World Academy of Science, Engineering and Technology
International Journal of Information and Communication Engineering
Vol:18, No:5, 2024

Exploring the Landscape of Information Visualization through a Mark Lombardi Lens

Alon Friedman, Antonio Sánchez Chinchón

Abstract—This bibliometric study takes an artistic and storytelling approach to explore the term ”Information visualization.” Analyzing over 1008 titles collected from databases that specialize in data visualization research, we examine the titles of these publications to report on the characteristics and development trends in the field. Employing a qualitative methodology, we delve into the titles of these publications, extracting leading terms and exploring the co-occurrence of these terms to gain deeper insights. By systematically analyzing the leading terms and their relationships within the titles, we shed light on the prevailing themes that shape the landscape of ”Information visualization” by employing the artist Mark Lombardi's techniques to visualize our findings. By doing so, this study provides valuable insights into bibliometrics visualization while also opening new avenues for leveraging art and storytelling to enhance data representation.

Keywords—Bibliometrics analysis, Mark Lombardi design, information visualization, qualitative methodology.

I. Introduction

DIGITAL storytelling, which emerged in the 1990s with the rise of the Internet, encompasses various forms of visualization to enhance the communication of information. Economou [1] identifies three types of visualization relevant to digital storytelling: data visualization, information visualization, and infographics. While bibliometric methods have traditionally been used in scientific research to evaluate the impact of researchers, papers, or topics within specific fields, recent advancements have leveraged natural language processing to analyze bibliographic data and create new visualizations [2]. Initially, bibliometric analysis focused primarily on scientific content and relied on manual methods to identify key topics. However, it has evolved to encompass diverse research domains and reveal interconnected terminologies. Previous research has extensively explored the visual aspects of bibliometrics using advanced computational techniques and visualization methods, but the artistic perspective remains largely unaddressed. Notably, a study by Kim et al. [3] compares the terms ”information visualization” and ”data visualization” in terms of their primary topics, finding cross-referencing terminologies within each field, but without delving into the artistic perspective associated with these topics. By examining the leading topics in scientific articles in the field of ”information visualization,” we can gain valuable insights that align with the benefits of visualizing the analysis results through an artistic lens and storytelling approach [4].

Mark Lombardi, renowned for his conceptual art and distinctive diagrammatic hand drawings, has gained international recognition. His work has been exhibited in esteemed museums worldwide, including the Museum of Modern Art, the Whitney Museum, and The Reina Sofia Museum. Lombardi referred to his diagrams as 'narrative structures,' endowing them with dual significance: depicting fraudulent activities and serving as a form of visual storytelling, as described by Hobbs [5]. Hobbs further elaborates on Lombardi's creative process, which involved meticulous research, condensing findings onto index cards, and refining multiple sketches before finalizing the artwork [5]. Following Lombardi's death, new researchers begun expanding upon his designs using computational programming languages. However, their work on Lombardi has not yet been examined through bibliometric analysis to uncover the narrative behind the results.

This study raises the following questions: A) Can Lombardi-inspired designs be generated based on bibliometric analysis? B) What insights do the results of Lombardi's designs offer about the term ”information visualization”?

A. Friedman is an Associate Professor with the University of South Florida, Tampa, FL, USA (e-mail: alonfriedman@usf.edu).

Antonio Sánchez is a Data Scientist with Telefónica (e-mail: aschinchon@gmail.com).

II. Background

The term “information visualization” encompasses design, development, and computer applications that produce interactive visual representations using generative algorithms. Chen [6] emphasized its components, dividing it into “information” and “visualization,” revealing its potential for extracting insights from data and enhancing cognitive abilities. Liu et al. [7] discussed the multidisciplinary nature of information visualization, addressing challenges and advancements across various disciplines. Munzner [8] focused on the process of writing information visualization manuscripts, describing different stages and decision-making processes.

However, empirical research exploring the composition of titles in bibliometrics related to ”information visualization” through an artistic design and qualitative research approach is lacking. This study aims to fill this gap by examining title compositions in bibliometric analysis using artistic design and qualitative research methods, thus providing valuable insights into the field of information visualization.

