Chapter 2
An Illuminated Path: The Impact of the Work of Jim Thomas

Chaomei Chen, Haiyan Hou, Zhigang Hu, and Shengbo Liu

Abstract   We provide a glimpse of the impact of the most influential work of Jim Thomas on the development of visual analytics. A quantitative visual analytic approach is applied to the analysis of three datasets of scholarly publications in order to highlight the nature and profoundness of the impact in terms of the structure and dynamics of the underlying intellectual landscape. We investigate and provide intriguing evidence of the impact of the groundbreaking book in visual analytics, *Illuminating the Path*, in terms of its role in shaping the scientific literature of relevant fields. Furthermore, we analyze the intellectual structure of visual analytics in a broader context in order to set the impact of *the Illuminating the Path* in perspective.

2.1 Introduction

Visual analytics has transformed not only how we visualize complex and dynamic phenomena in the new information age, but also how we may optimize analytical reasoning and make sound decisions with incomplete and uncertain information (Keim et al. 2008). Today's widespread recognition of the indispensable value of visual analytics as a field and the rapid growth of an energetic and interdisciplinary scientific community would be simply impossible without the remarkable vision and tireless efforts of Jim Thomas, his colleagues of the National Visualization and Analytics Center (NVAC) at Pacific Northwest National Laboratory (PNNL), and the growing community in visual analytics science and technology.

James J. Thomas (March 26, 1946–August 6, 2010), an internationally recognized visionary scientist and inspirational leader in the field of visualization, motivated researchers across the world by creating and promoting the visual analytics field and by expanding the boundaries of computer graphics, user interface software,

C. Chen (✉)
College of the Information Science and Technology, Drexel University, Philadelphia, PA, USA
e-mail: cc345@drexel.edu

H. Hou · Z. Hu · S. Liu
WISELab, Dalian University of Technology, Dalian, China

and visualization (Christopher Columbus Fellowship Foundation 2009; Ebert et al. 2010).

Jim Thomas started his professional career as a computer graphics researcher and worked on computer graphics systems to support the then new computer-aided design process and rose to the top of the computer graphics community within the next decade (Ebert et al. 2010; Wong 2010). He published extensively and invented continuously. His work has inspired many generations of researchers and practitioners. His contributions to the society were recognized by international and national awards for transferring research technology to industry and universities (Christopher Columbus Fellowship Foundation 2009; Kasik et al. 2010).

Jim was a pioneer and a leader in many areas. He was the founder and the first Director of the Department of Homeland Security National Visualization and Analytics Center (NVAC). He led major technology initiatives and research programs (Christopher Columbus Fellowship Foundation 2009; Kasik et al. 2010). He led a team of PNNL scientists and developed the groundbreaking system SPIRE (Wong 2010). Its successor, IN-SPIRE, is still in use today.

In 2004, he founded NVAC and opened a new research area, visual analytics, to the world. Visual analytics is the science of analytical reasoning facilitated by visual interactive interfaces that focuses on analytical reasoning facilitated by interactive visual interfaces (Thomas and Cook 2005; Wong and Thomas 2004). Visual analytics is a multidisciplinary field. It brings together several scientific and technical communities from computer science, information visualization, cognitive and perceptual sciences, interactive design, graphic design, and social sciences. It addresses challenges involving analytical reasoning, data representations and transformations, visual representations and interaction techniques, and techniques to support production, presentation, and dissemination of the results. Although visual analytics has some overlapping goals and techniques with information visualization and scientific visualization, it is especially concerned with sense-making and reasoning and it is strongly motivated by solving problems and making sound decisions.

Visual analytics integrates new computational and theory-based tools with innovative interactive techniques and visual representations based on cognitive, design, and perceptual principles. This science of analytical reasoning is central to the analyst's task of applying human judgments to reach conclusions from a combination of evidence and assumptions (Thomas and Cook 2005). Under Jim's influence, visual analytics centers were established throughout the world, including Canada, Germany, and the United Kingdom; and universities integrated visual analytics into their core information sciences curricula which made the new field a recognized and promising outgrowth of the fields of information visualization and scientific visualization (Wong 2010).

In order to supplement the recognition of Jim's contributions by a wide variety of domain experts and leaders in science and technology, we take a unique approach and highlight patterns and trends that can be discerned computationally from the scientific literature.

In this chapter, we present several studies of three datasets, which increasingly expand the scope of our study. We begin with papers authored by Jim Thomas. Then

we trace patterns found in papers that cited *Illuminating the Path*. Finally, we expand our analysis to a broader context defined by visual analytics and superimpose the impact of *Illuminating the Path* over the broader intellectual landscape.

2.2 Three Datasets

In this chapter, we analyzed three datasets: D1: 13 papers authored by Jim Thomas, D2: 135 papers that cited *the Illuminating the Path*, and D3: 2,238 papers on broadly defined visual analytics. The three datasets provide an increasingly broader view of the impact of the contributions made by Jim Thomas.

We retrieved relevant bibliographic records for D1 by using a simple topic search for publications of Jim Thomas in the Web of Science, which is the most authoritative source of scientific publications in multiple disciplines. We used the search query: AU="Thomas J*" and ZP=99352 or AU="Thomas J*" and AD="Pacific NW Natl Lab". This retrieved articles that are authored by Thomas J or Thomas JJ along with PNNL's name and zip code to ensure that this is the Jim Thomas from PNNL. According to Ebert et al. (2010), Kasik et al. (2010), PNNL was the address he used since 1976.

The topic search returned 13 papers contributed by Jim Thomas (or James Joseph Thomas), and all these papers were published after 1996. These 13 papers form the first dataset D1. The most cited one is a 2006 article entitled "A visual analytics agenda" by Thomas JJ and Cook KA. It was cited 31 times. Figure 2.1 shows a network of collaborating authors derived from the 13 papers. The network visualization was produced by CiteSpace, a freely available tool for visualizing and analyzing emerging patterns and changes in scientific literature (Chen 2004, 2006; Chen et al. 2009, 2010). Two authors are connected if they coauthored one of the 13 papers. Lines are color coded to indicate the time of collaboration. Blue lines are the earliest, followed by green, yellow, and orange. Labels in red are key phrases extracted from the titles of papers where they collaborated. For example, #1 visual analytics introduction is from the title of a paper coauthored by Jim Thomas with Joe Kielman at the Department of Homeland Security and Richard May at PNNL.

Figure 2.2 shows a network of keywords extracted from the titles and abstracts of the 13 papers. Keywords such as information visualization, visual analytics, and visualization rightly dominate the network. The labels in red characterize clusters of keywords. For example, #0 state of practice is from an article that focused on visual analytics, future, challenges, domains and applications, as well as the state of the practice.

