MethodsX 11 (2023) 102367

Contents lists available at ScienceDirect

MethodsX

journal homepage: www.elsevier.com/locate/methodsx

A method of Mapping Process for scientific production using the Smart Bibliometrics ☆

Vilker Zucolotto Pessin*, Celso Alberto Saibel Santos, Luciana Haure Yamane, Renato Ribeiro Siman, Roquemar de Lima Baldam, Valdemar Lacerda Júnior

Department of Informatics, Federal University of Espírito Santo, Av. Fernando Ferrari, 514, Vitória, Espírito Santo 29075-910, Brazil

ARTICLE INFO

Method name:
Scientific Mapping Process

Keywords:
Scientific method
Science mapping
Systematic literature reviews
Business intelligence
Making-decision
State-of-the-art
Scientific gaps

ABSTRACT

Big data launches a modern way of producing science and research around the world. Due to an explosion of data available in scientific databases, combined with recent advances in information technology, the researcher has at his disposal new methods and technologies that facilitate scientific development. Considering the challenges of producing science in a dynamic and complex scenario, the main objective of this article is to present a method aligned with tools recently developed to support scientific production, based on steps and technologies that will help researchers to materialize their objectives efficiently and effectively. Applying this method, the researcher can apply science mapping and bibliometric techniques with agility, taking advantage of an easy-to-use solution with cloud computing capabilities. From the application of the "Scientific Mapping Process", the researcher will be able to generate strategic information for a result-oriented scientific production, assertively going through the main steps of research and boosting scientific discovery in the most diverse fields of investigation.

- The Scientific Mapping Process provides a method and a system to boost scientific development.
- It automates Science Mapping and bibliometric analysis from scientific datasets.
- It facilitates the researcher's work, increasing the assertiveness in scientific production.

Specifications table

Subject area:	Computer Science
More specific subject area:	Computer systems, visualization, and graphics, data analysis
Name of your method:	Scientific Mapping Process
Name and reference of an original method:	Smart Bibliometrics: an integrated method of science mapping and bibliometric analysis.
Resource availability:	Supplementary material: https://data.mendeley.com/datasets/tbpd29zsg5/1 DOI:10.17632/tbpd29zsg5.1 Smart Bibliometrics system access: https://bit.ly/3GgYUZh

Method details

☆ **Direct Submission**: Pessin, V.Z., Yamane, L.H., Ribeiro, R.R., 2022. Smart Bibliometrics: an integrated method of science mapping and bibliometric analysis. Scientometrics. https://doi.org/10.1007/s11192-022-04406-6

* Corresponding author.
E-mail address: vilker.pessin@gmail.com (V.Z. Pessin).

https://doi.org/10.1016/j.mex.2023.102367
Received 27 February 2023; Accepted 5 September 2023
Available online 6 September 2023
2215-0161/© 2023 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

Introduction

Scientific discoveries take place in a dynamic, ongoing, complex environment, with significant changes, marked by a constant overcoming of research topics. These challenges require new skills and competencies, so researchers must apply a method and technology to achieve their research goals. Also, there is a considerable increase in scientific publications in various areas of knowledge, so researchers' choices need to be based on the analysis and interpretation of large amounts of data.

The "big data" context opens opportunities for the integration of new data analysis tools and methods to boost researchers' productivity with speed, efficiency, and accuracy in their scientific production, expanding the science horizons [10,42,51,58,72]. It is rare for researchers to start their studies without first applying bibliometric and scientometrics techniques to gain insights into the science landscape, identify knowledge gaps, and more accurately uncover the state-of-the-art in a particular research field [12,28,61,65,70].

Typically, the most well-known methods for scientific production recommend a preliminary analysis of the data available on official scientific portals, generating bibliometric information to guide the selection of theoretical repertoire [3,4,22,58,60,64].

The data available on scientific portals are valuable inputs for generating strategic information related to the global scientific production landscape, which opens space for the development of new methods that bring together a management process of activities and technology applications to drive scientific advancement [41]. Therefore, it is highly recommended that the researcher bases the scientific process on a consistent, managed method, starting from a series of well-directed, organized steps. A well-applied method has a high potential to guide scientific writing efficiently, at various steps of execution. It is also crucial for the researcher to choose a quality theoretical reference with good judgment from the beginning of their work, being aligned with the main discussions of the scientific community regarding a particular study object. In other words, it is highly advisable to choose a quality bibliographic repertoire from the beginning of the research, which increases the probability of publication success [4,22,33,55].

According to [3], Science Mapping (SM) and bibliometric analysis methods can be performed by combining different tools to analyze the evolution of the cognitive structure of a certain research topic, leading to the discovery of scientific boundaries. Despite being more consolidated in the field of Medicine/Health [44], the use of bibliometrics and SM is spreading to various areas, as mapping science is complex and difficult to manage, involving many steps and often requiring numerous bibliometric tools, not always free [6].

Over the years, these bibliometric tools have evolved along with information systems, either providing SM analyses with a focus on visualizations [6,34] or providing bibliometric analyses based on classification and statistical measures [10,18,20]. Today, the possibility of incorporating artificial intelligence (AI), with machine learning and natural language processing (NLP) resources into these processes is being discussed [42,58,71].

The first discussions on the mapping of scientific literature show the importance of methodological applications and systems that can perform bibliometric analyses facilitating the work of the researcher and promoting accuracy in scientific writing [1,11,52,66–68]. These methodologies and systems have evolved significantly since scientific databases began to emerge, opening opportunities for the application of technological solutions with the capacity to process massive amounts of data [2,30].

In the business and academic environment, research development must be based on data analysis, using methods and technologies that promote rationality to the scientific production process and that relate the knowledge boundaries under the prism of certain authors and journals around a certain issue, uncovering important nuances of science evolution through scientific data [6,15,19,22,31].

In this bibliographic review process, systematized research is a traditional procedure in the academic context to facilitate scientific writing [5,7,63], being recognized for following a transparent and replicable method of execution, so that it is possible to select the most relevant works and eliminate those that are least relevant for a specific search [54].

Systematic literature reviews (SLR) have significantly contributed to the advancement of knowledge in various fields by examining existing studies with attention to theoretical boundaries, units of analysis, data sources, study contexts and definitions, and operationalization of constructs, as well as research methods, with the aim of refining or revising an existing theory [25]. In other words, scientific production goes through a series of systematic steps and presupposes SM and bibliometric systems to assist the researcher in their decisions about which theoretical framework to use, giving rise to various methods [10,42,46,49,51,59,62,71,72].

It was from these opportunities in the field of scientometrics that [58] introduced "Smart Bibliometrics," a system developed on a Business Intelligence (BI) platform and freely accessible that aims to support the processes of SM (visualizations) and bibliometric (statistical measurements) analysis. As a differential, the "Smart Bibliometrics" includes cloud computing resources to embody the selection of the theoretical framework and it is constantly evolving to incorporate new analysis functions in tune with the concept of Open Science.

