledge domain's intellectual structure. *Computer* 2001, 34:65–71.
39. Bollen J, Sompel HVd, Hagberg A, Bettencourt L, Chute R, Rodriguez MA, Balakireval L. Clickstream data yields high-resolution maps of science. *PLoS ONE* 2009, 4:e4803.
40. Eccles R, Kapler T, Harper R, Wright W. Stories in GeoTime. *Inf Vis* 2008, 7:3–17.
41. MacEachren AM. *How Maps Work: Representation, Visualization, and Design*. New York: Guilford Press; 1995.
42. Chalmers M. Bead—an information visualization system. *Proc Asis Annu Meet* 1995, 32:249–249.
43. Jerding DF, Stasko JT. The information mural: a technique for displaying and navigating large information spaces. *IEEE Trans Vis Comput Graph* 1998, 4:257–271.
44. Shneiderman B. *The eyes have it: A task by data type taxonomy for information visualization*. In Proceedings of IEEE Workshop on Visual Language, Boulder, CO, 1996, 336–343.
45. Lehrer J. The Eureka Hunt: why do good ideas come to us when they do? *The New Yorker*, July 28, 2008, 40–45.
46. Sternberg RJ, Davidson JE, eds. *The Nature of Insight*. MIT Press; 1996.
47. Chen C, Chen Y, Horowitz M, Hou H, Liu Z et al. Towards an explanatory and computational theory of scientific discovery. *J Informetrics* 2009, 3:191–209.
48. Plaisant C, Fekete JD, Grinstein G. Promoting insight-based evaluation of visualizations: from contest to benchmark repository. *IEEE Trans Vis Comput Graph* 2008, 14:120–134.
49. Lam H. A framework of interaction costs in information visualization. *IEEE Trans Vis Comput Graph* 2008, 14:1149–1156.
50. Norman, DA. *The Design of Everyday Things*. New York: Basic Books; 2002.
51. Chen C, Yu Y. Empirical studies of information visualization: a meta-analysis. *Int J Hum Comput Stud* 2000, 53:851–866.
52. Julien CA, Leide JE, Bouthillier F. Controlled user evaluations of information visualization interfaces for text retrieval: Literature review and meta-analysis. *J Am Soc Inform Sci Technol* 2008, 59:1012–1024.
53. Kerren A, Stasko JT, Fekete J-D, North C. Workshop report: information visualization—human-centered issues in visual representation, interaction, and evaluation. *Inf Vis* 2007, 6:189–196.
54. Liu ZC, Nersessian NJ, Stasko JT. Distributed Cognition as a Theoretical Framework for Information Visualization. *IEEE Trans Vis Comput Graph* 2008, 14:1173–1180.
55. Eick SG. Information visualization at 10. *IEEE Comput Graph Appl* 2005, 25:12–14.
56. Gray AG, Moore AW. N-body problems in statistical learning. *Adv Neural Inf Process Syst* 2001, 13:521–527.
57. Moore A. Very fast EM-based mixture model clustering using multiresolution kd-trees. In: Kearns M, Cohn D, eds. *Advances in Neural Information Processing Systems*: Morgan Kaufman; 1999, 543–549.
58. Holten D. Hierarchical edge bundles: visualization of adjacency relations in hierarchical data. *IEEE Trans Vis Comput Graph* 2006, 12:741–748.
59. Cui W, Zhou H, Qu H, Wong PC, Li X. Geometry-based edge clustering for graph visualization. *IEEE Trans Vis Comput Graph* 2008, 14:1277–1284.
60. Dwyer T, Koren Y, Marriott K. IPSep-CoLa: An Incremental Procedure for Separation Constraint Layout of Graphs. *IEEE Trans Vis Comput Graph* 2006, 12:821–828.
61. Bottger J, Balzer M, Deussen O. Complex logarithmic views for small details in large contexts. *IEEE Trans Vis Comput Graph* 2006, 12:845–852.
62. Luboschik M, Schumann H, Cords H. Particle-based labeling: fast point-feature labeling without obscuring other visual features. *IEEE Trans Vis Comput Graph* 2008, 14:1237–1244.
63. Chen C. Tracking latent domain structures: an integration of pathfinder and latent semantic analysis. *AIDS Soc* 1997, 11:48–62.
64. Chen C, Morris S. Visualizing evolving networks: Minimum spanning trees versus Pathfinder networks. In *Proceedings of IEEE Symposium on Information Visualization*, Seattle, Washington, 2003, 67–74.

