Indian Journal of Applied Economics and Finance
Vol. 1 No. 1 (January 2025): pp: 96-106

Freeware Tools for Visualization of Collaboration and Citation Network in Scientometrics Research

Samyak Pratik Pattanaik[1] & Bikram Kishore Beura[2]

Abstract

Visualization in Scientometrics research have become very much necessary for the Interpretation and Presentation of complex patterns in large bibliographic datasets. Due to the increase in digital information and improvement in research output, visualization tools have become necessary in the analysis, interpretation, and presentation of complex data efficiently by the researcher. This paper collected and compiled selected freely accessible Visualization tools preferred by many academic researchers in the area of Scientometricss study. A detailed analysis of the features, capabilities and functionality of each tool have been carried out to identify the best suitable tool for Visualization of Collaboration Networks and Citation Patterns. Each is described with respect to its features, capabilities, and applications in research. It discusses advantages accrued from the use of such tools in improving data interpretation, pattern identification. It has been found that Vosviewer and CiteSpace tools are most appropriate freeware tools which can help the researchers to carry out advanced Scientometrics study and decision-making. Rest of the tools like SciMAT, Biblioshiny and others can be used Optimizely according to the researcher's requirement as all these tools have both pros and cons in Visualising the Scientometrics Study. The paper aims at helping researchers to select appropriate tools for specific purposes and will contribute to the advancement of scientometrics analysis.

Keywords: Scientometrics, Visualization Tools, Freeware, Collaboration Networks, Citation Networks

[1] Ph.D. Research Scholar (JRF), Dept. of Library & Information Science, Sambalpur University
[2] Assistant Professor, Dept. of Library & Information Science, Sambalpur University

Indian Journal of Applied Economics and Finance
Vol. 1 No. 1 (January 2025): pp: 96-106

Introduction

Scientometrics, have emerged as key approaches to understand the structure, dynamics, and impact of scientific research. Bibliometrics is understood to be the quantitative analysis of written publications in the form of books, articles, and other types of academic output. Its main focus lies in the measurement of pattern publications with the intent of assessing research impact, output, and dissemination across and within disciplines. Due to the fast pace of academic publication, there is now even more interest in bibliometric and scientometric analyses, especially when it comes to figuring out how citation and research cooperation affect academic impact (Bornmann & Leydesdorff, 2014).Common bibliometric indicators, such as the count of citations, the h-index, and impact factors, serve as proxies to help quantify the quality and impact of research outputs. Scientometrics, by contrast, is the development of bibliometrics beyond mere publication and citation counting. It involves an array of indicators and techniques that focus on the scientific enterprise in its entirety. Scientometric analysis also involves understanding the dynamics of research, such as the development of scientific fields, collaboration networks, funding effects, and knowledge dissemination. Scientometrics provides a wide-ranging perspective on the research landscape and its implications for society.

Recently, these visualization tools have become very powerful tools within the realm of b scientometrics, resolutely changing how researchers investigate and make meaning from data.Visualization tools are necessary to the analysis of these networks to find patterns, clusters, and important nodes in complicated datasets (Van Eck & Waltman, 2010). These are the tools that convert intricate datasets into visual representations such as maps, graphs, and networks to discern patterns, trends, and relationships. Specific examples include network visualizations of citations and co-authorship patterns, made possible by tools like VOSviewer and Gephi, which disclose the underlying structure and relationships in the scientific literature. In turn, CiteSpace and Pajek are software programs that help identify emerging research fronts and contributing authors in a given field. This has definitely enhanced the ability of researchers using bibliometrics and scientometrics to analyze large sets of data, gain new insights, and present findings in an effective manner. With the ever-growing number of research publications, these tools are now increasingly becoming essential in navigating the forever-expanding universe of scientific research.

Indian Journal of Applied Economics and Finance
Vol. 1 No. 1 (January 2025): pp: 96-106

Need and Advantages of Visualization Tools for Bibliometrics and Scientometrics

Large-Scale Data Management: With increasing research output across the globe, several millions of scholarly articles and citations result annually. Visualization tools play an important role in processing such large data sets and helping researchers to settle and analyze data in a much better way.