Therefore, the main objective of this paper is to present a method for scientific production, applying a systematic process, while providing the researcher with an information system that facilitates the attainment of their research objectives. It is expected that this method, called the "Scientific Mapping Process," brings accuracy to the researcher in their creative process of producing knowledge and contributes to the advancement of the science world wild.

Steps of the scientific mapping process

The scientific production method will follow a process, according to Fig. 1, predicting systematized steps to achieve the research objectives.

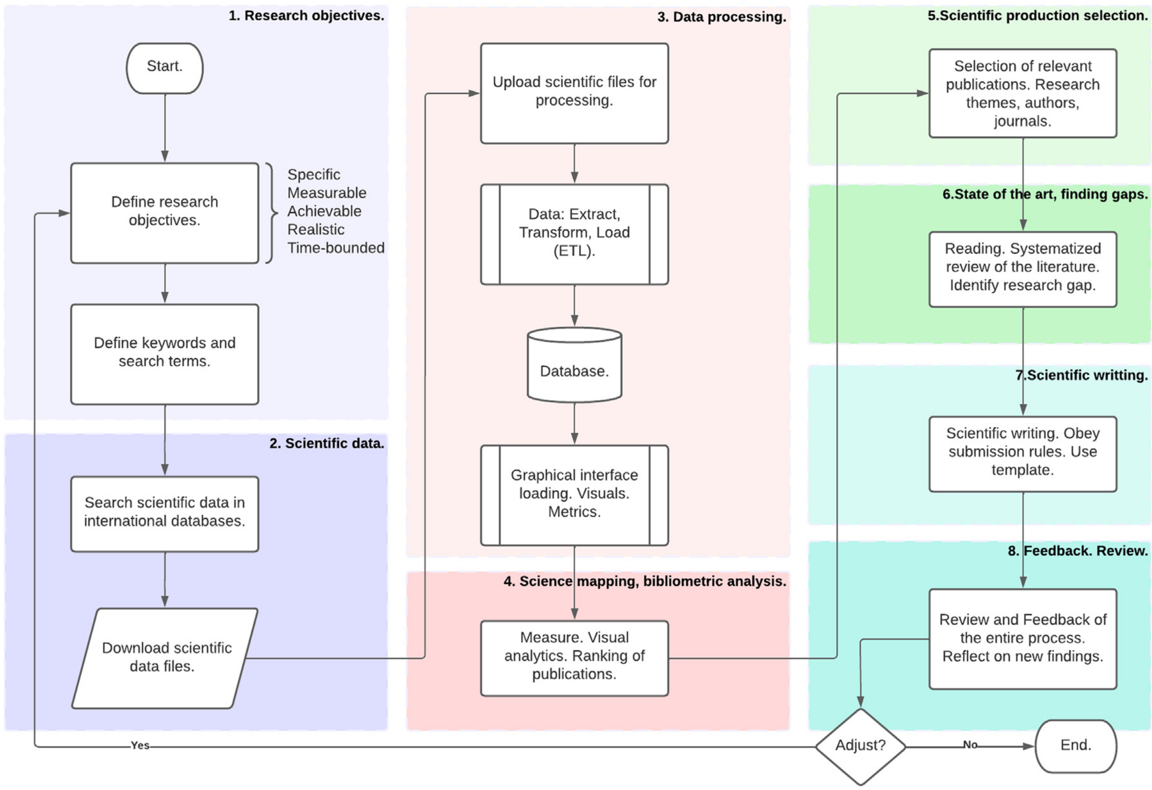

Fig. 1. Scientific mapping process method
Source: own authorship.

Step 1 – Research objectives

The first step for structuring research is to outline a problematization around a given object of study when the researcher has to define the objectives and scope of the research [22]. At this moment, the researcher defines the general context of the scientific problem in a general objective, that is, an aspiration of what is expected to materialize.

It is suggested that the researcher builds a mental map, outlining the overall research objective as a goal to be achieved, which should be broken down into secondary objectives and aligned with an execution methodology. Thus, the objectives guide the entire journey, requiring methods and protocols that are broken down into organized steps of investigation to successfully achieve the results of scientific research, such as the “Systematic Review Protocol” [38,39].

According to BJERKE; RENGER, (2017), for a goal to be achieved it is necessary to be Specific, Measurable, Attainable, Relevant, and Time-bound (SMART). It is a useful approach to establishing and achieving goals, with contextualized and effective objectives [53], because they provide a clear and defined goal to work for.

A SMART objective should be well-defined and specific, meaning it cannot be vague or too broad. The purpose of SMART must be very specific and definite, to increase the probability of achieving it [24]. Also, it should be possible to measure to track progress and advancement of steps, and yet, it should be broken down into specific goals and milestones to be achieved. To be an indicator, the specified goal measure needed to be a direct quantifiable variable of the outcome [53].

A SMART objective must be feasible and attainable, meaning it can be achieved, taking into consideration resources to achieve the goal. Additionally, the goals must be challenging. The goal must be achievable to be successful [24]. In addition, it must be relevant content that makes sense to move forward toward the goals to achieve objectives with realistic timeframes [8].

Finally, a SMART objective should have a timeline with step descriptions that mark deliveries, making it possible to determine execution time and create a sense of urgency and priority. The SMART objective must be time-limited, which means that a limiting term for the achievement of the objective should exist [24]. In general, SMART objectives set achievable goals since they provide a clear and defined goal to work for, and it is possible to materialize results.

Thus, from the definition of the objectives, the researcher will have an initial perception of the keywords related to the object of study. In short, the objectives of a bibliometric study must be related to a retrospective analysis of scientific discussions, which can reveal the bibliometric structure through the relationships among authors, journals, and educational institutions [22], at which point the researcher can bring their contribution.

So, in this preliminary step of setting goals, the researcher can choose representative search terms to locate the bibliographic repertoire related to the issue to be investigated. Therefore, it is important to define search terms (keywords) used with the original meaning of expressions in the international scope, giving special attention when the researcher's native language is not English, which is internationally adopted as the main language in the scientific community. With these search terms, the researcher can find the relevant scientific publications in international databases and proceed with an SLR in later steps.

Step 2 – Scientific data

After the research objectives and search terms (keywords) have been defined, it is time to search for discussions around the world about a particular study object. The researcher's first diligence at this point is to navigate international scientific portals and download scientific data [22].

The database query aims to extract available and accessible metadata files to generate a set of bibliometric information. There are various databases, some of which include Web of Science, Scopus, Crossref, Cochrane, and others, so the researcher can collect data from multiple sources [13,17,26,37].

As seen, in this step, the choice of representative keywords is a crucial step in finding related work directed to the research problem in focus. The combination of representative terms or keywords through a logical expression is the main query parameter in the search portals, as it will direct the results to the addressed topics. It is suggested that the search terms should be defined with equivalence to the technical terms in English, considering the research problem, area of application, and the objective of the academic work. In addition to keywords, other filter parameters can be applied to improve the results, such as date range, document types, and areas of knowledge, for example.

The search expression must align with technical terms used by scholars in the field to return global references on the topic. The query terms must be expressed in words that better represent the original meaning used in an international context, so the researchers need to be familiar with the jargon practiced worldwide for the study elements to investigate, considering also regionalities.