The second dataset D2 was generated to show the impact of the book *Illuminating the Path: The Research and Development Agenda for Visual Analytics*, coauthored by Thomas and Cook (2005). It is widely known as a landmark of visual analytics. We searched for papers that cited the work in the Web of Science using a citation search for Thomas J* and Illum*. The dataset D2 contains 135 citing papers. We analyzed a variety of patterns and properties of the dataset D2. Then we used the
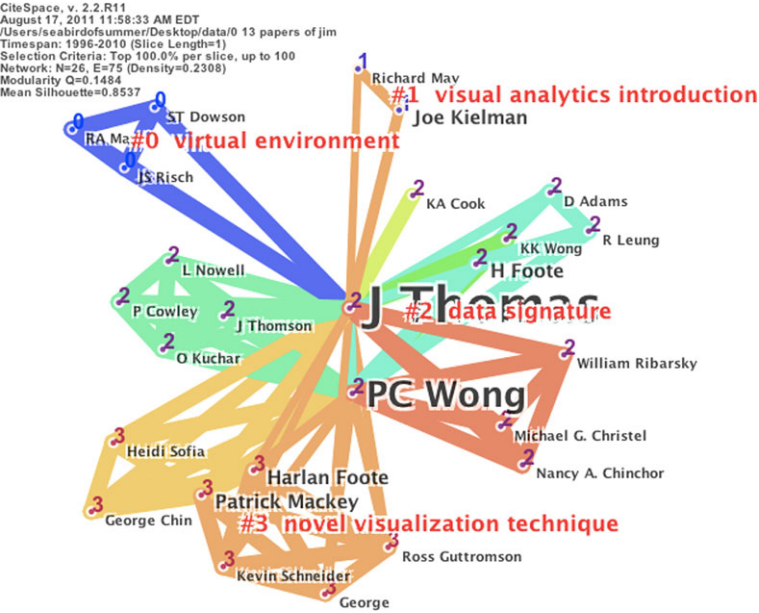
CiteSpace, v. 2.2.R11
August 17, 2011 11:58:33 AM EDT
/Users/seabirdofsummer/Desktop/data/0 13 papers of jim
Timespan: 1996-2010 (Slice Length=1)
Selection Criteria: Top 100.0% per slice, up to 100
Network: N=26, E=75 (Density=0.2308)
Modularity Q=0.1484
Mean Silhouette=0.8537

Richard May
#1 visual analytics introduction
Joe Kielman
ST Dowson
RA Ma #0 Virtual environment
JS Risch
KA Cook   D Adams
                 R Leung
            KK Wong
L Nowell   H Foote
P Cowley   #2 data signature
J Thomson   J Thomas
O Kuchar   William Ribarsky
            PC Wong
Heidi Sofia   Michael G. Christel
                     Nancy A. Chinchor
Harlan Foote
Patrick Mackey
George Chin   #3 novel visualization technique
                        Ross Guttromson
Kevin Schneider
          George

Fig. 2.1 A collaboration network of Jim Thomas

CiteSpace, v. 2.2.R11
August 18, 2011 11:40:23 AM EDT
/Users/seabirdofsummer/Desktop/data/0 13 papers of jim
Timespan: 1996-2010 (Slice Length=1)
Selection Criteria: Top 100.0% per slice, up to 100 visualization techniques and methodologies
Network: N=34, E=119 (Density=0.2121)
Modularity Q=0.505
Mean Silhouette=0.7973
            #4 visualization techniques and methodologies
            data and knowledge visualization
applications
            data visualization
            #2 data visualization
            google
information visualization
knowledge management
            future challenges
                         state of practice
visual para digital libraries visual an #0 ti state of practice
            #3 strain information visualization
            higher-order interaction
digital content and media   domains and applications
human-computer interaction
            graph analytics
            #1 future challenges visual analytics
visualization networks
fractal curve very large sparse graphs
digital image processing multiscale visualization
strain
            gestalt psychology
      pathogen helicobacter-pylori
sequence #5 chlamydia–pneumoniae
      chlamydia–pneumoniae   retrieval
whole genome alignment   #6 images
            bioinformatics   of-the-art
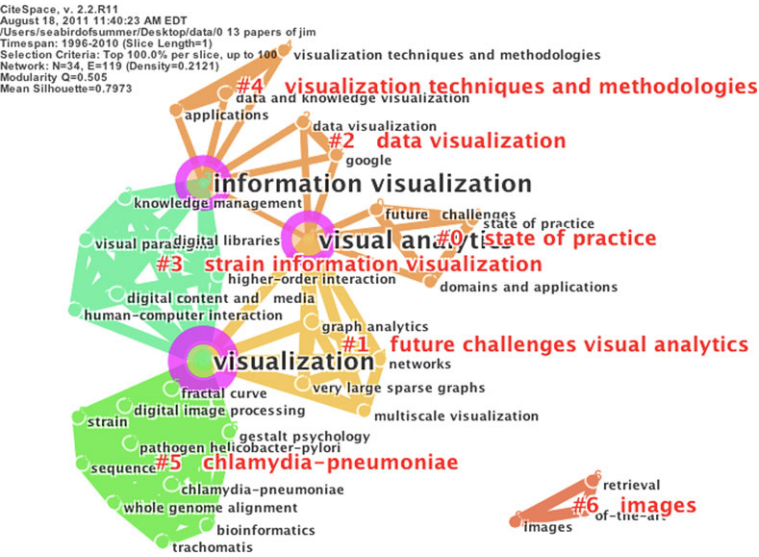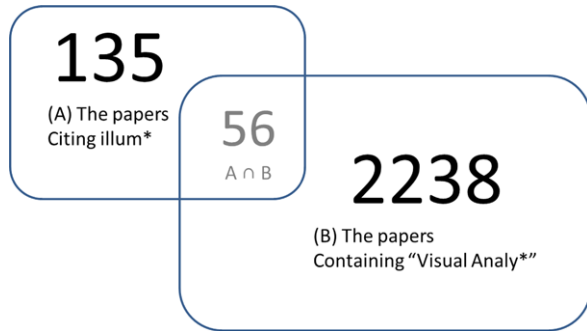   trachomatis   images

Fig. 2.2 Key topics in the 13 papers (D1)

CiteSpace software to generate a geospatial map of these authors to show coun-
tries where the impact of the book was evident. In order to obtain insights into the

Fig. 2.3 Datasets D2 (**A**) and D3 (**B**) analyzed in this chapter

collaborative activities across institutions and countries, we generated a co-author network from the 135 citing papers. Finally, a document co-citation network was constructed and analyzed to highlight how the influence of the book evolved and spread.