Revealing Hidden Patterns: While traditional statistical methods may overlook complex relationships in data, such insights become feasible through visualization tools, which reveal, among other aspects, citation networks, connections between co-authors, and nascent research trends that underpin the structure and evolution of scientific knowledge.

Simplify Data Interpretation: Large volumes of complicated, multivariate data are difficult and mostly forbidding to interpret. Visualization tools translate such datasets into intuitively understandable graphs, maps, and network diagrams that allow researchers to understand and analyze such complex information with ease.

Enhancing Analytical Depth: This is further deepened by the fact that visualization tools let the researcher interactively explore his data, create dynamic views, and examine different aspects of research, such as collaboration networks, research impact, and knowledge dissemination. This depth of analysis provides valuable insights for the identification of influential papers, authors, and institutions.

Facilitating Effective Communication: It thereby enhances the communication of research findings, making complex data more accessible to different classes of audiences such as academics, policymakers, and the general public. Visualization aids in portraying insights clearly and allows for more effective communication to support better comprehension and engagement.

Supporting Strategic Decision-Making: Visualization tools support strategic decision-making by underlining the relevant trends and metrics. They provide very important inputs for

researchers, institutions, and funding agencies for strategizing research, performance evaluation, and identification of areas to be developed in the future.

Visualisation tools

Visualization Tools	Year	Developer	Operative System	User interface	Visualisation
Bibexcel	2017	University of Umeå (Sweden)	Win	Desktop	External software
Biblioshiny	2019	University of Naples Federico II (Italy)	Runs in R	Web	Network, three-fields plot, word cloud, tree map, historiograph, strategic diagram, evolution map and world map
BiblioMaps	2018	University of Lyon (France)	Runs in Python	Web	Network
CiteSpace	2019	Drexel University (USA)	Win	Desktop	Tree ring, geospatial map
CitNet Explorer	2014	Leiden University (The Netherlands)	Win, OSX, Linux	Desktop	Network
SciMAT	2016	University of Granada (Spain)	Win, OSX, Linux	Desktop	Strategic diagram, cluster network, overlapping map, evolution map
Sci^2 Tool	2018	Cyberinfrastructure for Network Science Center (USA)	Win, OSX, Linux	Desktop	Temporal, geospatial map, topical, network
VOSviewer	2019	Leiden University (The Netherlands)	Win, OSX, Linux	Desktop	Network, overlay, density

Bibexcel

It is designed to create data files that can be imported into Excel or any other application which may process tab-delimited data files for these incoming data sources. Olle Persson was leading this process of tool development. This is a free program comprising several tools in one, while only some of them show up in the window, others are embedded into the menus. Where this flexibility is a feature of the software, it is for that reason not easy to learn for the beginner. Data can be imported into BibExcel from several bibliographic sources, including WoS and

Scopus, and export formats for ProCite. Even so, if the user is familiar with the standard Bibexcel various types of documents can be generated in advance of analysis relevant to the file structure. Taking advantage of this software's properties allows the possibility of creating different bibliometric networks. The two basic ones are bibliographic coupling and co-word, co-citation.and Co-author. based on any document area or a mix of some., one may produce various co-occurrence matrices. This normalizes the data so that the user may export the file and use some other utility to produce a matrix to perform MDS or some clustering operation on. Alternatively, it prepares the text data for entry through the incorporation of various pre-processing functions, which include an English word stemmer, document deduplication, text transformation, author preservation, comma-delimited address conversion, and many more. Although it does not have the feature to visualize, it exports into many formats that can be visualized with other software such as Pajek, and Ucinet. Depending on the characteristics of the unit of analysis, SPSS or VOS viewer will be used for the type of visualization.

Biblioshiny

It is developed in the R programming language utilizing Bibliometrix, where it may be integrated with other programs in R. It was created by Massimo Aria and Corrado Cuccurullo, from Luigi Vanvitelli at the University of Campania and the University of Naples in Italy. The data is almost entirely from Dimensions, WoS, and Scopus. The user interface is designed systematically and easily navigated. The developers used an SMA process to classify the main menu, including additional statistics and various graphs. Three levels of metrics were included: conceptual, intellectual, and social; source, author, and document; along with three architectural descriptions of knowledge may be included. The analytical and graphic representations mentioned previously demonstrate good variety analysis capability, divided into seven categories: 1) General overview 2) Authors 3) Documents 4) Sources 5) Intellectual structure 6) Social structure 7) Conceptual structures. The visual representation of the performance analysis can be exported to various formats such as Excel, PDF, and Pajek with HTML. The visual representation of the performance analysis can also be copied to clipboards or exported to Pajekwith HTML.