Regarding the types of documents, it is important to consider the importance of searching for different formats for different research objectives. Some databases still need new development to include records of reports or patents, in certain types of research. In databases, most records are scientific articles, books, and book chapters.

During the data extraction in the scientific portals, the researcher must generate data files that will be used as input for processing, to generate relevant information about the science landscape in each research area. After downloading the data, the researcher should analyze them, either by manually processing the scientific data or by using some system oriented to bibliometric and SM processes.

To use Smart Bibliometrics, after collecting the data in the scientific portals, the researcher must upload the data files to the platform for processing and database consolidation. After uploading data, the extraction, transformation, and loading (ETL) process will begin, also performing the exclusion of blank lines, duplicate records, and important corrections so that a standardized database is set and ready to generate information.

Step 3 – Data processing

The scientometric analysis is increasingly being applied to obtain quantitative insights in the development of research in specific domains of scientific investigation [43], which will require a system that provides interesting visuals and statistical metrics for comparing publications.

In the proposed method, we provide a cloud computing-based solution to automate the processing of scientific databases. If the researcher opts for a system in their analysis, they may use the Smart Bibliometrics system [58], which comes with a repository of analysis to support theoretical reference choices. The Smart Bibliometrics must be fed with data extracted from scientific databases to generate a set of information that brings insights to the researcher, providing SM and bibliometric analyses for increased efficiency in selecting a high-quality bibliographic repertoire.

Using Smart Bibliometrics, information is produced about the scientific production in the world, aggregating data visualizations and classification measures for the researcher to make good choices for the theoretical framework [58], based on the analysis of scientific data.

The cloud computing resources, processing capacity, and delivery of information with robust data analysis from the construction of intelligence for decision-making [23] are important and innovative aspects of the Smart Bibliometrics platform. Using the platform, the data will be automatically processed to generate strategic information that brings accuracy to the selection of a high-quality theoretical framework.

Upon completion of the data processing step, the researcher should interact with the system's functions to gain insights into the main publications. In these analyses, they will be able to determine relationships among authors, research problems, and journals interested in the discussions at hand and select a bibliographic repertoire to support scientific discussion.

Step 4 – Science mapping and bibliometrics analysis

Data visualization has become crucial in decision-making processes, as it can provide a much deeper understanding of a given problem, leading to a multi-faceted comprehension of any phenomenon. One of the key features of interactive data visualization is that users have direct control over how the information is presented during the decision-making process [47].

In the proposed method, with the support of Smart Bibliometrics, the researcher will be able to interact with the visuals in a logical sequence of information presented dynamically with the data, navigate through different analyses, apply data filters and slicers to obtain relevant insights and thus decide on the bibliographic repertoire to use in the research.

On the other hand, the classification metric is based on representative numerical variables to provide a measure that allows classification among publications, which include variables such as the number of citations, year of publication and impact factor of the journals, and other metrics to generate a representative index for classification. In Smart Bibliometrics, the classification metric is based on a citation rate per year, which improves the index by treating the publications more equitably by removing the bias of older publications that by their nature have a higher probability of citations [58].

Based on these interactive analyses, the researcher will have enough information to make smart choices about the "corpus" of the research [48], selecting the publications that directly address the research problem and that are indispensable references in the scientific theorization about the object of study.

Step 5 - Scientific production selection

This is one of the most crucial steps in the method when the researcher selects a theoretical framework to support discussion and build knowledge about the state-of-the-art of scientific knowledge related to the theme to explore.

Science mapping and bibliometric systems should offer solutions that allow statistical analysis or data visualization [6,19,27] to facilitate assertive choices by researchers. In this context, where we need to apply compelling and enlightening data visualizations as well as measure and quantify classification indicators, BI tools are widely used [58].

Smart Bibliometrics, developed on a BI platform, offers functionalities that allow for the selection of key scientific documents based on objective criteria, providing advanced and dynamic data visualization [58]. Thus, with a powerful data analysis tool, it will be intuitive to choose the main publications in the study area, significantly reducing the effort of the researcher in producing their theoretical discussion.

Step 6 – State-of-the-art and find gaps

Understanding the 'state-of-the-art ' of scientific knowledge in this step will be much more efficient if the researcher has completed the previous steps with assertiveness, selecting the main references that have dealt with discussions about the research problem, and providing a perspective on a certain object of study. The 'state-of-the-art ' refers to the starting point for innovation, in that it presents the updated knowledge of the latest scientific discussions to that point, bringing space for new contributions.

With the main publications related to a theme obtained after the execution of step 5, the researchers should then carry out a guided and systematic reading of each scientific document, at which point they will be able to grasp this 'state-of-the-art ' of scientific knowledge related to the theme. At this point, insights and discoveries may arise, supporting scientific writing aimed at solving questions that are still open.

Understanding the boundaries of scientific knowledge, the researcher will be able to make new contributions by identifying prominent research gaps for a focus on innovation. Individual contribution in a specific research gap opens space for the advancement of science in those areas that still lacking in depth and have a higher likelihood of publication by expanding the horizons of science.

Etapa 7 – Scientific writing

At this step, researchers have sufficient elements to materialize their work through scientific writing. It is time to put on paper their contribution, applying methodological rules of scientific writing.

Scientific writing requires a peer review process and the use of appropriate vocabulary [32]. It is also important to apply some rules and write scientifically [69].

In this line, scientific writing has its rite of development, which follows a series of steps and rules [36]. Therefore, it is recommended that the researcher should identify which publication vehicle will be sent his work to at the beginning of the production of the scientific text, as he/she must comply with the general rules of scientific writing and submission requirements applicable to the Editor. Thus, based on the SM and bibliometric analyses, the researcher will be successful in determining the target journal for publication and producing the content obeying the submission rules from the beginning [58].

Regarding the choice of scientific journal interested in the subject study, Smart Bibliometrics presents extensive visual resources to precisely determine the journals that are interested in the subject, providing strategic information to increase the chances of success in a publication [58].

At this step of scientific writing, the use of reference managers is also an important instrument to help the researcher's work. Several solutions help a lot in the management of scientific documents, automatically generating bibliographic references. The use of Mendeley, EndNote, or Zotero, for example, offers full management support allowing collection, organization, citation, and sharing of papers [14,16,50].

Step 8 – Feedback and review

One key attitude in the learning process is monitoring and feedback on previous steps. Along the way, learning is usually deepened by discoveries and insights that the researcher has obtained in previous steps. After an SLR, new keywords and technical terms may awaken new perspectives on other aspects, when it will be the opportune moment to feedback on the previous steps.