In order to investigate the influence of Jim Thomas in the broader context of visual analytics, we searched for topics in the Web of Science for TS="Visual Analy*". The wildcard * was used to cover variations such as analytics, analytical, and analysis. The search resulted in 2,238 papers relevant to visual analytics. The third dataset D3 consists of these 2,238 records. As attested by the 2,238 papers, the most productive institute is PNNL! The single work that has influenced 2,238 papers is, intriguingly, the *Illuminating the Path*! The most cited author is Jim Thomas! What a remarkable tribute to Jim Thomas (Fig. 2.3).

2.3 Illuminating the Path

Jim Thomas and Kristin A. Cook co-authored "Illuminating the Path: The Research and Development Agenda for Visual Analytics". This book motivated researchers and developers to help safeguard the nation by transforming information overload into insights through visual analytics. It has provided the roadmap of visual analytics ever since: the science of analytical reasoning facilitated by interactive visual interfaces. This book had a profound impact on the course of the development of visual analytics.

2.3.1 The Spread of the Impact

The geographic map of authors and collaborative authorship were generated by CiteSpace and can be viewed interactively in Google Earth. Figure 2.4 revealed the spread of the geographic distribution of the citing papers over time. A colored disk in the map represents the location of an author. The line connecting two locations indicates that authors at the two locations collaborated on a published paper.
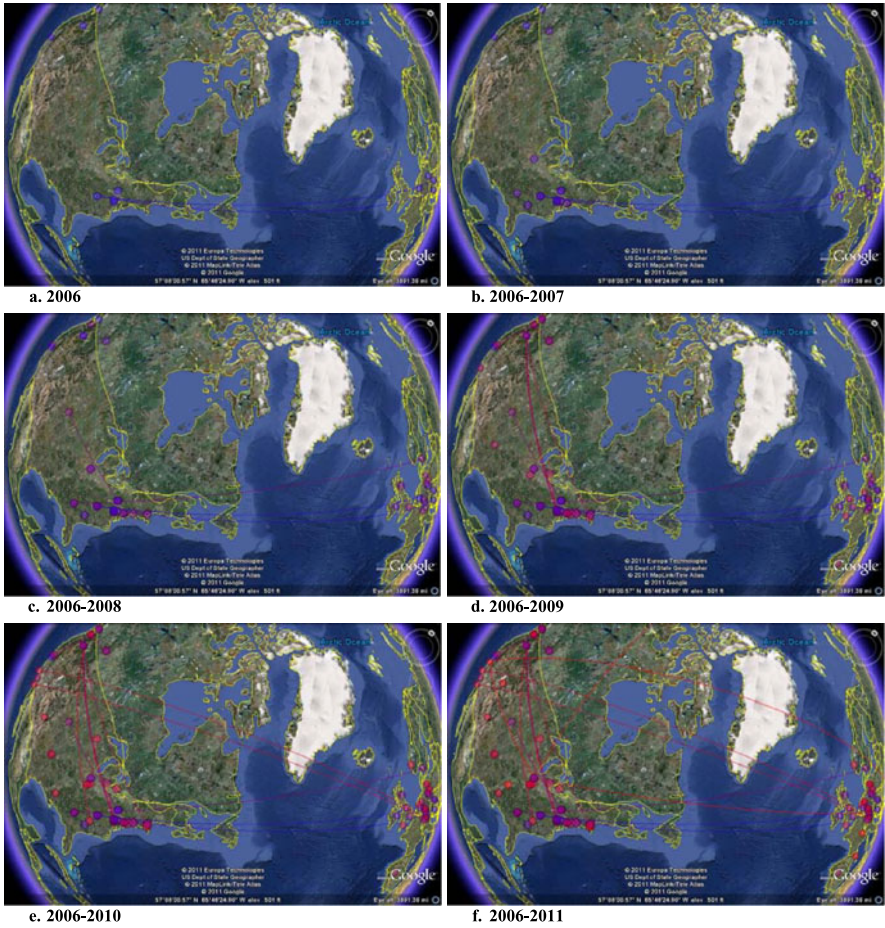
Fig. 2.4 Geographic maps of the growing influence of *Illuminating the Path*. Data source: D2

As shown in Fig. 2.4a, in 2006, this book was mainly cited on the east and west coasts of the United States, a small number of cities in European countries, such as Germany, Austria, and Netherlands. U.S.-European collaboration was evident by the blue lines across the Atlantic Ocean. In 2007, the impact of this book expanded to include more southern cities on the coast of the U.S., more European countries, and more cities in Canada (Fig. 2.4b). In the period 2008 and 2009, the influence of this book emerged in the middle of the United States with long-range collaboration links. Meanwhile, the impact of the work spread to more European cities in Sweden, Spain, England, Scotland, and Switzerland (Fig. 2.4c and 2.4d). More recently, the impact has reached more and more cities in the U.S. and other parts of the world. International collaboration increased over time with countries such as China, Israel, and some Europe countries (Fig. 2.4e and 2.4f).
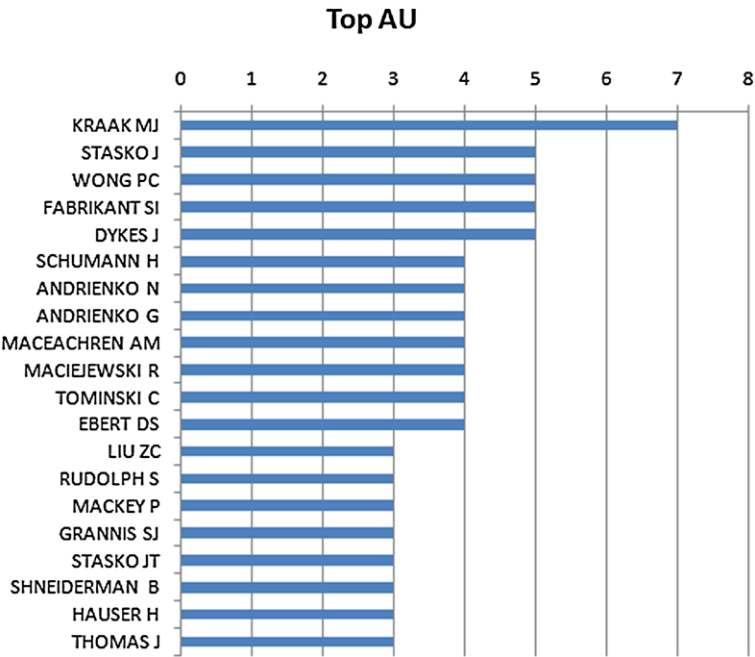
Top AU

Fig. 2.5 Top 20 authors in D2 who frequently cited *Illuminating the Path*

2.3.2 The Inspired Community

There are a total of 390 citing authors in the 135 papers in dataset D2. Figure 2.5 shows the top 20 most productive authors. Menno Jan Kraak, a researcher in geographic visualization, cited *Illuminating the Path* in 7 of his papers. John Stasko, Pak Chung Wong, and a few others have cited the book in 5 of their papers.