BiblioMaps

It isa collection of scripts for creating visualizations from bibliographic data. These scripts are in Python. The Biblio Maps web application was developed by author Sébastian Grauwin in 2016-2017. This software, which is intended for use only requires a knowledge of Biblio Maps, not Python because all of this makes it easier to deal with the issue of choosing the tools you want to work with. The user can alternatively modify this code to generate different results. The scripts perform the following functions using data from WoS and Scopus: i) data collection ii) data preparation, which consists of parsing the raw data and filtering it. iii) Data analysis, followed by identifying co-occurring patterns especially in bibliographic coupling and clusters iv) data visualization- present results or export it to another application, or visualized in BiblioMaps; and v) data report. you can export report of the data analysis results in LaTeX format. In BiblioMaps you can only see the network and other mapping types require you to use additional applications. This program has an "all_in_one" script which allows you to run all your analyses simultaneously but still gives you individual control of each step.

Citespace

The central focus of Citespace when analysing and visualizing tracks and patterns of specialty in a field of study is study therefore helps in the exploration of new knowledge trends. It was a product of the US, at Drexel University. Although this tool is not fully freely available, one can use this software on very limited basis with less dataThere is much to see in this tool to assist in understanding and interacting with the historic trends and network options, including the growth of a knowledge domain, key citations in the bodies of knowledge, The automatic labeling of various clusters determination utilizes the keywords of cited literature and network of partnerships with geospatial domains and international partnerships. CiteSpace can also merge perfectly with other databases, like Scopus, Web of Science, and Chinese Social Sciences Citation Index, plus citation-based approaches, thanks to free tools available from Dimensions and cross ref. The following points provide a brief overview of the CiteSpace workflow: Discovery of the knowledge domain, data collection, and extracted research Terms falling under the research front include time slicing, threshold selection, pruning and merging, visual review, confirmation of important points, and many more. It has many choices for analysis and visualization of the built networks. The important nodes are presented to the user in a visual representation that corresponds to their betweenness centrality.

Indian Journal of Applied Economics and Finance
Vol. 1 No. 1 (January 2025): pp: 96-106

CitNet Explorer

It is a framework applied as a program for the navigation and analysis of citation networks. The methodology of Garfield's Algorithmic Historiography was harnessed. The network building was accomplished through CitNet Explorer, including nodes for the original target studies, with the edges representing relationships of citations between two elements of the network. This network analysis, and its design, were introduced at the Leiden University Center for Science and Technology Studies (CWTS) in the Netherlands. Further, the WoS data is the only published source of information used to create the citation network elements; however, analysis is not limited to WoS data and Scopus data can be subjected to analysis concurrently. Once built, the citation network can be retrieved as a Pajek file. CitNetExplorer has two ways to approach data analysis of citation networks as described in the book Validators and Revitalizers: articles downloaded from a WoS dataset which represent one approach, and articles that were non-WoS (including the information from WoS) which represent the second approach. This method features three aspects worth noting: the selection of publications, filtering and expanding functionality, and the four techniques offered for building citation networks. CitNetExplorer features four different approach options regarding network analysis: extracting related elements, clustering publications, identifying main publications, and establishing the path from a publication to other publications. The user can additionally customize visualizations to depict both direct citation relations as well as indirect citation trails of higher order.