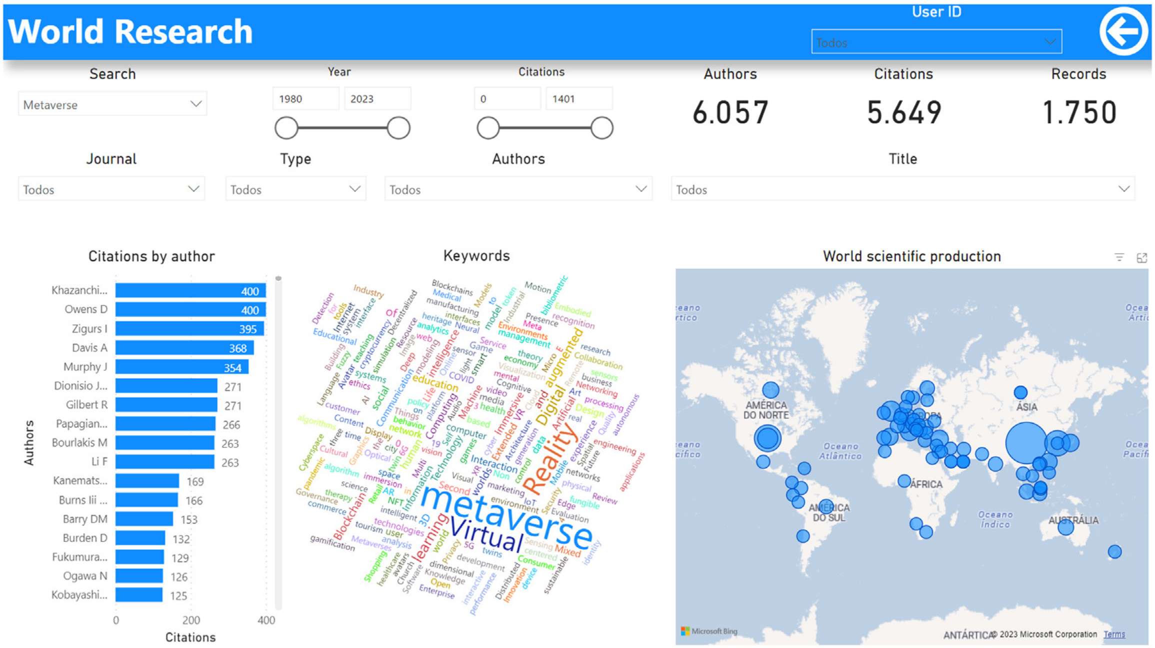

Fig. 2. Citations by author, word cloud, maps, and metrics
Source: adapted [58].

At this moment, based on the knowledge already acquired, the researcher remakes the management process, improving the objectives and looking for new publications that can add value to the work, perfecting the details. It is a cyclic process, of feedback, seeking to solidify the theoretical construction in such a way that new inputs are incorporated. Feedback is an important step in creative processes, opening space for changing prototypes, either to go deep or to stimulate refinement [35].

Feedback is critical to delivering a quality product, reflecting on the relationship between planning and product quality to be delivered [45]. Thus, in this step, it is suggested that the researcher return to the beginning of the process, review its objectives, define new search terms, and then continue with a refinement of its scientific production for the delivery of a quality product that promotes innovation in each research area.

Application for an emerging theme

The SM methodologies have evolved in recent years due to advances in information technology, following the evolution of scientific databases and the processing capability of systems [2,30]. Thus, new concepts such as AI and NLP find strong applicability in the academic context to produce automated analysis [40,56,58].

Specifically, regarding NLP to process large volumes of data, the results of the system made available to suggest the automation of a systematic review using NLP were compared, described in the article by [42], but with another analysis bias, thus allowing the selection of bibliographic repertoire dynamically.

According to Kwabena et al. [42], it is necessary to integrate several sets of data into a single database. In Smart Bibliometrics [58], there are several pre-established resources and analyses, including integrating various data into a consolidated database after uploading the scientific data. The Extraction, Transformation, and Loading (ETL) process of the data takes place automatically, integrating various data sets into a database [42,58]. The system is easy to update with cloud computing resources, allowing free access from any device with an internet connection and a web browser.

Conducting a brief analysis using the Smart Bibliometrics system [58], for example on an emerging topic "Metaverse," following the processing results from data extracted on November 2022. Fig. 2 shows the word cloud processing for bibliographic repertoire selection, with 1.750 records. Through interaction with the various visuals, the researcher will have sufficient resources to select a high-quality bibliographic repertoire.

The "Smart Graphs" function was developed to provide a theoretical reference download option that aligns with the research theme, detecting the main publications on the study topic in various specialized journals. In this dynamic report shown in Fig. 3, it is possible to view the main information from the latest publications to decide whether to use them as a reference. The report was designed to create a link among authors, journals, research topics, and scientific productions.

The simulation for these analyses on the "Metaverse" theme can be accessed by: http://bit.ly/3Jbtfgs. There are several analyses on this link, two analyzes will be highlighted as examples.

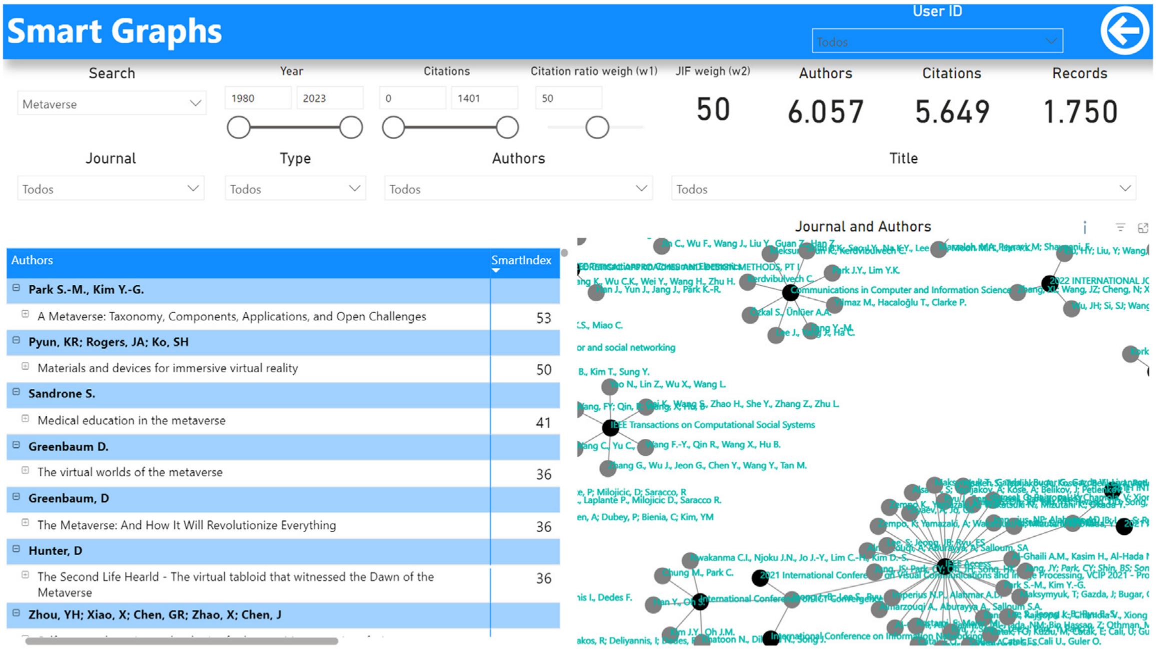

Fig. 3. Matrix and graph visualization mapping authors and journals' relationships
Source: adapted [58].