Figure 2.6 shows a collaboration network of authors who cited *Illuminating the Path*. A group of eight authors formed in 2010, including Menno Jan Kraak, Jason Dykes, Gennady Andrienko, Natalia Andrienko, and Sara Irina Fabrikant. Most of them are known for their work in geographic visualization. They are from different institutions in different countries such as Fraunhofer Institute for Intelligent Analysis and Information Systems, Germany, the Faculty of Geo-Information Science and Earth Observation of the University of Twente, the Netherlands, and the Department of Information Science of City University London, England. They all took part in a book entitled *Exploring Geovisualization*. These authors worked together in geovisualization in 2010. Pak Chung Wong, John Stasko, and William Ribarsky are well known leaders in visual analytics among other fields of their expertise.
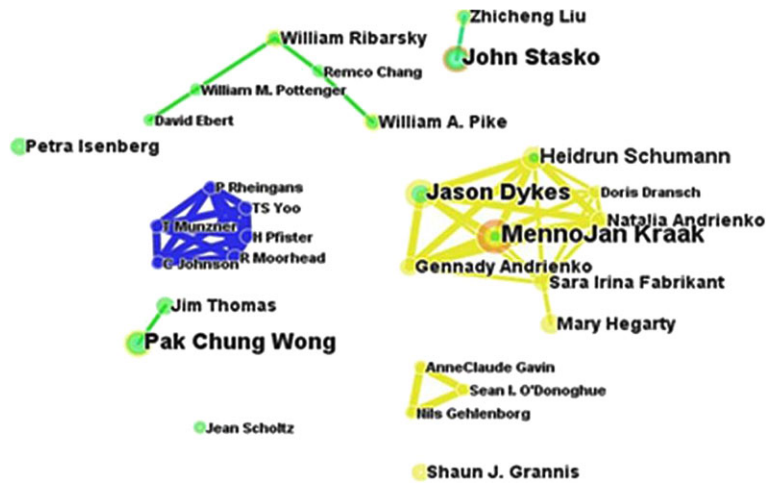
Fig. 2.6 A network of coauthors (2006–2011). Data source: D2

2.3.3 A Document Co-citation Analysis

Document co-citation analysis (DCA) is an established method to delineate the structure and dynamics of the intellectual landscape of subject areas in scientific literature. We conducted a DCA of articles that cited *Illuminating the Path* so that we could identify research topics that have been influenced by the work.

CiteSpace was used to generate the co-citation network of the citing papers. We used six one-year time slices between 2006 and 2011. The threshold was set by (c, cc, ccv) method. We set the threshold at (2, 2, 12), (3, 2, 20), (3, 2, 20). The pathfinder algorithm was used to obtain the important information on the network.

Figure 2.7 shows that the co-citation network revealed 11 topic areas with labels identified by noun phrases extracted from titles of papers that are responsible for each cluster. Labels were selected by log-likelihood ratio tests. The shaded areas highlight the boundaries of clusters. The largest node (most cited) in the visualization is *Illuminating the Path* on the upper left corner. It also has the strongest betweenness centrality, showing as a thick purple ring, which indicates its central role in the structure. It is not surprising because of the way the data was collected, but more interestingly it shows its role in forming the basis of visual analytics. And this book is direct related to cluster 0, 1, 2, 3, 4 and 10. The topic words in each cluster are listed in Table 2.1. The major topic areas are visualization, including information visualization, visualization design, and visualization model. System biology and geospatial visual analytic were also influenced by this book.
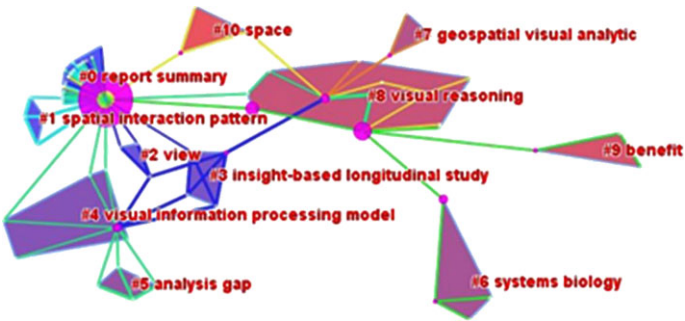
Fig. 2.7 A document co-citation network showing the impact of *Illuminating the Path*. Data source: D2

Table 2.1 Topic words of 11 topic areas influenced by *Illuminating the Path*

Cluster ID	Topic word 1	Topic word 2	Topic word 3
0	Report summary	Analysis	NIH-NSF visualization research
1	Spatial interaction pattern	Pandemic decision support	Predictive analytics approach
2	View	Visualization	Report summary
3	Insight-based longitudinal study	Function graph	Family
4	Visual information processing model	Interactive visualization environment	Visualization design
5	Analysis gap	Visual generation	Promoting insight-based evaluation
6	Systems biology	Omics data	Conjunctive visual form
7	Geospatial visual analytic	Mortality data	Construction management data visualization
8	Visual reasoning	Top-down perspective	Mental model
9	Benefit	Document collection	Co-located visual analytic
10	Space	Efficiency	Eye movement recording

2.3.4 Major Co-citation Clusters

Co-citation clusters represent codified knowledge in a field of study. The network generated from dataset D2 contains 11 co-citation clusters. These clusters are labeled by title terms from their own citers. The largest 4 clusters are summarized in Table 2.2.

The largest cluster (#0) has 21 members and a silhouette value of 0.949. It is labeled as a report summary by both LLR and TFIDF, and as visual analysis by MI.

Table 2.2 Summary of the 4 largest clusters

Cluster ID	Size	Silhouette	Label (TFIDF)	Label (LLR)	Label (MI)	Most Active Citer
0	21	0.949	(6.99) report summary	report summary (79.72, 1.0E-4)	visual analysis	visual analysis
8	13	0.809	(5.26) standards-based 3d geovisualization	visual reasoning (87.91, 1.0E-4)	interaction research	interaction research
4	6	0.673	(3.35) visual information processing model	visual information processing model (47.89, 1.0E-4)	visual information processing model	visual information processing model
6	5	0.93	(4.24) future	systems biology (39.69, 1.0E-4)	visualizing biological data-now	visualizing biological data-now

The most active citer to the cluster is *NIH-NSF visualization research challenges report summary*. The second largest cluster (#8) has 13 members and a silhouette value of 0.809. It is labeled as visual reasoning by LLR, a standards-based 3d geo-visualization by TFIDF, and interaction research by MI. The most active citer to the cluster is mental models, visual reasoning and interaction in information visualization: a top-down perspective. The third largest cluster (#4) has 6 members and a silhouette value of 0.673. It is labeled as a visual information processing model by both LLR and TFIDF, and as visual information processing model by MI. The most active citer to the cluster is a visual information processing model to characterize interactive visualization environments. The 4th largest cluster (#6) has 5 members and a silhouette value of 0.93. It is labeled as systems biology by LLR, future by TFIDF, and visualizing biological data-now by MI.