III. Bibliometric Data Analysis

Bibliometrics is the statistical analysis of data about the publication and citation of works. It is commonly used to measure the impact of scientific research outputs by counting how many times research publications are cited [9]. Traditionally, bibliometrics studies have focused on measuring scientific citations found in academic journals within a discipline to examine characteristics such as gender, institutional affiliation, productivity ranking, and format. This

approach is appropriate for examining how scientific disciplines develop through the productivity of individual researchers. However, it raises the question of how to measure the impact of titles as part of the scientific identification of a work. To address this, researchers utilize citation analysis, where they analyze the most frequent words in the titles of articles published within a specific field [10]. As part of their study on visualization, Isenberg et al. [11] examine the keywords that researchers use in their research papers.

A. Visualization Analysis of Bibliometric Data

In recent years, the visualization of bibliometric data has gained significant attention due to increased data availability and advancements in visualization techniques [2]. Visualization enables insights into research trends, impact, and collaborations. Bibliometric data can be visualized to track research trends over time, capturing the evolution of topics and publication patterns [2]. It can also showcase the impact of a paper by visualizing citation counts and its influence within the research community [12]. As researchers explore various visualization techniques for bibliometric data, such as network visualization [13], timeline visualization [14], and topic modeling [15], it is important to consider their strengths and weaknesses based on specific requirements. However, there is a lack of research methodology that combines the data artist lens with these techniques.

IV. The Artist: Mark Lombardi

The artist Mark Lombardi is known for his conceptual artwork, where he used diagrammatic hand drawings. He referred to his diagrams as "narrative structures" a term he used to convey the subject of his drawing on global deceptions orchestrated by politicians and influential figures. Lombardi approach to creating these narrative structures involved a structured three-act process: first, extensive research; followed by the synthesis of research findings onto index cards; and finally, a series of preliminary sketches leading up to the final rendition, according to Hobbs [5]. In the first stage of his design process, Lombardi conducted extensive research. He manually collected data from mass media ranging from the news agency United Press International to Spy magazine. In addition, he kept a daily diary of his findings in which he recorded precise citations. In the second stage, he transferred the data to index cards, where he summarized the content of the articles about his investigation by listing names, places, and topics. He arranged the cards in alphabetical order, reflecting the organization of a library card catalogue system. (Before moving to New York City, Lombardi worked as a reference librarian for the Fine Arts department in the Houston Public Library.) To organize his index cards, Lombardi used three colours to classify them: red, green, and yellow according to Friedman [16]. However, we did not find any documentation what these three categories represent in his work. Fig. 1 shows an early stage of Lombardi work titled George W. Bush, Harken Energy, and Jackson Stevens c. 1979-90 (2nd version) published by Hobbs R (2003) Mark Lombardi Global Networks. New York: Independent Curators. International (ICI). Upon reviewing Lombardi's work,

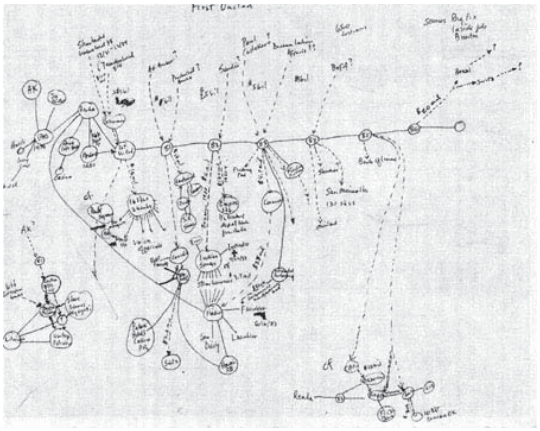

Fig. 1 Mark Lombardi: George W. Bush, Harken Energy, and Jackson Stevens c. 1979-90 (2nd version)

Fig. 2 The "road signs" in Mark Lombardi's art represent the connections among various fraudsters; Hobbs [5]

one can readily identify "nodes" and textual annotations that facilitate the viewer's comprehension of the networks depicted in his drawings. The notes in the networks introduce relevant people or institutions and are connected by lines. According to MacEachren [17], the road signs help the viewer discern levels of meaning in the map. A close examination of his work reveals that Lombardi used five graphic road sign elements to represent different types of data and their connections. These graphic elements are repeated throughout his work. Fig. 1 lists these elements.