SciMAT

SciMAT consists of one open visualization software application. The implications, measurement requirements, methodologies, and algorithms required to undertake and produce the different analyses and visualizations are embedded in one software application. It is a comprehensive tool designed to support every stage of the SMA process. It is now edited and maintained by the IntellSOK team at the University of Cadiz (Spain). SciMAT was developed by the Secaba Laboratory at the University of Granada (Spain). It brings together all of the required modules to accomplish each step in the SMA approach. It provides some essential assistance to the analyst at each step of the process, from importing the saved data, all the way through to visualizing and interpreting the results. The science maps and other generated

images can be enhanced with scientometric measures like the sum of citations, maximum and minimum of citations, and average of citations, all of which are based on citation-related index measures. It utilizes advanced bibliometric indices: G-index and h-index, and into the advanced bibliometric, aka bibliometric maps, measures an HG index and q2. The user can add the saved files of the data in RIS and WoS formats. A preprocessing module is offered where the analyst can conduct necessary tasks like data reduction (Reducing the number of articles contributed to the repository or analyzed), network reduction (restricting the study to a selection of related articles), time-sliced (time-based analysis), and deduplication are necessary functions. The second step of SciMAT is a high-level description of the analytic workflow: which is organized into 4 main phases: the first includes building the dataset; the second phase is to normalize the network; the third step is to employ a clustering method and produce the map to create any related clusters or sub-network that is presented on the map; the last four steps are the engagement of the network, performance, and longitudinal studies. Ultimately, SciMAT provides multiple ways to visualize the data in the forms of overlapping maps, cluster networks, strategic diagrams, and evolution maps. As an interesting output feature of the visualization module, the user may create reports either in HTML or LaTeX format.

Sci2 Tool

This is a modular toolbox that has been designed specifically to assist science research, to support the temporal, spatial, topical and network analysis as well as the representation of information, either at the micro or meso-levels with local or macro scope. Created by Indiana University USA's Cyberinfrastructure for Network Science Center, the program can read several types of bibliographic data from a variety of other sources, including EndNote export data and WoS, Scopus, GS, and Bitext (a peer-reviewed journal concerning higher education). The program will also process bibliographic data from grant funding sources from national institutes, such as NIH and NSF. Within the broader field of the academic knowledge-seeking domain, it can use data from social networking sites, such as Facebook, along with biographic records. It is important to state that the tool has affordances from knowledge sources, areas including modeling, visualizing, processing, and collecting, as well as analysis of and on data. It can enable analysis of a variety of network types and can conduct temporal, spatial, textual, and network studies. It produces networks of study that include co-author and co-PI, co-citation

of documents, co-citations of journals, co-citations of authors and bibliographic coupling, and author bibliographic coupling. It is important to note that it also links, the bibliographic treatment of journals. The program allows one to also build first-order networks for document-author, document-reference, author reference, and journal reference. It accurately builds networks to illustrate a variety of data and conceptually representation of that data, holds a lot of flexibility, and has an adequate range of options for visualization, including i) Temporal mapping; ii) Geospatial mapping; iii) proportional symbol mapping; iv) choropleth map;vi) network mapsand v) topical maps. This application provides the user a lot of flexibility, but a sufficient range to accurately portray the equally complicated nature of science.

VOSviewer

VOSviewer is a software application for constructing and visualizing bibliometric networks based on co-citation, bibliographic coupling, or co-authorship ties; the actors of those networks could be either journals, researchers, or individual publications. In addition, the software can generate co-occurring networks of salient terms by applying text-mining techniques to a corpus of scholarly literature. This software was developed by the Centre for Science and Technology Studies (CWTS) at Leiden University in the Netherlands, and it can easily extract bibliometric networks (i.e. co-authorship, co-occurrence, and citation-based metrics) from data stored in various formats, including bibliographic bibliometric metadata such as those from WoS, Scopus, Dimensions, PubMed, RIS, etc. More importantly, if Data is imported from crossref, Europe PMC or MA APIs, and if a list of DOIs is provided, data can be extracted from the Semantic Scholar, OpenCitations and WikiData APIs to be used with VOSviever. VOSViewer forms a co-occurrence matrix in to three distinguishable steps to create the map,i) similarity matrix - based on the degree of connection - the VOS mapping methodology has considerable coverage in the literature (ii) use the VOS mapping methodology advantages over other methods which use the map itself and produce a map that will reflect the utility measure among the items e.g., lastly, mathematical are needed again for each of the specifications and (iii) translation, rotation, and reflection. In terms of visualizations capabilities, there are three options for visualizations available: network, overlay, and density. The zooming and scrolling capabilities of VOSviewer, and the way that labels do not overlap in a complicated way, is impressive. This allows you to review the map and data in both bitmap and vector formats.