2.3.5 Landmark Papers

A number of landmark papers were co-cited with the *Illuminating the Path* in Fig. 2.7. These papers were ranked by citation counts, betweenness centrality, burst, and sigma (Chen et al. 2009, 2010).

2.3.5.1 Citation Counts

Table 2.3 lists top 10 references ranked by citation counts. The most cited one is CARD SK (1999) in Cluster #8, with a citation count of 19. This is the widely known "Readings in information visualization: using vision to think", a book of exemplar works compiled by Stuart Card, Jock Mackinlay and Ben Shneiderman. The second is SHNEIDERMAN B (1996) in Cluster #8, with a citation count of 15. The third is TUFTE E (2001) in Cluster #4, with a citation count of 10. The 4th is LARKIN JH (1987) in Cluster #10, with a citation count of 10. The 5th is Andrienko G (2007) in Cluster #7, with a citation count of 9.

2.3.5.2 Betweenness Centrality

Table 2.4 lists the top 10 ranked items by betweenness centrality. Articles with strong centrality scores are believed to be fundamental to the intellectual work of the community. The top ranked item by centrality is again CARD SK (1999) in Cluster #8, with a centrality of 0.67. The second is Mackinlay JD (2007) in Cluster #8, with a centrality of 0.53. The third is SHNEIDERMAN B (1996) in Cluster #8, with a centrality of 0.38. The 4th is MACKINLAY J (1986) in Cluster #4, with a centrality of 0.36. The 5th is Plaisant C (2008) in Cluster #6, with a centrality of 0.22.

Table 2.3 Top 10 ranked items by citation counts

Citation counts	References	Cluster #
19	CARD SK, 1999, READINGS INFORM VISU, V, P	8
15	SHNEIDERMAN B, 1996, P IEEE S VIS LANG, P336, P	8
10	TUFTE E, 2001, VISUAL DISPLAY QUANT, V, P	4
10	LARKIN JH, 1987, COGNITIVE SCI, V11, P65	10
9	Andrienko G, 2007, INT J GEOGR INF SCI, V21, P839	7
9	WARE C, 2004, INFORM VISUALIZATION, V, P	9
8	MACKINLAY J, 1986, ACM T GRAPHIC, V5, P110	4
8	Yi JS, 2007, IEEE T VIS COMPUT GR, V13, P1224	8
8	JOHNSON C, 2006, NIH NSF VISUALIZATION, V, P	4
8	PLAISANT C, 2004, P WORK C ADV VIS INT, P109, P	6

Table 2.4 Top 10 ranked items by betweenness centrality

Centrality	References	Cluster #
0.67	CARD SK, 1999, READINGS INFORM VISU, V, P	8
0.53	Mackinlay JD, 2007, IEEE T VIS COMPUT GR, V13, P1137	8
0.38	SHNEIDERMAN B, 1996, P IEEE S VIS LANG, P336, P	8
0.36	MACKINLAY J, 1986, ACM T GRAPHIC, V5, P110	4
0.22	Plaisant C, 2008, IEEE T VIS COMPUT GR, V14, P120	6
0.17	PLAISANT C, 2004, P WORK C ADV VIS INT, P109, P	6
0.16	Aigner W, 2008, IEEE T VIS COMPUT GR, V14, P47	10
0.14	WARE C, 2004, INFORMATION VISUALIZ, V, P	3
0.12	Andrienko G, 2007, INT J GEOGR INF SCI, V21, P839	7
0.12	Isenberg P, 2007, IEEE T VIS COMPUT GR, V13, P1232	9

2.3.5.3 Burst and Sigma

Articles with citation burst and sigma are regarded as having drawn a considerable amount of attention from the scientific community. There is only burst item in this network. It is WARE C (2004) in Cluster #9, with bursts of 2.38 and a sigma value 1.15.

2.3.6 Timeline View

Figure 2.8 shows a timeline visualization of the co-citation network. According to the colors of the co-citation lines, clusters #0, #1 to #5 were formed in the period 2006–2007. The co-cited papers are mostly before year 2000. They co-cited in the
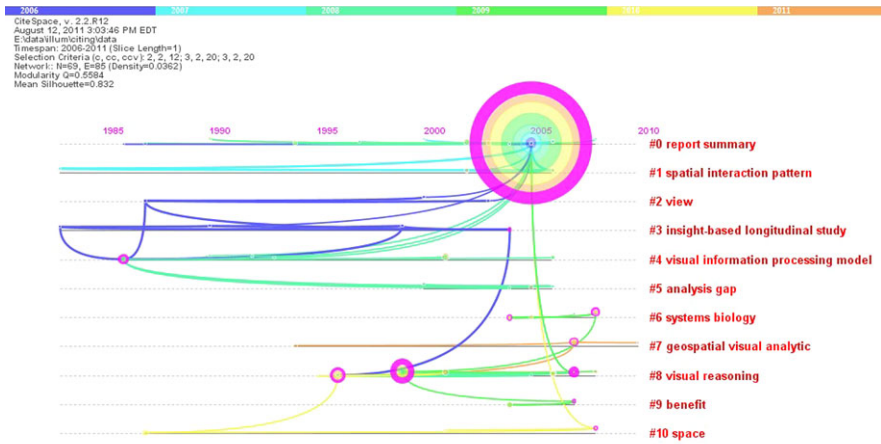
Fig. 2.8 A timeline visualization of the co-citation network. Data source: D2

field of visualization analysis and method. Clusters #6 to #10 are generated after 2007. The co-cited papers are around the year 2005. These research focii were on the application of visualization, including geospatial visual analytic, conjunctive visual form, and co-located with visual analytic.

2.4 A Broader Context

In order to set the impact of *Illuminating the Path* in a broader context, we further investigated patterns revealed by the third dataset D3, which contains 2,238 articles relevant to visual analytics. Based on this data, we identified the trend of growth, major source journals, hot topics, highly cited documents, and highly cited authors. In this way, the role of *Illuminating the Path* and the transformative impact of Jim Thomas' work can be better understood.