Scholars often discuss Lombardi's work in the field of visualization. Fry [18] reported on Lombardi's design in his blog: 'I think Lombardi's work encapsulates two themes that are important for the future of design. First is that we must maintain a humanist view of data, relying on our own faculties to tell a story. Second, to improve the discourse surrounding data, we must disavow our fascination with the intricate and complicated by learning how to throw things out'.

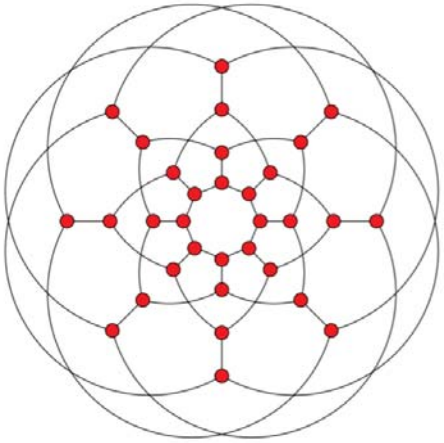

Fig. 3 Lombardi spirograph, according to Duncan et al. [19]

A. Exploring Software Developers' Adoption of Lombardi's Design Principles

Over the years, researchers have endeavoured to integrate Lombardi's design principles into their software development work, aiming to emulate his distinct style characterized by curved arcs representing edges, resulting in visually appealing and legible designs. Duncan et al. [19] defined Lombardi's drawings as computer-generated compositions that consist of circular arcs with edges connecting vertices positioned on a circle. They employed a deterministic algorithm that consistently produces the same output for a given input, ensuring predictability in the sequence of states. Their interpretation of Lombardi's design, known as Lombardi Spirograph, involves a group of symmetries based on rotations and reflections of a regular polygon. Fig. 2 is the Duncan et al. [19] work they called Lombardi's Spirograph.

Another perspective on Lombardi's design was presented by Tolksdrof [20], who meticulously documented all the relationships found in Lombardi's original drawings using scan images and photographs. This approach aimed to faithfully capture and reproduce the intricate connections present in Lombardi's works.

These various approaches exemplify the ongoing exploration and adaptation of Lombardi's design principles by software developers, as they strive to incorporate his aesthetic and conceptual elements into their own work. By building upon Lombardi's foundation, these researchers contribute to the evolution of visual representation and storytelling within the realm of software development.

Chernobelskiy et al. [21], in an analysis of Duncan et al. [19], reported: 'Their methods are deterministic, not force-directed, but, as they show, there are several types of graphs that cannot be drawn perfectly as Lombardi drawing'. A different approach to Lombardi's design was proposed by Tolksdrof [20] , who used scans and photographs of the original drawing and recorded all the relations found in Lombardi's original work. He then used GraphML files to produce the result shown in Fig. 3.

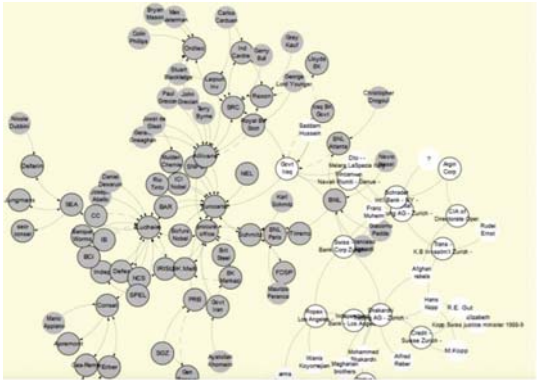

Fig. 4 Work based on images and photographs scanned from Lombardi. Source: Tolksdorf [20]