Two main visualizations are available: a matrix, with information about scientific works, ordered based on the ranking index and classification for each publication. There is a classification metric (Smart index) for ranking, which ranges from 0 to 100, helping to identify the most relevant publications.

The matrix is dynamic and allows expanding the data fields by clicking on the [+ icon], successively arranging hierarchically the author, title, summary, DOI, and document download link. This data matrix brings innovation by facilitating an analysis of the main bibliometric elements, making navigation very practical for the user, who can more quickly resolve the importance of the reference in the context of their research. The visuals are dynamic and will be updated to the filters that are applied, making a sequential analysis based on documents that meet the informed parameters. A cut in the year of publication, theme, and type of document will refine the result for agile and accurate analysis.

Thus, the SLR can be facilitated by this visual in which the main information of the articles that will make up the theoretical reference is analyzed. A preliminary reading of the summary helps in deciding whether to download the paper for use as a theoretical reference while updating the state-of-the-art knowledge. So, when navigating the matrix hierarchy and finding that the scientific production is relevant to the research to be done, simply copy the available link and paste it into the web browser. You will be automatically directed to the download, which will be made available by the publication source.

The second graph visual "Journals and Authors" shows the links among authors' groups and the main journals. It is formed by vertices and edges making the interconnection of journals and authors for each research area. The visual has a zoom feature to increase the capacity of the analysis and user interaction.

Each node represents a journal, and when selected, it will segment data in all visuals simultaneously. Similarly, by selecting only one of the authors, the work will be detailed with the information of the respective matrix visual. The systematic study of each author's contributions significantly increases the probability of identifying a relevant research gap, facilitated by a directed and focused reading, presenting the main publications in a scaled manner based on an assertive classification metric.

The rationale of the work, novelty, and further insights

A method to assist scientific production is essential to guide and structure the research process, ensuring efficiency, quality, and significant results [10,51,63]. This method, being rooted in innovation, seamlessly integrates the finest practices previously established in various approaches. By consolidating these practices into an agile framework, it places utmost importance on stages that enhance efficiency and effectiveness in the scientific production process.

This method breaks down scientific production into well-defined stages and provides a systematic and organized approach for researchers, describing a process to achieve scientific outcomes. Thus, the "Scientific Mapping Process" method innovates by complementarily providing a Business Intelligence system capable of extracting strategic information for assertive and results-oriented action.

By dividing scientific production into eight stages, it is possible to gain insights for researchers to visualize the overall research landscape and have a clearer and more directive understanding of the necessary steps to materialize scientific results. Within this perspective, the initial stage of this process encompasses the establishment of SMART objectives [8,53]. It serves as a crucial foundation for delineating the research scope and setting forth goals that are both attainable and practical. Positioned at the outset of the process, this stage plays a fundamental role in guiding researchers by ensuring that their efforts are rooted in objectives that are specific, measurable, attainable, relevant, and time-bound.

The Smart Bibliometrics system provides analyses of scientific data as a central and integrated stage throughout the research process, enabling researchers to gain insights into the global landscape of scientific research. This differentiated approach brings a new perspective by including the analysis of scientific data as a key factor for research success, helping researchers achieve their goals more efficiently.

Thus, by previously analyzing official science data, researchers can base their research on reliable and up-to-date information, applying a system that was designed for this purpose from the beginning of the process. In this context, Smart Bibliometrics plays a fundamental role, integrating the Scopus [26] and Web of Science [13] databases, and opens opportunities for future insights by integrating other databases, such as Lens [9,57], Dimensions [21], Cochrane [37], for example. Through this system, researchers have access to a wide range of scientific data, allowing for comprehensive and evidence-based analysis.

Bibliometric and science mapping analysis find an organized process in this method to bring agility and effectiveness, as it assists researchers in identifying trends, patterns, and connections between different research groups. This analysis provides valuable insights and contributes to the development of more impactful and innovative studies, indicating potential specialized journals interested in various subjects, and strategic information to leverage success in publications.

In this method, the selection of a bibliographic repertoire is a crucial step for the researcher, as it involves choosing relevant references that underpin the research from the beginning and identifying key authors and publications in the field of study. With the support of Smart Bibliometrics, researchers have access to a wide range of information about scientific articles, their citations, and the impact they have had on defining the pillars of knowledge in a particular area.

Following this line of reasoning, the identification of research gaps becomes facilitated, driving discoveries in areas that have not yet been explored or present opportunities for new studies. Thus, this preliminary analysis of scientific data offers insights into less-explored areas or areas that require further development.

With this roadmap in mind, materializing a rational praxis, in this stage, the researcher is in a better position to proceed with scientific writing to communicate the research results. These results will be enhanced through a final review process, a necessary step to ensure the quality and consistency of the work.

In summary, the proposed method of breaking down scientific production into organized and systematized stages, with the central element being the analysis of scientific data, brings innovation to the research process and boosts researcher productivity. Smart Bibliometrics acts as a supporting tool, providing strategic information from bibliometric and science mapping analyses, putting the researcher on the right path to advance the horizons of science.

Limitations

Smart Bibliometrics has undergone several improvements since its launch, but there is a vast horizon for new developments. It is a trend that new computing features are included, such as the development of AI and NLP functions, as well as the integration of other databases.

Another limitation, in the context of research, is the absence of a complete and consistent database, with standardized data and more descriptive variables to expand the power of analysis and automation. There is space for a true evolution of these databases, integrating scientific production globally, and gathering the most varied subjects of knowledge in a single and validated database by the scientific community.

It is also important to consider some pragmatic limitations of the method, acknowledging that these approaches may entail a partial view and a simplified representation of the complexity of scientific production, depending on a series of factors. For example, user interpretations and procedures used for system updates, limitations of the databases, and the availability of accurate and comprehensive data on scientific production can generate practical limitations.

Under this perspective, limitations may arise from the interaction and interpretation of users with these databases to extract data samples for analysis. The definition of terms and keywords, applied filters, journals, and areas of study, among other factors, can bias the results. Properly defining the terms and keywords used in science mapping analyses is essential for obtaining accurate results, considering the international jargon used and the syntax of the terms. Different authors and researchers may use similar terms with regional expressions and different interpretations, which can lead to ambiguities and imprecise results. Inadequate keyword selection can result in the exclusion of relevant studies or the inclusion of irrelevant studies in the mappings.

The process of interpreting the results itself may contain pragmatic limitations, as science mapping analyses provide visualizations and metrics that can be interpreted differently by researchers. Interpreting the results requires knowledge and understanding of the context of the research area, as well as the specific purpose of the analysis in line with the research objectives. An inadequate understanding of the results can lead to different conclusions depending on the experience and expertise of each researcher.