2.4.1 The Trend of Growth

Although the earliest article in dataset D3 dated as early as 1948, the majority of the domain began to emerge in the early 1990s. About 95 % of the articles (2,125 out of 2,238) in D3 were published after 1990. Based on the different growth rates, the overall trend since 1990 can be divided into three phases (Fig. 2.9). During the first phase (1990–1999), the number of publications per year increased from less than 50 to around 100. During the second phrase (1999–2004), the productivity per year was relatively stable, producing about 90 publications per year. During the third phase, starting from year 2005, the increase became much faster. For example,
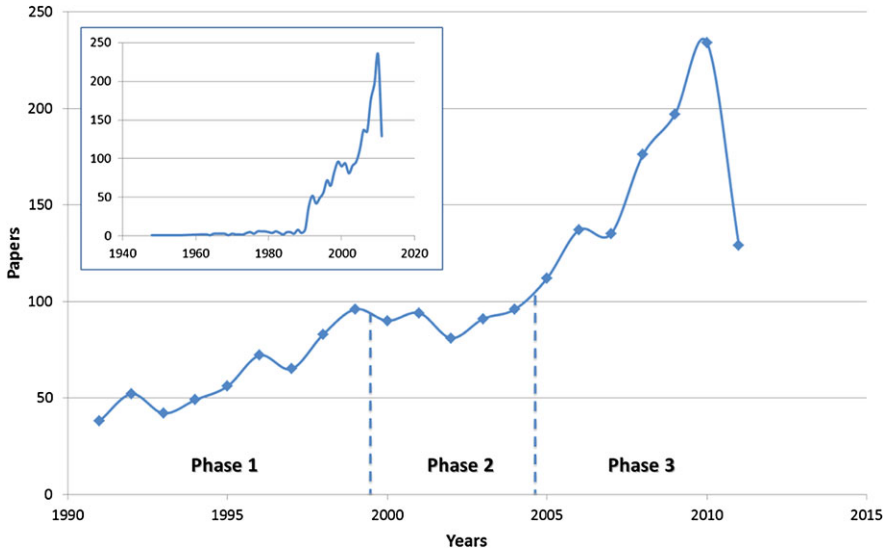
Fig. 2.9 The growth of the domain of visual analytics. Data source: D3

234 papers were published in 2010 alone, which more than doubled the number of papers published in these areas in 2004. We analyzed these three phases separately in order to characterize each phase further.

2.4.2 Major Source Journals and Hot Topics

We were interested in the following questions: what are the major intellectual works in the domain and how did Illuminating the Path transform the intellectual landscape? Articles containing the words 'visual' and 'analysis' are not necessarily relevant to today's visual analytics. Nevertheless, they provide a broader context to locate where exactly the impact of *Illuminating the Path* was.

Table 2.5 lists the top sources and the hot keywords of the 2,125 publications between 1991 and 2011. The *Journal of nuclear medicine* (*JNM*) is the major source of publication (71 papers) in this period. It is one of the top medical imaging journals worldwide with a high impact factor of 6.424 according to the newly released 2009 Journal Citation Reports. The second source journal close behind is *IEEE Transactions on Visualization and Computer Graphics* (*TVCG*). 70 papers were published in this journal. TVCG is a scholarly archival journal published monthly started in 1995. It involves subjects related to computer graphics and visualization techniques, systems, software, hardware, and user interface issues. TVCG is in computer science. Obviously, these two journals represent two quite different research fields.

Other top-ranked computer science source journals include Information Visualization, Computer Graphics Forum, and IEEE Computer Graphics and Applications.

JNM, on the other hand, belongs to Nuclear Medicine Imaging, with relevant journals such as Nuclear Medicine Communications, European Journal of Nuclear Medicine and Molecular Imaging, and Neuroimage. Is JNM a part of the modern visual analytics? Or is it connected purely by association of the words used in the query, namely visual and analy*? If it is not conceptually connected to visual analytics per se, would it be viable to consider potential connections in the future?

It should be noted that these two fields are distinct in terms of their emergence and evolution. Overall, the journals in Computer Science appeared later than the journals in the field of Medical Imaging. The latter ranked highly since 1991, while the former appeared only in the recent phase. In the top source journal TVCG for example, 67 of 71 papers in this journal were published between 2005 and 2011. Similarly, in Information Visualization and Computer Graphics Forum, all the publications sourced from them were published after 2005.

As far as the evidence from the literature is concerned, the year 2005 was a watershed for this domain even if it was probably not yet a single and conceptually integrated one. Since 2005, not only the scale of the literature expanded rapidly, but also new research directions emerged. According to the analysis of keywords, we can draw similar conclusions. Before 2005, the keywords were visual analysis, PET, SPECT, Epilepsy, and Magnetic Resonance Imaging (MRI). After 2005, the hot topics switched to Visual Analytics, Visual Analysis, Information Visualization, and Visualization. Remarkably, the new term "Visual Analytics" began to appear in 2005, and finally it became the most frequently used keyword (71 times). It is evident that *Illuminating the Path* defined, or redefined, whatever was meant by visual analysis before and after 2005.

2.4.3 Highly Cited Documents and Authors

As we know, the mission and agenda of Visual Analytics was laid down in *Illuminating the Path*. So we expect that a conceptual turning point should be detectable in the literature. To test this hypothesis, we conducted a citation analysis. A list of the top cited documents and authors are shown in Table 2.6.

In Table 2.6, Jim Thomas is the highest cited author with 82 citations according to these 2,125 publications, and *Illuminating the Path* is the most cited publication at 56 times. It confirms that *Illuminating the Path* is indeed a driving force of the emerging domain.

2.4.4 Mapping the Visual Analytics Domain

As a tool to find critical points in the development of a field or a domain, especially intellectual turning points and pivotal points, CiteSpace was used in this study to conduct a Document Co-citation Analysis (DCA) and Author Co-citation Analysis (ACA), and thus the role of Jim Thomas and his publication can be further clarified.

Table 2.5 Top source journals and hot keywords

Rank	1991–2011		1991–1999		2000–2004		2005–2011	
	Journal	Freq	Journal	Freq	Journal	Freq	Journal	Freq
1	Journal of Nuclear Medicine	71	Journal of Nuclear Medicine	31	Journal of Nuclear Medicine	17	IEEE Transactions on Visualization and Computer Graphics	69
2	IEEE Transactions on Visualization and Computer Graphics	70	Journal of The American College of Cardiology	12	Neuroimage	9	Information Visualization	38
3	Information Visualization	38	European Journal of Nuclear Medicine	10	European Journal of Nuclear Medicine and Molecular Imaging	9	Computer Graphics Forum	30
4	Nuclear Medicine Communications	35	International Journal of Remote Sensing	10	European Journal of Nuclear Medicine	9	Journal of Nuclear Medicine	23
5	Computer Graphics Forum	30	Nuclear Medicine Communications	8	Nuclear Medicine Communications	9	Nuclear Medicine Communications	18
6	European Journal of Nuclear Medicine and Molecular Imaging	24	Radiology	7	Journal of Nuclear Cardiology	7	IEEE Computer Graphics and Applications	17
7	Neuroimage	19	Epilepsia	6	Neurology	5	European Journal of Nuclear Medicine and Molecular Imaging	15
8	IEEE Computer Graphics and Applications	19	Neurology	6	Clinical Nuclear Medicine	5	BMC Bioinformatics	12
9	European Journal of Nuclear Medicine	19	Journal of Nuclear Cardiology	5	Optometry and Vision Science	5	Computers & Graphics-UK	11