65. Quirin A, Cordon O, Santamaria J, Vargas-Quesada B, Moya-Anegon F. A new variant of the Pathfinder algorithm to generate large visual science maps in cubic time. *Inform Process Manage* 2008, 44:1611–1623.

66. Schvaneveldt RW, ed. *Pathfinder Associative Networks: Studies in Knowledge Organization*. Norwood, NJ: Ablex Publishing Corporations; 1990.

67. Healey CG, Kocherlakota S, Rao V, Mehta R, Amant RS. Visual perception and mixed-initiative interaction for assisted visualization design. *IEEE Trans Vis Comput Graph* 2008, 14:396–411.

68. Chen H, Nunamaker J Jr, Titkova O. Information visualization for collaborative computing. *Computer* 1998, 31:75–82.

69. Isenberg P, Carpendale S. Interactive tree comparison for co-located collaborative information visualization. *IEEE Trans Vis Comput Graph* 2007, 13:1232–1239.

70. Wattenberg M, Kriss J. Designing for social data analysis. *IEEE Trans Vis Comput Graph* 2006, 12:549–557.

71. Kostelnick C. The visual rhetoric of data displays: the conundrum of clarity. *IEEE Trans Professional Commun* 2008, 51:116–130.

72. Rhyne T-M, Hibbard B, Johnson C, Chen C, Eick S. Can we determine the top unresolved problems of visualization? In *Proceedings of IEEE Visualization 2004*, Austin, Texas, 2004, 563–566.

73. Chen C. Top 10 unsolved information visualization problems. *IEEE Comput Graph Appl* 2005, 25:12–16.

74. Johnson CR, Moorhead R, Munzner T, Pfister H, Rheingans P et al., eds. *NIH-NSF Visualization Research Challenges Report*: IEEE Press; 2006.

75. Munzner T, Johnson C, Moorhead R, Pfister H, Rheingans P, Yoo T. NIH-NSF visualization research agenda report summary. *IEEE Comput Graph Appl* 2006, 26:20–24.

76. Burkhard RA, Andrienko G, Andrienko N, Dykes J, Koutamanis A et al. Visualization summit 2007: ten research goals for 2010. *Inf Vis* 2007, 6:169–188.

77. Clauset A, Newman MEJ, Moore C. Finding community structure in very large networks. *Phys Rev E* 2004, 70:066111.

78. Berry MW, Browne M, Langville AN, Pauca VP, Plemmons RJ. Algorithms and applications for approximate nonnegative matrix factorization. *Comput Stat Data Anal* 2007, 52:155–173.

79. Lee DD, Seung HS. Learning the parts of objects by nonnegative matrix factorization. *Nature* 1999, 401:788–791.

80. Dunbar K. Concept discovery in a scientific domain. *Cogn Sci* 1993, 17:397–434.

81. Soofi ES, Retzer JJ. Information indices: unification and applications. *J Econom* 2002, 107:17–40.

82. Bartels L. Issue voting under uncertainty: an empirical test. *Am J Pol Sci* 1988, 30:709–728.

83. Itti L, Baldi P. A Principled Approach to Detecting Surprising Events in Video. *Proceedings of IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2005, 631–637.

FURTHER READING

Card S, Mackinlay J, Shneiderman B, eds. *Readings in Information Visualization: Using Vision to Think*. San Francisco, CA: Morgan Kaufmann; 1999.

Chen C. *Information Visualization: Beyond the Horizon*. 2nd ed. London: Springer; 2004.

Chen C. *Mapping Scientific Frontiers: The Quest for Knowledge Visualization*. London: Springer-Verlag; 2003.

Kraak MJ. Geovisualization illustrated. *ISPRS J Photogramm Remote Sensing* 2003, 57: 390–399.

MacEachren AM. *How Maps Work: Representation, Visualization, and Design*. New York: Guilford Press; 1995.

Spence B. *Information Visualization*. London: Addison-Wesley; 2001.

Thomas J, Cook K. *Illuminating the Path, the Research and Development Agenda for Visual Analytics*: IEEE CS Press; 2005.

Tufte ER. *Visual Explanations*. Cheshire, CT: Graphics Press; 1997.

Tufte ER. *Envisioning Information*. Cheshire, CT: Graphics Press; 1990.

Ware C. *Visual Thinking for Design*: Morgan Kaufmann; 2008.