V. Methodology

Our methodology for transforming bibliometric data into Lombardi design involved six distinct steps. Firstly, we collected a comprehensive corpus of books, book chapters, journal articles, and conference proceeding titles from 1985 to 2020, utilizing resources such as IEEE Explore [22], Google Scholar [23], ACM [24], and Vispubdata [11]. The second step encompassed data preprocessing, where we ensured data consistency and accuracy through techniques like data normalization and deduplication. To focus our analysis, we filtered the entries to include only those with the term "Information Visualization" in their titles. In the third step, we applied Lombardi's methodology to categorize the data into three groups: red, green, and yellow. The red category represented the leading terms in the titles, while the green category contained terms associated with the leading terms. The yellow category explored the relationships between the top terms. To visualize the results effectively, we employed the R programming language and its associated libraries in the fourth step. R provided powerful tools for transforming and analyzing the data, enabling us to create visual representations that effectively conveyed the intricate relationships between terms and their connections. By following these six steps, we successfully transformed bibliometric data into a visually compelling Lombardi design, capturing the interconnections and thematic patterns present in the scholarly literature. Our methodology provided a systematic and rigorous approach, offering valuable insights into bibliometrics through an artistic lens.

To illustrate our examination of the green category, which contains terms associated with the leading terms, we present an example. Fig. 5 showcases the title of the work "Fluid Interaction for Information Visualization" by Niklas et al. [25]. Under this study analysis, the term "interaction" is classified as an occurrence within a 2-word distance from the term "information visualization."

VI. Results

In the initial phase of our analysis, we compiled a comprehensive dataset consisting of books, book chapters, articles, and conference titles that included the term

Fig. 5 Study identification of occurrence within a 2-word distance from the term ”information visualization” based on Niklas et al. [25]

TABLE I
STUDY SUMMARY BASED ON LOMBARDI'S DESIGN AND THE ANALYSIS OF THE TERM ”INFORMATION VISUALIZATION”

Ranking	Lead terms (red)	%	Co-occurrence matrix within 2 word distance (green)	%	The most unpopular terms (yellow)	%
1	Interface	16%	Interface	4.90%	Concept	2%
2	Aesthetics	14.50%	Design	4.80%	Discovery	1.90%
3	User experience	12%	Knowledge	3.60%	Low-level	1.20%
4	Evaluation	10%	Software development	3.10%	Empirical	1.10%
5	Graph	9%	Art	2.60%	Mental	0.40%
6	Theory	7%	Collaboration	2.10%	Casual	0.20%
7	Network	6%	Support	1.90%	Tangible	0.10%
8	Communication	5%	Cognitive	1.50%	Mondrian	0.10%
9	User characteristics	4.1%	Artifacts	1.30%	Adaptive	0.10%
10	3D	4%	Multimedia	0.2%	Perceptual	0.10%

”information visualization” in their title. Out of the more than 5000 titles we reviewed across these publication types, 1008 titles met our inclusion criteria. Upon analyzing the data, we observed variations in the average length of titles across different publication types. Specifically, book and chapter titles exhibited an average length of 4.5 words, while journal articles had an average length of 9.5 words, and conference proceedings had an average length of 8.5 words. We proceeded to categorize the collected data into three distinct groups: the red category consisted of the leading terms identified in titles, the green category encompassed terms that co-occurred with the term ”information visualization” within a 2-word distance, and the yellow category focused on exploring less frequently used terms. Among the leading terms identified throughout the data collection, ”Interface” appeared most frequently with a count of 151 out of 1,008 titles. Following closely, the term ”Aesthetics” appeared 136 times, and ”User Experience” ranked third with 110 occurrences. Additionally, in the co-occurrence analysis, the term ”Interface” emerged as the leading term. We also examined the frequency with which terms appeared together in titles to measure their co-occurrence. Table I summarizes the key results derived from our analysis.