It is worth noting that Scopus and Web of Science themselves may provide an insufficient quantity of publications within the scope of certain areas of study, such as engineering, among others. However, in the case of medical fields, it would be necessary to complement them with data from other specialized databases such as PubMed and Cochrane or others [29]. It is also important to highlight that these databases may have their limitations, such as limited coverage of specific journals, conferences, or thematic areas. Additionally, the information available in the databases may be inconsistent or contain errors, which can affect the accuracy of the analyses.

Another point to highlight is that not all data may be accessible or easily obtained, especially when it comes to ongoing research, non-indexed publications, or unstructured data. This can lead to an incomplete or biased representation of the field of study.

The availability of this data, combined with the constant changes occurring in the field of science, reveals that the research environment is dynamic, with discoveries, approaches, and emerging areas of study continuously emerging. This, too, can be seen as a pragmatic limitation. Therefore, science mapping analyses may face difficulties in capturing the dynamics and rapid changes that occur in the scientific field, considering the time it takes for studies to develop. As a result, the results can quickly become outdated or may not fully reflect the current state of research, another consequence of “big data.”

Finally, it is important to highlight another significant limitation regarding the concept of open science. With the advancement of information technologies, we are evolving toward the global connection of researchers and knowledge, allowing for the borderless exchange of science.

An open science concept is being worked on. The horizons are still wide open to promote an inclusive science, allowing for unrestricted data and tool availability that integrates researchers around the world for collective and integrated action, promoting synergies and driving scientific discovery.

Conclusions

Staying updated has become a challenge for any knowledge professional, especially in the innovation market and academic environment where the constant surpassing of research topics in a complex discussion environment is observed. Producing relevant scientific material requires a lot of dedication and time, and the systematic methods of SM and bibliometric are valuable to optimize scientific production with a focus on scientific results and publication. As a result of this movement, innovative technologies, and methods arise to support these complex processes, opening fields for new developments to overcome future challenges.

In the business environment, data analysis to support decision-making has been widely applied, as well as in the scientific environment where research is complex and requires total alignment with a well-defined problem. A solid theoretical framework is what provides support for a solid investigation, worthy of defense before the scientific community. The researcher in the knowledge era needs to act intelligently, selecting up-to-date and high-quality bibliographic repertoire to keep up with constant scientific publications.

Specialized journals have become increasingly demanding, so scientific production requires a consistent method based on consistent data analysis. Having a broad overview of the state-of-the-art of any topic, at the beginning of the research, is a predominant factor that may affect the decision of a publication to occur.

The scientific contributions of these authors, grouped by topic and target journal, in addition to revealing the discussions of the scientific community, have a high potential to bring clarity to the expectations of the various publication vehicles about their research field. It provides a logical sequence of the importance of articles based on assertive metrics that investigate the relationship among journals, authors, and research areas. Such analysis can direct the researcher’s work with great assertiveness, converting intelligence into results-oriented writing with scientific innovation and publication.

Therefore, the “Scientific Mapping Process” method presented has the potential to facilitate the researcher’s work, increasing the assertiveness in scientific production, and in a complementary way, contributing to the expansion of knowledge horizons, being the main contribution of this paper.

Ethics statements

Ethical considerations have been considered throughout the research process. The authors acknowledge their responsibility to adhere to the highest ethical standards in conducting and reporting their research.

Supplementary material and additional information

This solution is constantly evolving, and future developments will be available through this repository with supplementary material: https://data.mendeley.com/datasets/tbpd29zsg5/1. DOI:10.17632/tbpd29zsg5.1.

Smart Bibliometrics System access through this link: https://bit.ly/3GgYUZh.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Vilker Zucolotto Pessin: Conceptualization, Methodology, Software, Resources, Data curation, Writing – original draft. **Celso Alberto Saibel Santos:** Validation, Supervision, Writing – review & editing. **Luciana Haure Yamane:** Validation, Supervision, Writing – review & editing. **Renato Ribeiro Siman:** Validation, Supervision, Writing – review & editing. **Roquemar de Lima Baldam:** Validation, Supervision, Writing – review & editing. **Valdemar Lacerda Júnior:** Validation, Supervision, Writing – review & editing.

Data availability

Data will be made available on request.

Acknowledgments

We would like to express our sincere appreciation to the editors, the dedicated editorial team, and the anonymous reviewers for their valuable contributions during the review process of our manuscript. Their expertise, insightful comments, and constructive feedback have significantly enhanced the quality and impact of our research. We are grateful for their time, effort, and dedication in evaluating our work and for their instrumental role in shaping the final version of our paper.