Table 2.5 (*Continued*)

Rank	1991–2011 Journal	Freq	1991–1999 Journal	Freq	2000–2004 Journal	Freq	2005–2011 Journal	Freq
10	Journal of The American College of Cardiology	17	Electroencephalography and Clinical Neurophysiology	5	Radiology	4	International Journal of Geographical Information Science	11

Rank	1991–2011 Keyword	Freq	1991–1999 Keyword	Freq	2000–2004 Keyword	Freq	2005–2011 Keyword	Freq
1	Visual Analytics	71	Epilepsy	16	Visual Analysis	16	Visual Analytics	71
2	Visual Analysis	66	Visual Analysis	15	Magnetic Resonance Imaging	8	Visual Analysis	35
3	Pet	42	Coronary Artery Disease	14	Myocardial Perfusion	8	Information Visualization	30
4	Information Visualization	31	PET	12	PET	8	Visualization	22
5	Positron Emission Tomography	30	SPECT	11	Positron Emission Tomography	8	PET	22
6	Epilepsy	30	Magnetic Resonance Imaging	10	Spect	7	Positron Emission Tomography	16
7	Visualization	26	Positron Emission Tomography	6	Alzheimer's Disease	6	Eeg	15
8	Spect	26	Temporal Lobe Epilepsy	6	Diagnosis	6	Fmri	12
9	Magnetic Resonance Imaging	25	Hippocampus	6	Tc-99m-Tetrofosmin	6	Breast Cancer	12
10	Coronary Artery Disease	25	Eeg	5	Coronary Artery Disease	6	Fdg	11

Table 2.6 Highly cited documents and highly cited authors

	1991–2011		1991–1999		2000–2004		2005–2011	
	Citation publication	Freq	Citation publication	Freq	Citation publication	Freq	Citation publication	Freq
1	THOMAS JJ, 2005, ILLUMINATING PATH RE	56	DEPROSPERO A, 1979, J APPL BEHAV ANAL, V12, P573	11	BONDA E, 1996, J NEUROSCI, V16, P3737	8	THOMAS JJ, 2005, ILLUMINATING PATH RE	56
2	BLAND JM, 1986, LANCET, V1, P307	41	DEPASQUALE EE, 1988, CIRCULATION, V77, P316	10	HACHAMOVITCH R, 1996, CIRCULATION, V93, P905	8	BLAND JM, 1986, LANCET, V1, P307	28
3	KAZDIN AE, 1982, SINGLE CASE RES DESI	24	KAZDIN AE, 1982, SINGLE CASE RES DESI	10	RECHTSCHAFFEN A, 1968, MANUAL STANDARDIZED	8	SHNEIDERMAN B, 1996, P IEEE S VIS LANG, P336	19
4	MATYAS TA, 1990, J APPL BEHAV ANAL, V23, P341	22	DIAMOND GA, 1979, NEW ENGL J MED, V300, P1350	9	TALAIRACH J, 1988, COPLANAR STEREOTAXIC	8	CARD SK, 1999, READINGS INFORM VISU	17
5	DEPROSPERO A, 1979, J APPL BEHAV ANAL, V12, P573	21	JACKSON GD, 1990, NEUROLOGY, V40, P1869	9	DECETY J, 1997, BRAIN, V120, P1763	7	INSELBERG A, 1985, VISUAL COMPUT, V1, P69	16
6	RECHTSCHAFFEN A, 1968, MANUAL STANDARDIZED	21	JONES RR, 1978, J APPL BEHAV ANAL, V11, P277	9	FOLSTEIN MF, 1975, J PSYCHIAT RES, V12, P189	7	KEIM DA, 2002, IEEE T VIS COMPUT GR, V8, P1	16
7	CERQUEIRA MD, 2002, CIRCULATION, V105, P539	20	CASCINO GD, 1991, ANN NEUROL, V30, P31	8	FRISTON KJ, 1995, HUMAN BRAIN MAPPING, V2, P189	7	CERQUEIRA MD, 2002, CIRCULATION, V105, P539	15
8	OLDFIELD RC, 1971, NEUROPSYCHOLOGIA, V9, P97	20	JACK CR, 1990, RADIOLOGY, V175, P423	8	HACHAMOVITCH R, 1998, CIRCULATION, V97, P535	7	ANDRIENKO G, 2007, INT J GEOGR INF SCI, V21, P839	14

Table 2.6 (*Continued*)

#	1991–2011		1991–1999		2000–2004		2005–2011	
	Citation publication	Freq	Citation publication	Freq	Citation publication	Freq	Citation publication	Freq
9	SHNEIDERMAN B, 1996, P IEEE S VIS LANG, P336	19	JACK CR, 1992, ANN NEUROL, V31, P138	8	BERMAN DS, 1993, J AM COLL CARDIOL, V22, P1455	6	HAVRE S, 2002, IEEE T VIS COMPUT GR, V8, P9	12
10	TALAIRACH J, 1988, COPLANAR STEREOTAXIC	19	MATYAS TA, 1990, J APPL BEHAV ANAL, V23, P341	8	BLAND JM, 1986, LANCET, V1, P307	6	KANWISHER N, 1997, J NEUROSCI, V17, P4302	12
	Citation Author	Freq	Citation Author	Freq	Citation Author	Freq	Citation Author	Freq
1	THOMAS JJ	82	JACK CR	33	FRIED PA	34	THOMAS JJ	82
2	KEIM DA	73	ENGEL J	32	RIZZOLATTI G	26	KEIM DA	72
3	SHNEIDERMAN B	67	DOLLFUS A	30	FRISTON KJ	24	SHNEIDERMAN B	67
4	RIZZOLATTI G	65	JACKSON GD	26	BERMAN DS	24	PARKER RI	46
5	BERMAN DS	56	OTTENBACHER KJ	22	GOLDENBERG G	21	ANDRIENKO G	45
6	FRISTON KJ	55	CENDES F	22	GERMANO G	19	MACEACHREN AM	39
7	ENGEL J	54	GENTILUCCI M	22	HACHAMOVITCH R	18	HEER J	38
8	COHEN J	49	WAHL RL	21	DECETY J	17	PELPHREY KA	37
9	MACEACHREN AM	49	THEODORE WH	20	CENDES F	16	TUFTE ER	36
10	JACK CR	48	BERMAN DS	19	TALAIRACH J	16	SEO J	36

Fig. 2.10 A co-citation network showing a broader context of visual analytics, including depicting nuclear medicine imaging (*upper half*) and visual analytics (*lower half*). Data source: D3

Figure 2.10 shows a co-citation network created by CiteSpace. In this network, the top 1 % of citations for each slice (each slice equal 5 years) were chosen as the nodes. The size of the node means the number of cited times of the responding publications. The nodes are linked with each other if they were co-cited. Cluster analysis was then conducted and the label for each cluster was computationally generated by CiteSpace using an algorithm based on likelihood-ratio tests (LLR).