This makes placing the analytical result in digital or printed form much easier. In addition, the created webpage can also be used to open VOSviewer directly. Lastly, it is possible to save any visualization with several different graphical file types.

Discussion and Conclusions

WoS and Scopus represent the most relevant accepted databases and the majority of existing studies are grounded on these two databases, nearly all software tools and libraries apply in bibliometric and scientometric studies is set up to use downloaded data from these two databases. In addition, it is also possible to import Dimensions data through Biblioshiny, CiteSpace, and VOSviewer. Using its API, CiteSpace and VOSviewer allow for direct data retrieval using MA. It is important to note that the only instruments that can be used in conjunction are instruments. Finally, it is worth underlining that many software products' compatibility with dissimilar data source data varies between the software available. Because of this, we want to emphasize the bifurcation in bibliometric. The totality of data may be handled using programs such as CiteSpace and VOSviewer. Ultimately, the researcher must select well in advance what software will cover each unique analysis. It should also be defined as stated before that each tool we have discussed has exogenous reasons to choose or dismiss software. Bibliometrix and their Shiny instrument probably have the most utilized it for an array of approaches. And since the interface is easy to work with, it is a great choice, all around we suppose. VOSviewer has great visualization features and the capability to download or import data from an array of sources. SciMAT appears to have the pre-processing and exporting features. The data analyst might be able to showcase a definitive study focus with sufficient visualization to illustrate the strategic diagram and thematic areas of study.

References

Biblioshiny. (n.d.). Bibliometrix. Retrieved August 29, 2024, from https://www.bibliometrix.org/home/index.php/layout/biblioshiny

Bibliotools / bibliomaps. (n.d.). Retrieved August 29, 2024, from https://www.sebastian-grauwin.com/bibliomaps/

Bornmann, L., &Leydesdorff, L. (2014). Scientometrics in a changing research landscape: Bibliometrics has become an integral part of research quality evaluation and has been

changing the practice of research. *EMBO Reports*, *15*(12), 1228–1232. https://doi.org/10.15252/embr.201439608

Citespace home. (n.d.). CiteSpace. Retrieved August 29, 2024, from https://citespace.podia.com/

CitNetExplorer—Analyzing citation patterns in scientific literature. (n.d.). CitNetExplorer. Retrieved August 29, 2024, from https://www.citnetexplorer.nl//

Cobo, M. J., López-Herrera, A. G., Herrera-Viedma, E., & Herrera, F. (2011). Science mapping software tools: Review, analysis, and cooperative study among tools. Journal of the American Society for Information Science and Technology, 62(7), 1382–1402. https://doi.org/10.1002/asi.21525

Cobo, M. J., López-Herrera, A. G., Herrera-Viedma, E., & Herrera, F. (2012). scimat: A new science mapping analysis software tool. Journal of the American Society for Information Science and Technology, 63(8), 1609–1630. https://doi.org/10.1002/asi.22688

Moral-Muñoz, J. A., Herrera-Viedma, E., Santisteban-Espejo, A., & Cobo, M. J. (2020). Software tools for conducting bibliometric analysis in science: An up-to-date review. El Profesional de La Información, 29(1). https://doi.org/10.3145/epi.2020.ene.03

Nikolić, D., Ivanović, D., & Ivanović, L. (2024). An open-source tool for merging data from multiple citation databases. Scientometrics, 129(7), 4573–4595. https://doi.org/10.1007/s11192-024-05076-2

Sci2 tool: A tool for science of science research and practice. (n.d.). Retrieved August 29, 2024, from https://sci2.cns.iu.edu/user/index.php

Van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, *84*(2), 523–538. https://doi.org/10.1007/s11192-009-0146-3

VOSviewer—Visualizing scientific landscapes. (n.d.). VOSviewer. Retrieved August 29, 2024, from https://www.vosviewer.com//

(N.d.). Retrieved August 29, 2024, from https://homepage.univie.ac.at/juan.gorraiz/bibexcel/

Indian Journal of Applied Economics and Finance
Vol. 1 No. 1 (January 2025): pp: 96-106