Next, we conducted our data analysis and designed the visualization based on the criteria proposed by Duncan et al. [19]. This approach involved representing edges as circular arcs and placing vertices on a circle, with a creative arrangement of both edges and vertices. However, during our data analysis, we observed that the connection between the arcs and edges was not clearly defined. To address this issue and establish a more coherent connection between the arcs and edges, we

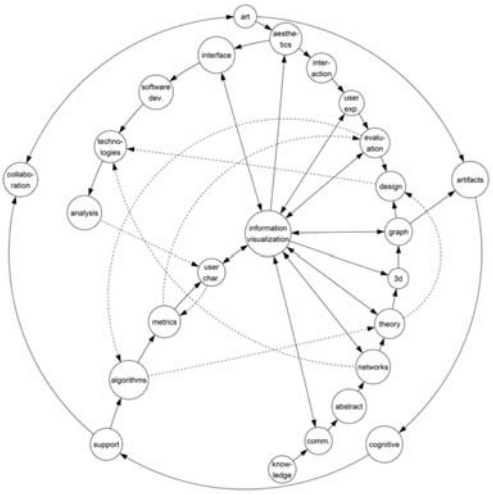

Fig. 6 Study visual summary based on Lombardi's design

employed the scales option to control the appearance of the data on the graph. By adjusting the size and shape of the lines, we were able to improve the visual representation. However, due to the limitations of this scaling approach, we were unable to extend the visual display beyond ten words and capture the entire collection of words. To generate this graph, we developed an algorithm to scale the location of the nodes based on their coordinates in the graph generated by this study. It's important to note that this algorithm is not based on a statistical foundation.

Fig. 5 is a graphic representation of the result of our conversion to bibliometric analysis and visualization based on Mark Lombardi. The figure showcases our efforts to capture significant top ranking terms and emphasize the co-occurrence matrix within a 2-word distance, as depicted in Fig. 6.

For the code of this work, please refer to the following GitHub repository: https://github.com/aschinchon/lombardi.

VII. CONCLUSION

This study explored the application of art visualization in analyzing bibliometric data. While there are various visualization techniques available for bibliometric analysis, we discovered a gap in the literature regarding the utilization of art visualization methods. By following the artistic approach of Mark Lombardi, not only in design but also in methodology, we aimed to bridge this gap and offer a unique perspective. The findings of our study have practical implications for researchers in the field of information visualization, particularly when it comes to selecting appropriate titles for their work. Through the identification of key terms ”information visualization” in the titles, we contribute to a better understanding of the commonly used language and terminology in this domain. This knowledge can aid researchers in crafting titles that accurately represent the content and attract relevant audiences. Future research should further explore the integration of visual art and data visualization techniques to gain more insights into bibliometric data. By combining artistic elements

with traditional analytical approaches, we can unlock new perspectives and uncover hidden patterns within the data. This approach has the potential to enrich the field of bibliometrics and enhance our understanding of scholarly communication. Using visual arts as a lens, we hope researchers will look at innovative ways to represent and interpret bibliometric data.

ACKNOWLEDGMENT

We would like to thank the University of South Florida for allowing us to continue studying Mark Lombardi's work in data visualization and storytelling. In addition, we would like to thank Joe Amrhein, from the PIEROGI gallery in NY, for allowing us to use Lombardi's drawing.