In Fig. 2.10, it became obvious that two separated parts emerged. The upper half corresponds to research in nuclear medicine imaging from 1991 (in blue) to 2010 (in yellow). The lower half reveals the structure of visual analytics, which is what *Illuminating the Path* promoted. Surprisingly, the two parts were still connected and it is interesting to see which ones formed the unusual bridge that connects the two distinct conceptual sub-domains. The weak connections were due to two books, *Algorithms for clustering* on the visual analytics side and *Pattern classification* on the nuclear medicine imaging side. With hindsight, this path of connection makes sense because clustering and pattern classification are essential techniques for both areas.

The role of *Illuminating the Path* can also be seen directly from the network visualization. In Fig. 2.10, the node representing *Illuminating the Path* is the largest one in size (lower left) because, as mentioned above, it is the most cited work. The betweenness centrality of each node was also calculated based on this network structure, which measures the importance in connecting the entire network together. The centrality of this node is 0.01 because it is obviously overshadowed by the two books that provided even longer-range bridges between the upper-lower boundaries.
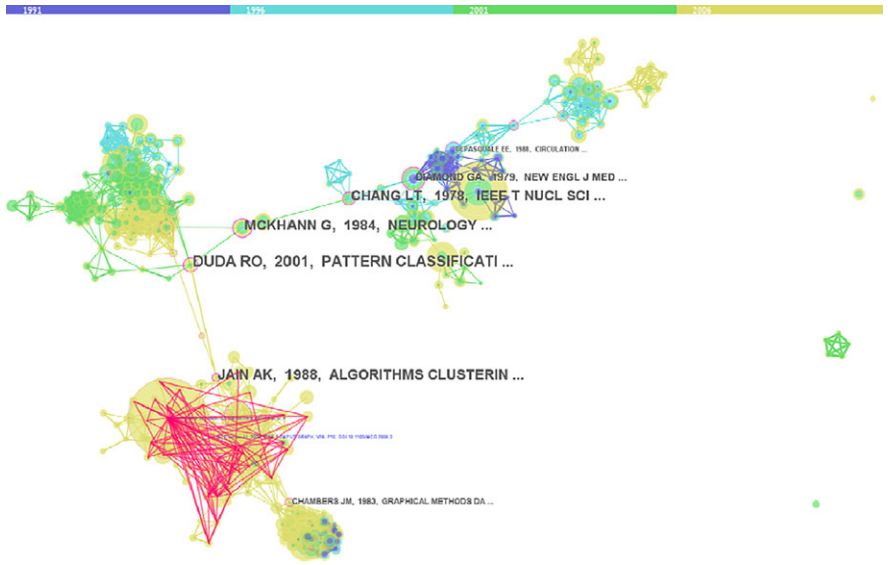
Fig. 2.11 The network overlay of the impact of *Illuminating the Path* (*red lines*) over the network of visual analysis

2.4.5 An Overlay of Network D2 in Network D3

The network shown in Fig. 2.7 was generated from the dataset D2, which consists of 135 articles that cited *Illuminating the Path*. The network shown in Fig. 2.10 was generated from dataset D3, which represents a broader context defined by over 2,200 articles relevant to visual analysis and visual analytics.

In Fig. 2.11, we superimposed the D2 network over the D3 network to highlight the network structure of citers to *Illuminating the Path* in contrast to the broader context of visual analytics. The red lines represent co-citation structures derived from the citers of Illuminating the Path. As shown in Fig. 2.11, the red sub-structure covers much of the lower half of the broader context network. The overlay network also reinforces the key position of *Illuminating the Path* in the visual analytics network.

The observed pattern from the overlay is consistent with our earlier discussion of how the domain is structured. There are two dominating areas in the domain about visual analysis. One is the domain of nuclear medicine imaging. We can see key papers in the 1970s and it evidently continues to evolve. The other is the field of visual analytics, founded by Jim Thomas and his colleagues since 2005. The overlay in Fig. 2.11 shows that currently the impact of *Illuminating the Path* is limited to the lower half of the network. Is it conceivable that the two parts will integrate more tightly in the future? Are there things that both sides could learn from each other?

2.5 Conclusion

The groundbreaking contributions made by Jim Thomas and their profound impacts
on the development of visual analytics, information visualization, and scientific vi-
sualization, are evident in scholarly publications. *Illuminating the Path* has been the
most intriguing source of inspiration that has attracted researchers and practition-
ers from more and more countries to devote their talents to solving challenging and
real-world problems through visual analytics.

References

Chen, C. (2004). Searching for intellectual turning points: progressive knowledge domain visu-
    alization. *Proceedings of the National Academy of Sciences of the United States of America*,
    101(suppl), 5303–5310.
Chen, C. (2006). CiteSpace II: Detecting and visualizing emerging trends and transient patterns in
    scientific literature. *Journal of the American Society for Information Science and Technology*,
    57(3), 359–377.
Chen, C., et al. (2009). Towards an explanatory and computational theory of scientific discovery.
    Journal of Informetrics, *3*(3), 191–209.
Chen, C., Ibekwe-SanJuan, F., & Hou, J. (2010). The structure and dynamics of co-citation clusters:
    a multiple-perspective co-citation analysis. *Journal of the American Society for Information
    Science and Technology*, *61*(7), 1386–1409.
Christopher Columbus Fellowship Foundation (2009). Homeland Security Award Columbus
    Scholar 2009. http://www.columbusfdn.org/homelandsecurity/2009.php.
Ebert, D. S., Dill, J., & Kasik, D. J. (2010). In memoriam: illuminating our paths—James (Jim)
    Joseph Thomas. In *IEEE symposium on visual analytics science and technology (VAST)* (pp.
    xiv–xiv).
Kasik, D., et al. (2010). Jim Thomas, 1946–2010. In memoriam. *IEEE Computer Graphics and
    Applications*, *30*(6), 10–13.
Keim, D., et al. (2008). *Visual analytics: scope and challenges* (pp. 76–90). Visual Data Mining.
Thomas, J. J., & Cook, K. (2005). *Illuminating the path: the R&D agenda for visual analytics*.
    IEEE.
Wong, P. C. (2010). The four roads less traveled—a tribute to Jim Thomas (1946–2010).
    http://vgtc.org/JimThomas.html.
Wong, P., & Thomas, J. (2004). Visual analytics. *IEEE Computer Graphics and Applications*,
    24(5), 20–21.