References

[1] P. Ahlgren, B. Jarneving, Bibliographic coupling, common abstract stems and clustering: a comparison of two document-document similarity approaches in the context of science mapping, Scientometrics 76 (2008) 273–290, doi:10.1007/s11192-007-1935-1.
[2] D. Akdur, V. Garousi, O. Demirörs, A survey on modeling and model-driven engineering practices in the embedded software industry, J. Syst. Archit. 91 (2018) 62–82, doi:10.1016/j.sysarc.2018.09.007.
[3] L. Alcaide–Muñoz, M.P. Rodríguez–Bolívar, M.J. Cobo, E. Herrera–Viedma, Analysing the scientific evolution of e-Government using a science mapping approach, Gov. Inf. Q. 34 (2017) 545–555, doi:10.1016/j.giq.2017.05.002.
[4] L. Alcaide–Muñoz, M.P. Rodríguez–Bolívar, M.J. Cobo, E. Herrera–Viedma, Analysing the scientific evolution of e-Government using a science mapping approach, Gov. Inf. Q. 34 (2017) 545–555, doi:10.1016/j.giq.2017.05.002.
[5] E. Alemdag, K. Cagiltay, A systematic review of eye tracking research on multimedia learning, Comput. Educ. 125 (2018) 413–428, doi:10.1016/j.compedu.2018.06.023.
[6] M. Aria, C. Cuccurullo, Bibliometrix: an R-tool for comprehensive science mapping analysis, J. Informetr. 11 (2017) 959–975, doi:10.1016/j.joi.2017.08.007.
[7] O. Bak, S. Shaw, C. Colicchia, V. Kumar, A systematic literature review of supply chain resilience in small– medium enterprises (SMEs): a call for further research, IEEE Trans. Eng. Manag. 70 (2020) 328–341, doi:10.1109/TEM.2020.3016988.
[8] M.B. Bjerke, R. Renger, Being smart about writing SMART objectives, Eval. Program Plan. 61 (2017) 125–127, doi:10.1016/j.evalprogplan.2016.12.009.
[9] Cambia, Explore global science and technology knowledge, Lens. Org. (2023). https://www.lens.org/ (accessed July 4, 2023).
[10] A. Carrera-Rivera, W. Ochoa, F. Larrinaga, G. Lasa, How-to conduct a systematic literature review: a quick guide for computer science research, MethodsX 9 (2022) 101895, doi:10.1016/j.mex.2022.101895.
[11] D. Chavalarias, J.P. Cointet, Bottom-up scientific field detection for dynamical and hierarchical science mapping, methodology and case study, Scientometrics 75 (2008) 37–50, doi:10.1007/s11192-007-1825-6.
[12] H. Chen, R. H.L.Chiang, V. C Storey, Business intelligence and analytics: from big data to big impact, MIS Q. 36 (2018) 1165–1188.
[13] Clarivate, Register for your free Web of Science profile, Web of Science TM. (2023). https://www.webofscience.com/wos/woscc/basic-search (accessed January 23, 2023).
[14] Clarivate, Accelerate your research, EndNote TM. (2023). https://endnote.com/ (accessed January 23, 2023).
[15] M.J. Cobo, F. Chiclana, A. Collop, J. De Ona, E. Herrera-Viedma, A bibliometric analysis of the intelligent transportation systems research based on science mapping, IEEE Trans. Intell. Transp. Syst. 15 (2014) 901–908, doi:10.1109/TITS.2013.2284756.
[16] Corporation for Digital Scholarship, Zotero, (2023). https://www.zotero.org/ (accessed January 23, 2023).
[17] Crossref, You are Crossref, Crossref. (2023). https://crossref.org/ (accessed January 23, 2023).
[18] E.A.R. de Campos, R.N. Pagani, L.M. Resende, J. Pontes, Construction and qualitative assessment of a bibliographic portfolio using the methodology Methodi Ordinatio, Scientometrics 116 (2018) 815–842, doi:10.1007/s11192-018-2798-3.
[19] G.D.G. De Carvalho, C.C. Sokulski, W.V. Da Silva, H.G. De Carvalho, R.V. De Moura, A.C. De Francisco, C.P. Da Veiga, Bibliometrics and systematic reviews: a comparison between the Proknow-C and the Methodi Ordinatio, J. Informetr. 14 (2020), doi:10.1016/j.joi.2020.101043.
[20] H. Dervis, Bibliometric analysis using bibliometrix an R package, J. Scientometr. Res. 8 (2019) 156–160, doi:10.5530/JSCIRES.8.3.32.
[21] Digital Science & Research Solutions Inc., Linked research data from idea to impact, Dimensions. (2023). https://www.dimensions.ai/ (accessed September 10, 2023).
[22] N. Donthu, S. Kumar, D. Mukherjee, N. Pandey, W.M. Lim, How to conduct a bibliometric analysis: an overview and guidelines, J. Bus. Res. 133 (2021) 285–296, doi:10.1016/j.jbusres.2021.04.070.
[23] Y. Duan, J.S. Edwards, Y.K. Dwivedi, Artificial intelligence for decision making in the era of Big Data – evolution, challenges and research agenda, Int. J. Inf. Manag. 48 (2019) 63–71, doi:10.1016/j.ijinfomgt.2019.01.021.
[24] G. Duca, Design of the smart objectives system in the management of a research project, Adv. Intell. Syst. Comput. (2021), doi:10.1007/978-3-662-47241-5.
[25] C.F. Durach, J. Kembro, A. Wieland, A new paradigm for systematic literature reviews in supply chain management, J. Supply Chain Manag. 53 (2017) 67–85, doi:10.1111/jscm.12145.
[26] Elsevier B.V, Welcome to Scopus Preview, Scopus Preview. (2023). https://www.scopus.com/home.uri (accessed January 23, 2023).
[27] Encyclopedia of Library and Information Sciences, Bibliometrics: tools and software, Post Campus Library. (2022). https://liu.cwp.libguides.com/c.php?g=225325&p=4966525 (accessed August 21, 2023).
[28] M. Favaretto, E. De Clercq, B.S. Elger, Big Data and discrimination: perils, promises and solutions. A systematic review, J. Big Data 6 (2019), doi:10.1186/s40537-019-0177-4.
[29] T.F. Frandsen, M.B. Eriksen, D.M.G. Hammer, J.B. Christensen, J.A. Wallin, Using Embase as a supplement to PubMed in Cochrane reviews differed across fields, J. Clin. Epidemiol. 133 (2021) 24–31, doi:10.1016/j.jclinepi.2020.12.022.
[30] S. Frezza, M. Daniels, A. Wilkin, Assessing students' IT professional values in a global project setting, ACM Trans. Comput. Educ. 19 (2019) 1–34, doi:10.1145/3231710.
[31] M.C.B. Galvão, I.L.M. Ricarte, Revisão sistemática da literatura: conceituação, produção e publicação, Filosofia da informação 6 (2019) 57–73, doi:10.21728/logeion.2019v6n1.
[32] W.H. Guilford, Teaching peer review and the process of scientific writing, Adv. Physiol. Educ. 25 (2001) 167–175, doi:10.1152/advances.2001.25.3.167.