REFERENCES

[1] D. Economou and G. Chatzidamianos, *Research and Advanced Technology for Digital Libraries*, 1st ed. Springer, 2015.
[2] K. Börner, C. Chen, and K. W. Boyack, "Visualizing Knowledge Domains," *Annual Review of Information Science and Technology*, vol. 37, no. 1, pp. 179–255, 2003, to appear.
[3] M. C. Kim, Y. Zhu, and C. Chen, "How Are They Different? A Quantitative Domain Comparison of Information Visualization and Data Visualization (2000–2014)," *Scientometrics*, vol. 107, pp. 123–165, 2016.
[4] N. Gershon and W. Page, "What Storytelling Can Do for Information Visualization?" *Communications of the ACM*, vol. 44, no. 8, pp. 31–37, 2001.
[5] R. Hobbs, *Mark Lombardi Global Networks*, 1st ed. New York: Independent Curators International (ICI), 2003.
[6] C. Chen, "Science Mapping: a Systematic Review of the Literature," *Journal of Data and Information Science*, vol. 2, no. 2, pp. 1–49, 2017.
[7] S. Liu, W. Cui, Y. Wu, and M. Li, "A Survey on Information Visualization: Recent Advances and Challenges," *The Visual Computer*, vol. 30, pp. 1373–1393, 2014.
[8] T. Munzner, "Process and Pitfalls in Writing Information Visualization Research Papers," *Information Visualization: Human-Centered Issues and Perspectives*, pp. 134–153, 2008.
[9] V. Durieux and A. G. Pierre, "Bibliometric Indicators: Quality Measurements of Scientific publication," *Radiology*, vol. 255, no. 2, pp. 342–351, 2010.
[10] J. Ravenscroft, M. Liakata, A. Clare, and D. Duma, "Measuring Scientific Impact Beyond Academia: An Assessment of Existing Impact Metrics and Proposed Improvements," *PloS one*, vol. 12, no. 3, p. e0173152, 2017.
[11] P. Isenberg, T. Isenberg, M. Sedlmair, J. Chen, and T. Möller, "Visualization as Seen Through its Research Paper Keywords," *IEEE Transactions on Visualization and Computer Graphics*, vol. 23, no. 1, pp. 771–780, 2017.
[12] R. Smith, "Mark Lombardi, 48, an Artist Who was Inspired by Scandals," *The New York Times*, vol. 25, 2000.
[13] L. Barbu, D. Mihaiu, R. Şerban, A. Opreana, and A.M., "Knowledge Mapping of Optimal Taxation Studies: a Bibliometric Analysis and Network Visualization," *Sustainability*, vol. 14, no. 2, 2022, to appear.
[14] N. J. Van Eck and L. Waltman, "Visualizing Bibliometric Networks," *Measuring scholarly impact: Methods and practice*, pp. 285–320, 2014.
[15] A. Chaney and D. Blei, "Visualizing Topic Models," in *Proceedings of the International AAAI Conference on Web and Social Media*. AAAI Association for the Advancement of Artificial Intelligence, 2012, pp. 419–422.
[16] A. Friedman, "Mark Lombardi's Visualization Discovery," in *Proceedings of the ADS-VIS 2011: Making visible the invisible: art, design and science in data visualisation*, 2011, pp. 12–16.
[17] A. M. MacEachren, *How Maps Work: Representation, Visualization, and Design*, 1st ed. New York: Guilford Press, 2004.
[18] B. Fry, "Learning from Lombardi's Blog Posting," www.benfry.com/exd09/, 2014.
[19] C. A. Duncan, D. Eppstein, M. T. Goodrich, S. G. Kobourov, and M. Nöllenburg, "Lombardi Drawings of Graphs," *Journal of Graph Algorithms and Applications*, vol. 16, no. 1, pp. 37–83, 2010.
[20] R. Tolksdorf, "Initial Report on Digitally Researching the Network Drawings of Mark Lombardi," Freie Universität Berlin, Tech. Rep. TR-B-13-03, 2013.
[21] R. Chernobelskiy, K. I. Cunningham, M. T. Goodrich, S. G. Kobourov, and L. Trott, "Force-Directed Lombardi-Style Graph Drawing," in *Graph Drawing: 19th International Symposium, GD 2011, Eindhoven*. Springer, 2012, pp. 320–331.
[22] "IEEE Xplore," https://ieeexplore.ieee.org/Xplore/home.jsp, Accessed 2000, accessed: May 24, 2024.
[23] "Google Scholar," https://scholar.google.com/, Accessed 1966, accessed: May 24, 2024.
[24] ACM DL, "Association for Computing Machinery (ACM) Digital Library," https://dl.acm.org/, 2005, accessed: May 24, 2024.
[25] N. Elmqvist, A. V. Moere, H.-C. Jetter, D. Cernea, H. Reiterer, and T. Jankun-Kelly, "Fluid Interaction for Information Visualization," *Information Visualization*, vol. 10, no. 4, pp. 327–340, 2011.

Alon Friedman Alon Friedman currently works at the School of Information Science, University of South Florida. Alon is an interdisciplinary researcher with a focus on visualization theory and the visualization of data. He uses different theoretical perspectives to organize data through a visualization lens.

Antonio Sánchez Chinchón Antonio Sánchez Chinchón is data scientist at Telefónica, the leading telecommunications company in Spain. He published his work on his blog called Fronkonstin at https://fronkonstin.com/about/.