[33] P. Hallinger, Science mapping the knowledge base on educational leadership and management from the emerging regions of Asia, Africa and Latin America, 1965–2018, Educ. Manag. Adm. Leadersh. 48 (2020) 209–230, doi:10.1177/1741143218822772.
[34] P. Hallinger, J. Kovačević, A bibliometric review of research on educational administration: science mapping the literature, 1960 to 2018, Rev. Educ. Res. 89 (2019) 335–369, doi:10.3102/0034654319830380.
[35] S.H. Harrison, E.D. Rouse, An inductive study of feedback interactions over the course of creative projects, Acad. Manag. J. 58 (2015) 375–404, doi:10.5465/amj.2012.0737.
[36] S. Hotaling, Simple rules for concise scientific writing, Limnol. Oceanogr. Lett. 5 (2020) 379–383, doi:10.1002/lol2.10165.
[37] Inc. John Wiley & Sons, Cochrane library, Wiley. (2023). https://www.cochranelibrary.com/ (accessed July 06, 2023).
[38] B. Kitchenham, D. Budgen, P. Brereton, M. Turner, S. Charters, S. Linkman, Large-scale software engineering questions - Expert opinion or empirical evidence? IET Softw. 1 (2007) 161–171, doi:10.1049/iet-sen:20060052.
[39] B. Kitchenham, S. Charters, Guidelines for performing Systematic Literature Reviews in Software Engineering, EBSE Technical Report EBSE-2007-01, Keele, Durham, 2007.
[40] S. Kiv, S. Heng, Y. Wautelet, S. Poelmans, M. Kolp, Using an ontology for systematic practice adoption in agile methods: expert system and practitioners-based validation, Expert Syst. Appl. 195 (2022) 116520, doi:10.1016/j.eswa.2022.116520.
[41] T. Kolajo, O. Daramola, A. Adebiyi, Big data stream analysis: a systematic literature review, J. Big Data 6 (2019), doi:10.1186/s40537-019-0210-7.
[42] A.E. Kwabena, O.B. Wiafe, B.D. John, A. Bernard, F.A.F. Boateng, An automated method for developing search strategies for systematic review using Natural Language Processing (NLP), MethodsX 10 (2023), doi:10.1016/j.mex.2022.101935.
[43] J. Li, F. Goerlandt, G. Reniers, An overview of scientometric mapping for the safety science community: methods, tools, and framework, Saf. Sci. 134 (2021) 105093, doi:10.1016/j.ssci.2020.105093.
[44] C. Liang, S. Qiao, B. Olatosi, T. Lyu, X. Li, Emergence and evolution of big data science in HIV research: bibliometric analysis of federally sponsored studies 2000–2019, Int. J. Med. Inform. 154 (2021) 104558, doi:10.1016/j.ijmedinf.2021.104558.
[45] J.W. Liu, C.Y. Ho, J.Y.T. Chang, J.C.A. Tsai, The role of sprint planning and feedback in game development projects: implications for game quality, J. Syst. Softw. 154 (2019) 79–91, doi:10.1016/j.jss.2019.04.057.
[46] K. Lo, M. Stephenson, C. Lockwood, Extending the hierarchical decision matrix to incorporate a dominance ranking score for economic systematic reviews, MethodsX 7 (2020) 101047, doi:10.1016/j.mex.2020.101047.
[47] W. Luo, User choice of interactive data visualization format: the effects of cognitive style and spatial ability, Decis. Support Syst. 122 (2019) 113061, doi:10.1016/j.dss.2019.05.001.
[48] M.W. Bauer, G. Gaskell, Qualitative Researching with Text, Image and Sound: a Practical Handbook, 1st ed., Sage Publications Ltd, London, 2000.
[49] K. Mdingi, S.Y. Ho, Literature review on income inequality and economic growth, MethodsX 8 (2021) 101402, doi:10.1016/j.mex.2021.101402.
[50] Mendeley Ltd, Mendeley brings your research to life, so you can make an impact on tomorrow, Mendeley. (2023). https://www.mendeley.com/?interaction_required=true (accessed January 23, 2023).
[51] W. Mengist, T. Soromessa, G. Legese, Method for conducting systematic literature review and meta-analysis for environmental science research, MethodsX 7 (2020) 100777, doi:10.1016/j.mex.2019.100777.
[52] E.C. Noyons, H. Moed, A.F. van Raan, Integrating research performance analysis and science mapping, Scientometrics 46 (1999) 591–604, doi:10.1016/S0138-9130(00)86853-7.
[53] O. Ogbeiwi, Why written objectives need to be really SMART, Br. J. Healthc. Manag. 23 (2017) 324–336, doi:10.12968/bjhc.2017.23.7.324.
[54] R.N. Pagani, J.L. Kovaleski, L.M. Resende, Methodi Ordinatio: a proposed methodology to select and rank relevant scientific papers encompassing the impact factor, number of citation, and year of publication, Scientometrics 105 (2015) 2109–2135, doi:10.1007/s11192-015-1744-x.
[55] F. Pallottino, M. Biocca, P. Nardi, S. Figorilli, P. Menesatti, C. Costa, Science mapping approach to analyze the research evolution on precision agriculture: world, EU and Italian situation, Precis. Agric. 19 (2018) 1011–1026, doi:10.1007/s11119-018-9569-2.
[56] R. Papa, T. Gioia, S. Alseekh, E. Frascaroli, C. Costa, S. Carpentier, U. Schurr, F. Loreto, P. Menesatti, Plant phenotyping research trends, a science mapping approach, Front. Plant Sci. (2019) 1 www.frontiersin.org, doi:10.3389/fpls.2018.01933.
[57] R. Penfold, Using the Lens database for staff publications, J. Med. Libr. Assoc. 108 (2020), doi:10.5195/jmla.2020.918.
[58] V.Z. Pessin, L.H. Yamane, R.R. Ribeiro, Smart bibliometrics : an integrated method of science mapping and bibliometric analysis, Scientometrics 22 (2022), doi:10.1007/s11192-022-04406-6.
[59] V. Quan-Hoang, L. Anh-Vinh, L. Viet-Phuong, H. Phuong-Hanh, H. Manh-Toan, Making social sciences more scientific: literature review by structured data, MethodsX 7 (2020) 100818, doi:10.1016/j.mex.2020.100818.
[60] M.P. Rodríguez-Bolívar, L. Alcaide-Muñoz, M.J. Cobo, Analyzing the scientific evolution and impact of e-Participation research in JCR journals using science mapping, Int. J. Inf. Manag. 40 (2018) 111–119, doi:10.1016/j.ijinfomgt.2017.12.011.
[61] S. Sahoo, Big data analytics in manufacturing: a bibliometric analysis of research in the field of business management, Int. J. Prod. Res. 0 (2021) 1–29, doi:10.1080/00207543.2021.1919333.
[62] S. Scheuer, J. Jache, L. Sumfleth, T. Wellmann, D. Haase, Creating accessible evidence bases: opportunities through the integration of interactive tools into literature review synthesis, MethodsX 8 (2021) 101558, doi:10.1016/j.mex.2021.101558.
[63] F. Schreiber, C. Cramer, Towards a conceptual systematic review: proposing a methodological framework, Educ. Rev. 0 (2022) 1–22, doi:10.1080/00131911.2022.2116561.
[64] P.B. Seddon, D. Constantinidis, T. Tamm, H. Dod, How does business analytics contribute to business value? Inf. Syst. J. 27 (2017) 237–269, doi:10.1111/isj.12101.
[65] J. Sheng, J. Amankwah-Amoah, X. Wang, Technology in the 21st century: new challenges and opportunities, Technol. Forecast. Soc. Chang. 143 (2019) 321–335, doi:10.1016/j.techfore.2018.06.009.
[66] H. Small, Update on science mapping: creating large document spaces, Scientometrics 38 (1997) 275–293, doi:10.1007/BF02457414.
[67] S. Soós, The Functional Anatomy of Science mapping: Katy Börner: Atlas of science: Visualizing what We Know, The MIT Press, Cambridge, MA/London, UK, 2011 2010, US$20. Scientometrics 89, 723–726, doi:10.1007/s11192-011-0480-0.
[68] S. Soós, G. Kampis, Beyond the basemap of science: mapping multiple structures in research portfolios: evidence from Hungary, Scientometrics 93 (2012) 869–891, doi:10.1007/s11192-012-0713-x.
[69] B.E.C. Timmermana, D.C. Strickland, R.L. Johnson, J.R. Paynec, Development of a "universal" rubric for assessing undergraduates' scientific reasoning skills using scientific writing, Assess. Eval. High. Educ. 36 (2011) 509–547, doi:10.1080/02602930903540991.
[70] V.H. Trieu, Getting value from Business Intelligence systems: a review and research agenda, Decis. Support Syst. 93 (2017) 111–124, doi:10.1016/j.dss.2016.09.019.
[71] T. Weißer, T. Saßmannshausen, D. Ohrndorf, P. Burggräf, J. Wagner, A clustering approach for topic filtering within systematic literature reviews, MethodsX 7 (2020), doi:10.1016/j.mex.2020.100831.
[72] K.K. Westphaln, W. Regoeczi, M. Masotya, B. Vazquez-Westphaln, K. Lounsbury, L. McDavid, H.N. Lee, J. Johnson, S.D. Ronis, From Arksey and O'Malley and beyond: customizations to enhance a team-based, mixed approach to scoping review methodology, MethodsX 8 (2021) 101375, doi:10.1016/j.mex.2021.101375.